

ASTR 792: The Interstellar Medium

Elisabeth A.C. Mills

October 31, 2023

Contents

1 Glossary of Fundamental Concepts	7
1.1 Number density – n	7
1.2 Mean Molecular Weight – μ	8
1.3 Collisional Cross section – σ	8
1.4 Mean free Path – τ	9
1.5 Two-body Collisional Rate Coefficient – $\langle \sigma v \rangle$	10
1.6 Maxwell-Boltzmann Statistics and Kinetic Temperature – T_{kin}	11
1.7 Internal Energy States	11
1.8 Boltzmann Distribution and Excitation Temperature – T_{ex}	12
1.9 Partition Functions – Z	12
1.10 LTE: Local Thermodynamic Equilibrium	14
1.11 An Aside on the Concept of an Ideal Gas	14
2 Radiative Transfer	17
2.1 Definition of Standard Quantities	17
2.1.1 Energy	17
2.1.2 Intensity	18
2.1.3 Flux	19
2.2 The Equation of Radiative Transfer	19
2.3 Brightness Temperature	22
3 Continuum Emission Mechanisms	25
3.1 Thermal Plasma	25
3.1.1 Free-Free Emissivity and Absorptivity	25
3.1.2 Emission Measure	26
3.2 Continuum Radiation	27
3.2.1 Synchrotron Radiation	28
3.2.2 Free-Free Radiation	30
3.2.3 Dust Emission	32
4 Atomic Structure	35
4.1 Electron Orbital Structure	35
4.1.1 Fermi Statistics	35
4.1.2 Energy States	35
4.1.3 Quantum Numbers	36
4.2 Spectroscopic Terms	37

4.3	Fine Structure (Spin-Orbit Coupling)	38
4.4	Parity, Degeneracy, and Multiplicity (oh my)	39
4.5	Electric Dipole Selection Rules and Forbidden Transitions	40
4.6	Hyperfine Structure	42
4.7	Applications	42
5	Molecular Structure	45
5.1	Molecular Quantum Numbers	45
5.1.1	Electronic Transitions	45
5.1.2	Rotational Transitions	46
5.1.3	Vibrational Transitions	46
5.1.4	Spectroscopic Terms	47
5.2	Symmetric Top molecules	49
5.3	Ammonia (NH_3)	52
5.3.1	Inversion Transitions	52
5.3.2	Ortho- and Para- NH_3	55
5.3.3	Level Degeneracies	56
5.3.4	Selection Rules	56
5.3.5	The Ammonia Thermometer	56
5.4	Asymmetric Top Molecules	57
5.5	Permanent Electric Dipole Moment	58
6	Excitation Rates	61
6.1	Rate Coefficients	61
6.1.1	Radiative Transition Rates	61
6.1.2	Collisional Rate Coefficients	64
6.2	The Two-Level Atom	65
6.2.1	Full Rate Equation	65
6.2.2	Steady-State Solution	66
6.3	Critical Densities and Thermalization	67
6.4	Boltzmann Plots	68
6.5	The Three-Level Atom	69
7	Spectral Lines	71
7.1	Line Profiles	71
7.1.1	Natural Broadening (Lorentzian profile)	71
7.1.2	Collisional (Pressure) Broadening	72
7.1.3	Doppler Broadening (Gaussian profile)	72
7.1.4	Combined effects (Voight profile)	72
7.1.5	Damping Wings	72
7.2	Absorption	72
7.2.1	Oscillator Strength	72
7.2.2	Optical Depth	73
7.2.3	Equivalent Width and Column Density	75
7.2.4	Optically-thin vs. Saturated Lines	76
7.2.5	Curve of Growth	77
8	Radiative Trapping	79

8.1 Escape Probability Approximation	79
8.2 Geometry: Spherical cloud	81
8.3 Velocity Field: Large Velocity Gradient	82
9 Astrochemistry	85
9.1 Astrochemistry 101	85
9.1.1 Important environments for astrochemistry	86
9.1.2 Some general classes of molecules	87
9.1.3 Types of Reactions	88
9.1.4 Reaction Rates	88
9.1.5 What's what in the ISM?	89
9.2 H ₂	90
9.2.1 Observational Challenges	90
9.2.2 Gas-Phase Formation	91
9.2.3 Grain-Surface Formation	92
9.2.4 Lyman and Werner Bands	93
9.2.5 Photodissociation	93
9.2.6 UV Pumping	94
9.2.7 H ₂ Self-Shielding	94
9.2.8 H ₂ Cooling	94
10 Dust	95
10.1 Dust Properties	95
10.1.1 Composition	95
10.1.2 Grain Sizes	98
10.2 Dust Heating	99
10.2.1 Radiative Heating	99
10.2.2 Collisional Heating	100
10.3 Dust Cooling and Emission	101
10.3.1 Radiative Cooling	101
10.3.2 Dust Emission	102
10.3.3 Steady-State Temperature	102
10.4 Dust Destruction	102
10.4.1 Sputtering	102
10.4.2 Erosion	103
10.4.3 Fragmentation	103
10.5 Drag and Radial Drift	104
10.5.1 Drag and the Meter-Sized Barrier	104
10.5.2 Poynting-Robertson Effect	104
10.6 Additional Barriers to Grain Growth	104
10.6.1 Bouncing Barrier	105
10.6.2 Charging Barrier	105
10.7 Theories for Grain Growth	105
10.7.1 Fragmentation + Mass Transfer	105
10.7.2 Dust Traps	105
11 The Neutral ISM	107
11.1 21 cm emission and absorption	108

11.2 HI Heating	110
11.2.1 Cosmic Ray Heating	110
11.2.2 Photoionization Heating	110
11.3 HI Cooling	111
11.4 The Two-Phase ISM	111
12 Ionized Gas Processes	113
12.1 Ionization	113
12.1.1 Photoionization	113
12.1.2 Collisional Ionization	114
12.1.3 Cosmic Ray Ionization	114
12.2 Recombination	115
12.2.1 Case A Recombination	115
12.2.2 Case B Recombination and Lyman α Resonant Scattering	116
12.3 Heating and Cooling Processes	117
12.3.1 HII Region Heating	119
12.3.2 HII Region Cooling	120
12.3.3 Thermal Equilibrium	120
12.4 Nebular Diagnostics	122
12.4.1 Temperature-Sensitive Line Ratios	123
12.4.2 Density-Sensitive Line Ratios	126
12.4.3 BPT Diagram	127
13 Fluid Dynamics	129
13.1 The Basic Assumptions	129
13.2 Conservation Laws	130
13.2.1 Conservation of Mass	131
13.2.2 Conservation of Momentum	131
13.2.3 Conservation of Energy	132
13.2.4 Adiabatic Flow	132
13.2.5 Isothermal Flow	132
13.3 Sound Waves and Propagation	132
13.4 Shocks	135
13.4.1 Shock Formation	135
13.4.2 Examples of shocks in the ISM:	135
13.4.3 Shock Structure	136
13.4.4 Jump Conditions	137
13.4.5 Strong Shocks	137
13.4.6 Fixed Reference Frame	138
13.4.7 Cooling Time and Cooling Length	139
13.4.8 Shocks and the 3-Phase ISM	140

Chapter 1

Glossary of Fundamental Concepts

When we talk about the interstellar medium, we are dealing with macroscopic systems of microscopic particles (and photons, which are a little more complicated). Despite the incredible difference in densities between terrestrial gas and interstellar gas, we are still talking about systems in which the individual components/particles interact regularly through both short-range and long-range forces.

While Draine Ch 2 deals with defining ways to characterize the frequencies of these interactions, Draine Ch 3 is all about the energies & specifically the thermodynamics. This means considering a lot of different distributions of energies in our system of gas particles, from macroscopic (like translational kinetic energy) to microscopic (where the internal energy states are now small, so much so that they are quantized).

Throughout this course we will focus on different energy regimes, and the different relative importance of long-range forces for different types of particles/conditions. This will help us define some of the different subsets of the ISM we will deal with in this course (plasmas, ionized gas, neutral/molecular gas, and dust).

1.1 Number density – n

A fundamental quantity when talking about a gas is the number density: the number of particles in a specified volume. In chemistry, a related concept would be the concentration of a given reactant. Note that (for whatever the reason) very common volume units encountered in astronomy are cm^{-3} .

We can define the number density n in several ways:

$$n = \frac{N}{V} \tag{1.1}$$

where N is the total number of particles and V is the volume.

We can also relate a number density back to the mass density ρ :

$$n = \frac{\rho}{\mu m_H} \tag{1.2}$$

Here I am making this expression extremely general, and able to encompass a heterogeneous mixture of particles for which the mean particle mass is given by \bar{m}_H .

This requires defining a new quantity: The mean molecular weight μ .

1.2 Mean Molecular Weight – μ

The mean molecular weight is a dimensionless quantity that can be expressed as

$$\mu = \frac{\bar{m}}{m_H} \quad (1.3)$$

\bar{m} can be either the mass of a single particle, or the average mass of an ensemble of particles. As an example, let's calculate the mean molecular mass of air.

If we know the concentration of different species by number, we can write:

$$\bar{m} = \frac{\%A(m_A) + \%B(m_B) + \%C(m_C) + \dots + \%N(m_N)}{100\%} \quad (1.4)$$

(Note that if instead you know the *mass fraction* of a species, your math will be more complicated!)

For dry air, the concentrations of the most abundant species are: N₂ (78.08%), O₂ (20.95%), and Ar (0.934%).

Taking the masses of these species, we can compute

$$\bar{m} = 0.7808(28.013 \text{ amu}) + 0.2095(31.999 \text{ amu}) + 0.00934(39.938 \text{ amu})$$

$$\bar{m} = 28.9 \text{ amu}$$

The mean molecular weight of air is then

$$\mu = \frac{28.9 \text{ amu}}{1.00784 \text{ amu}}$$

$$\mu = 28.9$$

1.3 Collisional Cross section – σ

Now that we have defined the concentration of particles that are present in our gas, we have to consider the possibility (inevitability) that they will interact. To do this, we need to know not just their number density, but a typical particle size. If there were no long-range forces to consider, this would just be determined by a physical radii of the interacting particles (the hard sphere approximation).

In this case, we can define a purely geometric collisional cross section (essentially, the cross-sectional area in space taken up by a particle as it speeds around in some direction). To consider whether a collision will occur, we have to determine the likelihood that a particle will come into contact, which we define as a distance less than the sum of the radius of both particles. For identical particles:

$$\sigma = \pi(2r)^2 = \pi d^2 \quad (1.5)$$

where σ has the expected units of area.

If the particles are not identical, we can more generally state

$$\sigma_{AB} = \pi(r_A + r_B)^2 \quad (1.6)$$

However, in many cases we will consider, the effective particle radius is actually determined by the distance at which an interaction occurs due to a long-range force (resulting in the particle substantially deviating from its original course). In these cases, the collisional cross section becomes substantially more complicated (and illustrating cases in which this is true is one of the main purposes of Draine Ch. 2).

1.4 Mean free Path – τ

The mean free path τ is the distance a typical particle in a gas travels before undergoing an encounter (collision) with another particle.

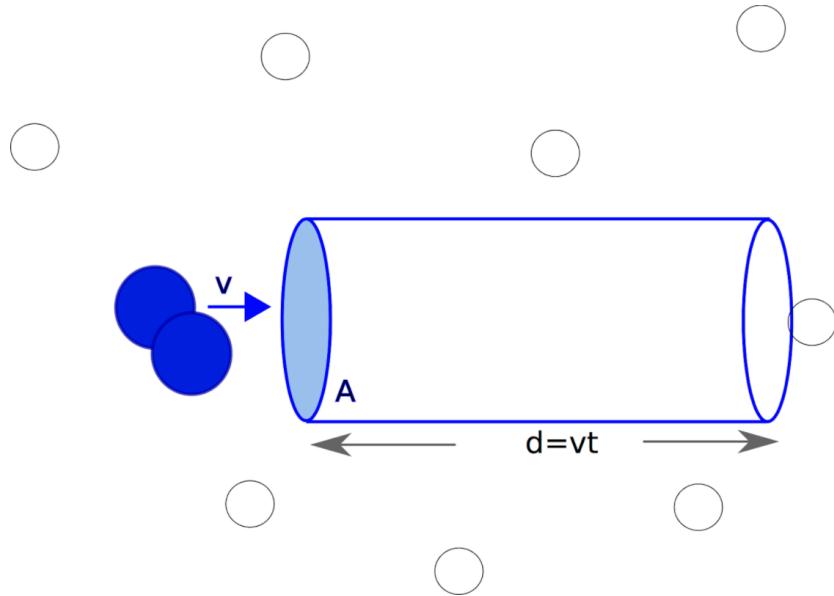


Figure 1.1: collisions.png

This quantity has units of length, and geometrically (using the figure above, in which the area A is essentially equivalent to our collisional cross section σ) we can express it as

$$\tau = \frac{1}{n\sigma} \quad (1.7)$$

Essentially, we are looking for a distance $d = \tau$ that the particle must travel in order for the volume it sweeps out to contain a single particle.

1.5 Two-body Collisional Rate Coefficient – $\langle \sigma v \rangle$

Let's consider a case in which two particles of interest (A and B) interact in some way (e.g., collide). Drawing from chemistry, we can describe this process with a rate equation:



For a two-body interaction, we can describe the rate of collisions *per volume* as

$$\text{Rate}_{AB} = n_A n_B \langle \sigma v \rangle_{AB} \quad (1.9)$$

where v is the typical *relative* velocity of the particles.

Looking at the units, you can see that Rate_{AB} has units of per volume, per time (e.g., $\text{cm}^{-3} \text{ s}^{-1}$)

(Note that you will sometimes see the rate variable written as Z , but we are not doing this in order to avoid confusion with the partition function, which will be discussed later).

In some cases, you also may want to know the rate of collisions experienced by a single particle (for example, relating this rate back to the mean free path, which described how far a single particle can travel before undergoing a collision). In this case, we can say:

$$\text{Rate}_A = n_B \langle \sigma v \rangle_{AB} \quad (1.10)$$

with units of collisions per time (e.g., s^{-1})

The quantity $\langle \sigma v \rangle$ has special significance: this is the two-body collisional rate coefficient. Rate coefficients are usually represented with the variable k :

$$k_{AB} = \langle \sigma v \rangle_{AB} \quad (1.11)$$

This yields the most typical form of a rate equation, in which the rate (in this case, the two-body collision rate) is parameterized by a rate coefficient which has units of $\text{cm}^3 \text{ s}^{-1}$:

$$\text{Rate}_{AB} = k_{AB} n_A n_B \quad (1.12)$$

The rate coefficient can be thought of as the volume in which collisions can occur that particles take up as they move around in the gas. Multiplying by the number densities of both particles then gives the rate at which these particle motions (given their effective sizes) result in a collision.

Practically, this parameterization allows for complex dependencies (for some interactions, the cross-section is velocity-dependent and the velocity itself is often some function of temperature) to be bundled together into a single "constant". Determining the collisional coefficient can then require a somewhat complex integral over the velocity, taking into account the probabilistic distribution of velocities of all the interacting particles. Because of this, values of these rates coefficients

are often tabulated over some relevant range of physical conditions. A number of forms of $\langle \sigma v \rangle$ taking into account interactions via long-range forces are given in Draine Ch 2. For now, we will focus on where exactly this temperature dependency arises in the 'hard sphere' approximation, where the cross section is geometrically determined and only the relative particle velocity v must be determined.

Note that the relatively simple forms of this collisional rate equation will re-appear over and over again, as the two-body collisional rate coefficient is a fundamental component of rates like chemical reaction rates and recombination rates.

1.6 Maxwell-Boltzmann Statistics and Kinetic Temperature – T_{kin}

For our purposes in studying gas in the interstellar medium, the gas will essentially always be well-described by Maxwell-Boltzmann statistics. That is to say:

- (1) Individual gas particles are distinguishable (Unlike for Bose-Einstein statistics, in which individual particles are not distinguishable)
- (2) There is no limit on the number of particles in a system that can share the same internal energy state (Unlike for Fermi statistics, in which each energy state of the system can be occupied by at most one particle).

From the kinetic theory of gases, assuming that the motion of the particles is homogeneous in 3 dimensions, we can then write a probability function that describes the likelihood of finding a particle in our system having a velocity v :

$$f(v) dv = \left(\frac{\bar{m}}{2\pi kT} \right)^{3/2} 4\pi v^2 \exp \left(-\frac{\bar{m}v^2}{2kT} \right) \quad (1.13)$$

where k is the Boltzmann constant.

By definition, the temperature in this equation is the gas kinetic temperature T_{kin} .

This distribution has some similarities with a Gaussian, and its mean value can be similarly used to describe the average random speed of a particle (e.g., the expectation value for the velocity distribution of particles moving in 3 dimensions– note however that this is not the same as the most *probable* speed for a particle in this ensemble, which is the mode of the velocity distribution). This quantity is often called the thermal velocity, and it can be expressed as

$$v_{th} = \sqrt{\frac{8kT}{\pi\bar{m}}} \quad (1.14)$$

1.7 Internal Energy States

So far we have been considering our system of gas particles relatively macroscopically, looking at their speeds and interaction frequency.

However, now we have to consider that collisions between gas particles are not perfectly elastic, and may not only alter the trajectory of a particle, but may change its internal energy state (or even result in some combination of the two particles).

Internal energy states of atoms and molecules are discrete and quantized values.

For individual atoms, internal energy states are simply (or not so simply) governed by electron orbitals.

For molecules, the internal energy states become more complicated. In addition to the electronic energy states from the electron orbitals of each atom in the molecule, the molecule also has a set of rotational energy states due to the quantization of angular momentum around non-degenerate rotational axes of the molecule. Finally, depending on the complexity of the molecule, it will have one set (or many sets) of vibrational energy states due to different vibrational modes (bending and stretching and overtones thereof) of the intramolecular bonds.

Details of these different states, including quantum selection rules, will be discussed later in the semester.

1.8 Boltzmann Distribution and Excitation Temperature – T_{ex}

For gas obeying Boltzmann statistics, we can describe the probability p that a particle in the gas is in a given internal energy state i as

$$p_i \propto \exp\left(-\frac{E(i)}{kT}\right) \quad (1.15)$$

where p_i is the probability of being in state i and $E(i)$ is the energy of the i^{th} state.

We can construct a ratio of probabilities for the population of two different energy states as:

$$\frac{p_i}{p_j} = \frac{g_i}{g_j} \exp\left(-\frac{E(j) - E(i)}{kT}\right) \quad (1.16)$$

where g_i and g_j are the degeneracies of each state

(basically, the degeneracy is the number of distinct quantum states of the atom/molecule having identical energies– states which, under 'typical' conditions, are observationally indistinguishable)

The ratio of probabilities can also be related to the number of atoms/molecules in each state, and this equation can then be rewritten as

$$\frac{N_i}{N_j} = \frac{g_i}{g_j} \exp\left(-\frac{E(j) - E(i)}{kT_{ex}}\right) \quad (1.17)$$

This equation (I will sometimes refer to it as the Boltzmann Equation) defines the observed Excitation Temperature (T_{ex}) of a system.

1.9 Partition Functions – Z

To exactly determine the probability that an atom (or molecule) is in a specific energy state requires knowledge of the full probability distribution for all allowed energy states of the atom. This

quantity is called the Partition Function. We can define the partition function by requiring that the sum of the probabilities of occupying all possible internal energy states s of the atom is 1:

$$\sum_s p_i = 1$$

This then allows us to normalize our original expression for p_i :

$$p_i = \frac{1}{\sum_s \exp\left(-\frac{E(s)}{kT}\right)} \exp\left(-\frac{E(i)}{kT}\right) \quad (1.18)$$

We then define this normalizing factor as the partition function:

$$Z = \sum_s \exp\left(-\frac{E(s)}{kT}\right) \quad (1.19)$$

Note that the partition function is separable into different components (e.g., the probabilities for the occupation of the electronic, rotational, and vibrational energy states):

$$Z_{int} = Z_e Z_{rot} Z_{vib} \quad (1.20)$$

As may be seen later, these can be written as:

$$Z_e = \sum_s \exp\left(-\frac{E_e(s)}{kT}\right) \quad (1.21)$$

$$Z_{rot} = \sum_{J=0}^{\infty} (2J+1) \exp\left(-\frac{BJ(J+1)}{kT}\right) \quad (1.22)$$

where J is the angular momentum quantum number and B is the rotational constant of the molecule.

$$Z_{vib} = \sum_{v=0}^{\infty} \exp\left(-\frac{h\nu_0(v + \frac{1}{2})}{kT}\right) \quad (1.23)$$

where v is the vibrational quantum number of the molecule

We also ignored in our original formulation the translational energy of the atoms/molecules. As this distribution of energies is a continuous rather than a discrete function, it can be determined from an integral over the Maxwell-Boltzmann distribution.

$$Z_{trans} = \frac{(2\pi mkT)^{3/2}}{h^3} V \quad (1.24)$$

where V is the volume of the gas

(note that here m is not the mean mass, but is instead the mass of just the single species being considered).

We can then express the total partition function

$$Z_{tot} = \frac{(2\pi mkT)^{3/2}}{h^3} V Z_{int} \quad (1.25)$$

1.10 LTE: Local Thermodynamic Equilibrium

While in Draine Ch. 3 all of these equations referred to temperature generically as T , the truth is more complicated: in general, in the ISM, we do not expect all of these temperatures to be the same. It may be that

$$T_{kin} \neq T_{ex} \neq T_i$$

The condition in which all of the temperatures are identical is known as Local Thermodynamic Equilibrium, or LTE. This requirement encapsulates a wide range of temperatures, including some that we have not discussed (the radiation temperature, which is often parameterized as the effective (Blackbody) temperature T_{eff} defined by the Planck Function, is an important one that we will come to soon).

Luckily for us, while LTE is not generally an overall valid approximation, we generally can find that individual processes (e.g., chemistry or ionization or excitation) are in equilibrium, and we can use these to determine a valid and representative temperature that describes this aspect of the physics of the system.

As an example of where we experience non-LTE conditions in real life, consider the sun shining outside right now. Its radiation has a Blackbody temperature of $T_{eff} = 5800$ K – but luckily for us, this is not the same as the kinetic temperature of gas in the atmosphere, which sets the temperature we experience when we walk outside.

1.11 An Aside on the Concept of an Ideal Gas

Many of the variables we have discussed so far (e.g., Temperature, Volume and Number of particles) can be related together with a thermodynamic equation of state. A common equation of state that you may have encountered is the Ideal Gas Law:

$$PV = NkT \quad (1.26)$$

or

$$P = nkT \quad (1.27)$$

It turns out that there are a number of assumptions underlying this relation that need to be satisfied in order for it to hold true. These include:

- (1) The gas consists of a large number of molecules that are in random motion and obey Newton's laws of motion.
- (2) The particles are far enough apart (so that the volume of the individual particles is much smaller than the volume of the gas)
- (3) The particles are not TOO far apart (so that they undergo many collisions before crossing a region, and the velocity distribution can be described by kinetic temperature).
- (4) All collisions are perfectly elastic.
- (5) Apart from brief elastic collisions, there are no intramolecular forces.

Chapter 2

Radiative Transfer

2.1 Definition of Standard Quantities

The light emitted from or passing through objects in space is almost the only way that we have to probe the vast majority of the universe we live in. The most distant object object to which we have travelled and brought back samples, besides the moon, is a single asteroid. Collecting solar wind gives us some insight into the most tenuous outer layers of our nearby star, and meteorites on earth provide insight into planets as far away as Mars, but these are the only things from space that we can study in laboratories on earth. Beyond this, we have sent unmanned missions to land on Venus, Mars, another asteroid and a comet (kind of). To study anything else in space we have to interpret the radiation we get from that source. As a result, understanding the properties of radiation, including the variables and quantities it depends on and how it behaves as it moves through space, is then key to interpreting almost all of the fundamental observations we make as astronomers.

2.1.1 Energy

To begin to define the properties of radiation from astronomical objects, we will start with the energy that we receive from an emitting source somewhere in space. Consider a source of radiation in the vacuum of space (for familiarity, you can think of the sun). At some point in space away from our source of radiation we want to understand the amount of energy dE that is received from this source. What is this energy proportional to?

As shown in Figure 2.1, our source of radiation has an intensity I_0 (we will get come back to this in a moment) over an apparent angular size (solid angle) of $d\Omega$. Though it may give off radiation over a wide range of frequencies, as is often the case in astronomy we only concern ourselves with the energy emitted in a specific frequency range $\nu + d\nu$ (think of using a filter to restrict the colors of light you see, or even just looking at something with your eyeball, which only detects radiation in the visible range). At the location of detection, the radiation passes through some area dA in space (an area perhaps like a spot on the surface of earth) at an angle θ away from the normal to that surface. The last property of the radiation that we might want to consider is that we are detecting it over a given window of time (and many astronomical sources are time-variable). You might be wondering why the distance between our detector and the source is not being mentioned yet: we will get to this.

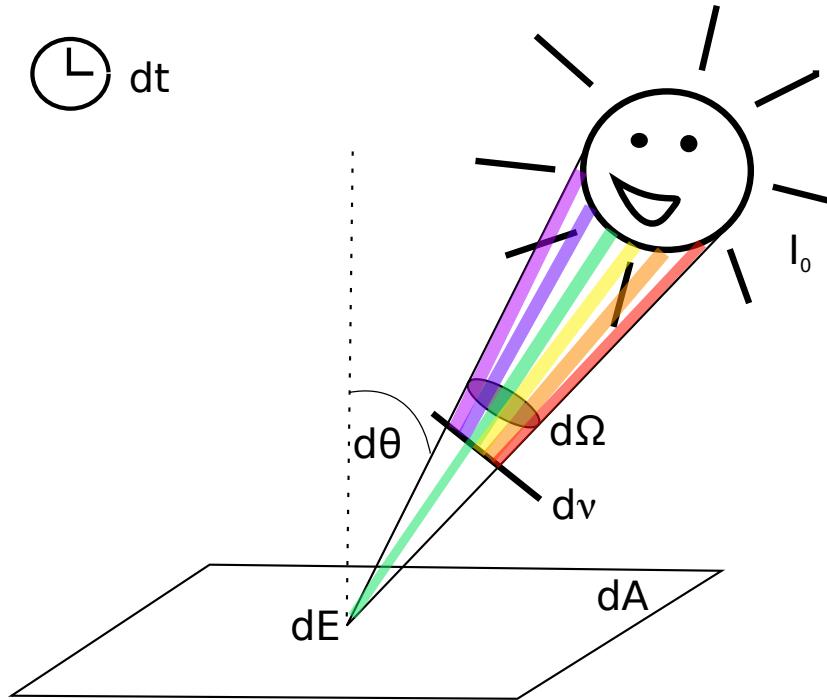


Figure 2.1: Description of the energy detected at a location in space for a period of time dt over an area dA arriving at an angle θ from an object with intensity I_0 , an angular size $d\Omega$, through a frequency range $\nu, \nu+dv$

Considering these variables, the amount of energy that we detect will be proportional to the apparent angular size of our object, the range of frequencies over which we are sensitive, the time over which we collect the radiation, and the area over which we do this collection. The constant of proportionality is the specific intensity of our source: I_0 . Technically, as this is the intensity just over a limited frequency range, we will write this as $I_{0,\nu}$.

In equation form, we can write all of this as:

$$dE_\nu dv = I_{0,\nu} \cos\theta dA d\Omega dv dt \quad (2.1)$$

Here, the $\cos\theta dA$ term accounts for the fact that the area that matters is actually the area "seen" from the emitting source. If the radiation is coming straight down toward our unit of area dA at an angle of 0, it "sees" an area equal to that of the full dA ($\cos 0 = 1$). However, if the radiation comes in at a different angle θ , then it "sees" our area dA as being tilted: as a result, the apparent area is smaller ($\cos\theta < 1$). You can test this for yourself by thinking of the area dA as a sheet of paper, and observing how its apparent size changes as you tilt it toward or away from you.

2.1.2 Intensity

Looking at Equation 2.1, we can figure out the units that the specific intensity must have: energy per time per frequency per area per solid angle. In SI units, this would be $\text{W Hz}^{-1} \text{ m}^{-2} \text{ sr}^{-1}$. Specific intensity is also sometimes referred to as surface brightness, as this quantity refers to the brightness over a fixed angular size on the source. Technically, the specific intensity is a

7-dimensional quantity: it depends on position (3 space coordinates), direction (two more coordinates), time (in general, we won't worry too much about this) as well as frequency (it's freaking complicated).

2.1.3 Flux

The flux from a source is defined as the total energy of radiation received from all directions at a point in space, per unit area and per unit time. Given this definition, we can modify equation 2.1 to give the flux at a frequency ν :

$$F_\nu = \int_{\Omega} \frac{dE_\nu d\nu}{dAdt} = \int_{\Omega} I_\nu \cos\theta d\Omega \quad (2.2)$$

The total flux at all frequencies (the bolometric flux) is then:

$$F = \int_{\nu} F_\nu d\nu \quad (2.3)$$

As expected, the SI units of flux are W m^{-2}

The last, related property that one should consider (particularly for spatially well-defined objects like stars) is the Luminosity. The luminosity of a source is the total energy emitted per unit time. The SI unit of luminosity is W .

Luminosity can be determined from the flux of an object by integrating over its entire surface:

$$L = \int F dA \quad (2.4)$$

Having defined these quantities, we now ask how the flux you detect from a source varies as you increase the distance to the source. Looking at Figure 2.2, we take the example of our happy sun, and imagine two shells or bubbles around the sun: one at a distance R_1 , and one at a distance R_2 . The amount of energy passing through each of these shells per unit time is the same: in each case, it is equal to the luminosity of the sun. However, as $R_2 > R_1$, the surface area of the second shell is greater than the first shell. Thus, the energy is spread thinner over this larger area, and the flux (which by definition is the energy per unit area) must be smaller for the second shell. Comparing the equations for surface area, we see that flux decreases proportional to $1/d^2$.

We have showed that the flux obeys an inverse square law with distance from a source. How does the specific intensity change with distance? The specific intensity can be described as the flux divided by the angular size of the source. We have just shown that the flux decreases with distance, proportional to $1/d^2$. What about the angular source size? It happens that the source size also decreases with distance, proportional to $1/d^2$. As a result, the specific intensity is a quantity that is independent of distance.

2.2 The Equation of Radiative Transfer

We can use the fact that the specific intensity does not change with distance to begin deriving the radiative transfer equation. For light traveling in a vacuum along a path length s , we say that the intensity is a constant. As a result,

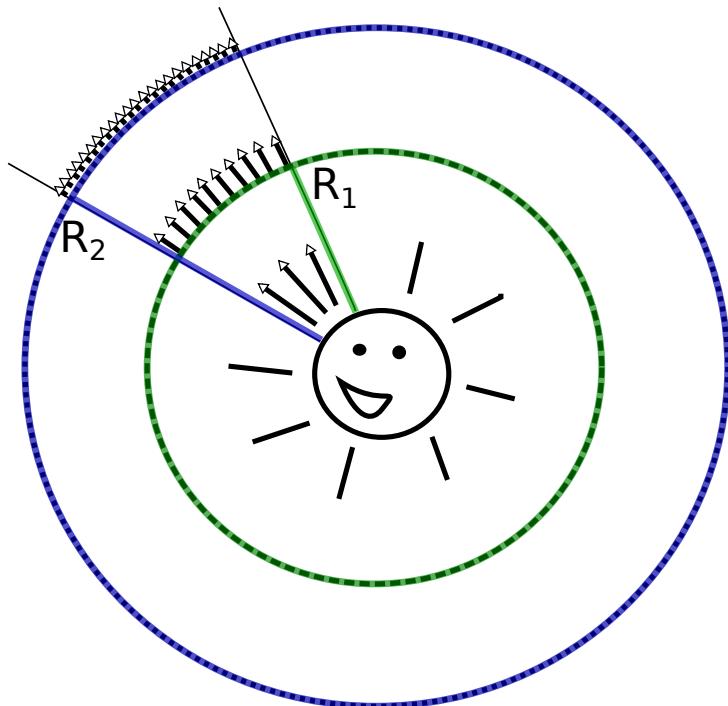


Figure 2.2: A depiction of the flux detected from our sun as a function of distance from the sun. Imagining shells that fully enclose the sun, we know that the energy passing through each shell per unit time must be the same (equal to the total luminosity of the sun). As a result, the flux must be less in the larger outer shell: reduced proportional to $1/d^2$

$$\frac{dI_\nu}{ds} = 0 \text{ (for radiation traveling through a vacuum)} \quad (2.5)$$

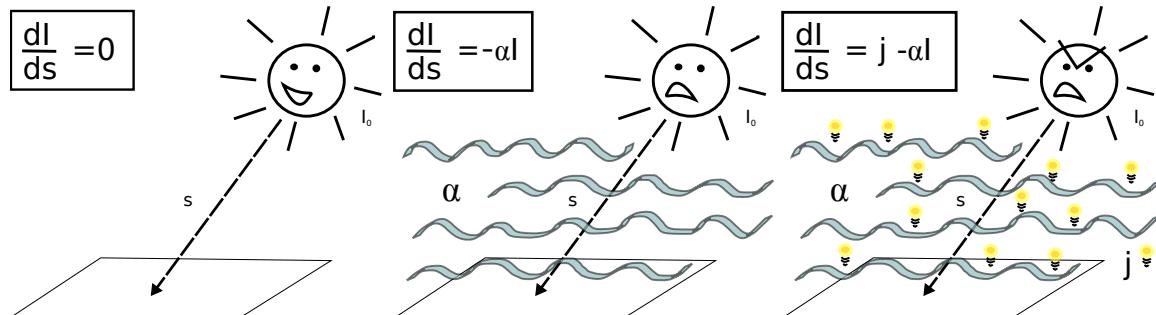


Figure 2.3: The radiative transfer equation, for the progressively more complicated situations of: (left) radiation traveling through a vacuum; (center) radiation traveling through a purely absorbing medium; (right) radiation traveling through an absorbing and emitting medium.

This case is illustrated in the first panel of Figure 2.3. However, space (particularly objects in space, like the atmospheres of stars) is not a vacuum everywhere. What about the case when there is some junk between our detector and the source of radiation? This possibility is shown in the second panel of Figure 2.3. One quickly sees that the intensity you detect will be less than it was at the source. You can define an extinction coefficient α for the space junk, with units of extinction (or fractional depletion of intensity) per distance (path length) traveled, or m^{-1} in SI units. For our

purposes right now, we will assume that this extinction is uniform and frequency-independent (but in real life of course, it never is). Now, our equation of radiative transfer has been modified to be:

$$\frac{dI_\nu}{ds} = -\alpha I_\nu \text{ (when there is some junk between us and our source)} \quad (2.6)$$

As is often the case when simplifying differential equations, we then find it convenient to try to get rid of some of these pesky units by defining a new unitless constant: τ , or optical depth. If α is the fractional depletion of intensity per path length, τ is just the fractional depletion. We then can define

$$d\tau = \alpha ds \quad (2.7)$$

and re-write our equation of radiative transfer as:

$$\frac{dI_\nu}{d\tau} = -I_\nu \quad (2.8)$$

Remembering our basic calculus, we see that this has a solution of the type

$$I = I_0 e^{-\tau} \quad (2.9)$$

So, at an optical depth of 1 (the point at which something begins to be considered optically thick), your initial source intensity I_0 has decreased by a factor of e.

However, radiation traveling through a medium does not ALWAYS result in a net decrease. It is also possible for the radiation from our original source to pass through a medium or substance that is not just absorbing the incident radiation but is also emitting radiation of its own, adding to the initial radiation field. To account for this, we define another coefficient: j_ν . This emissivity coefficient has units of energy per time per volume per frequency per solid angle. Note that these units (in SI: $\text{W m}^{-3} \text{ Hz}^{-1} \text{ sr}^{-1}$) are slightly different than the units of specific intensity. Including this coefficient in our radiative transfer equation we have:

$$\frac{dI_\nu}{ds} = j - \alpha I_\nu \quad (2.10)$$

or, putting it in terms of the dimensionless variable τ , we have:

$$\frac{dI_\nu}{d\tau} = j/\alpha - I_\nu \quad (2.11)$$

The quantity j/α is typically defined as the Source function S. Making this substitution, we arrive at the final form of the radiative transfer equation:

$$\frac{dI_\nu}{d\tau} = S_\nu - I_\nu \quad (2.12)$$

What is the solution of this equation? For now, we will again take the simplest case, and assume that the medium through which the radiation is passing is uniform. In this case,

$$I_\nu = I_{\nu,0} e^{-\tau_\nu} + S_\nu (1 - e^{-\tau_\nu}) \quad (2.13)$$

What happens to this equation when τ becomes large? In this case, the term $e^{-\tau_\nu}$ becomes negligible, and you arrive at the result

$$I_\nu = S_\nu \text{ (for an optically-thick source)} \quad (2.14)$$

This makes some sense: travel far enough through an absorbing medium, and the background radiation I_0 is totally blocked, and the only radiation that makes it out is from the emission of the medium itself. So, what the hell is this source function thing?

For a source in thermodynamic equilibrium, you can show that $S_\nu = B_\nu$, the blackbody function.

$$B_\nu = \frac{2h\nu^3}{c^2} \left[e^{\frac{h\nu}{kT}} - 1 \right]^{-1} \quad (2.15)$$

For an optically-thick source (say, a star like our sun) we can use Equation 2.14 to then say that $I_\nu = B_\nu$.

This gives us the ability to define key properties of stars – like their flux and luminosity – as a function of their temperature. Using equations 2.2 and 2.3, we can integrate the blackbody function to determine the flux of a star (or other blackbody) as a function of temperature:

$$F = \sigma T^4 \quad (2.16)$$

This fundamental result is known as the Stefan-Boltzmann law.

Another classic result, the peak frequency (or wavelength) at which a star (or other blackbody) radiates, based on its temperature, can be found by differentiating the blackbody equation with respect to frequency (or wavelength). The result must be found numerically, but the peak wavelength can be expressed as

$$\lambda_{peak} = \frac{2.898 \times 10^{-3}}{T} \quad (2.17)$$

This relation is known as Wien's law.

2.3 Brightness Temperature

Radio and millimeter astronomers use two main units to talk about specific intensity.

The first is the Jansky:

$$1\text{Jy} = 10^{-26}\text{Wm}^{-2}\text{Hz}^{-1}$$

You should note that a Jansky alone is not a unit of specific intensity, as there is no dependence on source solid angle. Radio astronomers typically get around this by image specific intensity in terms of Jy/beam, where the beam has units of solid angle (In radio astronomy, a beam is roughly equivalent to the point spread function: it represents the resolution limit of the telescope. It is generally specified as an ellipse, with major and minor axis length, and a position angle)

Radio astronomers also often describe the specific intensity of a source using a quantity known as Brightness temperature. This quantity relates the specific intensity I_ν of an observed source to that of a blackbody function (B_ν) at a temperature T .

As Draine correctly notes, this is only a linear relation in the Rayleigh-jeans limit ($h\nu \ll kT$). You can see this visually if you plot the Planck function as a function of frequency, in log-space: the specific intensity at a given temperature is proportional to ν^2 which leads to a constant (linear) slope of +2 in the plot.

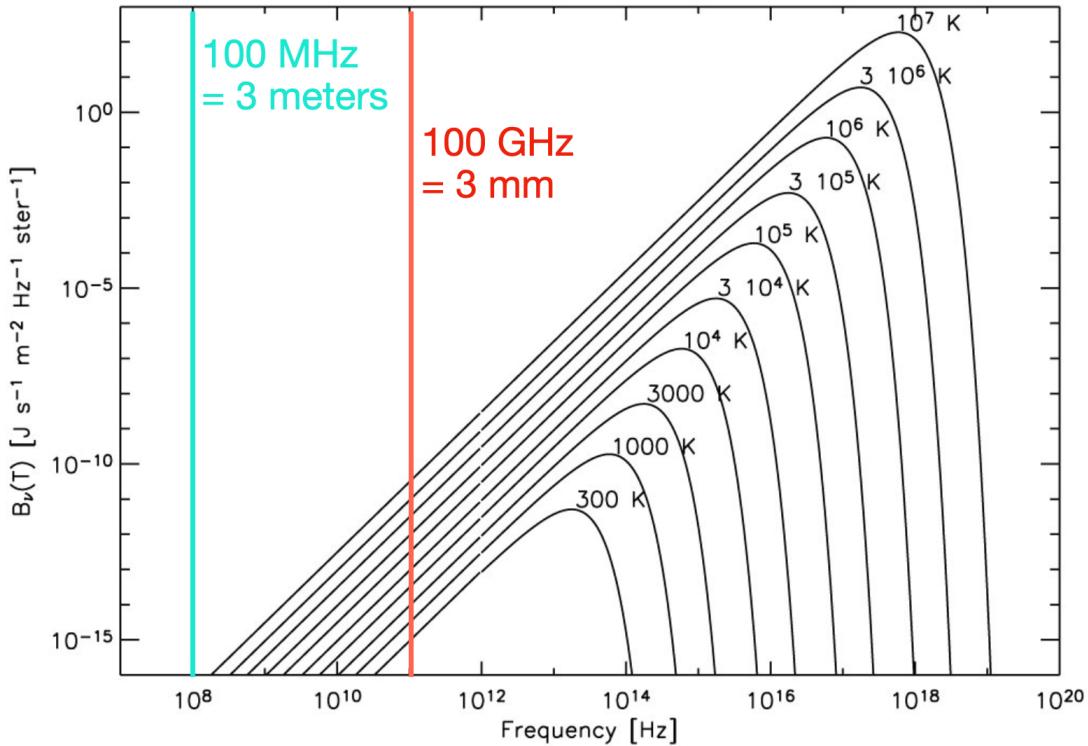


Figure 2.4: *RJ_limit.png*

This quantity is then typically only used by radio and millimeter astronomers, for observations at wavelengths where the Rayleigh-Jeans limit applies. In this case:

$$B_\nu = \frac{2\nu^2 k_B T}{c^2} \quad (2.18)$$

and we can define the brightness temperature as

$$T_B = \frac{c^2}{2k_B \nu^2} I_\nu \quad (2.19)$$

IMPORTANT NOTE: This is the same as Draine's definition of Antenna Temperature (T_A), which appears to be totally wrong. In practice, Antenna temperature is an analogous but more observationally/instrumentally-oriented quantity: the temperature of a blackbody (in practice, usually a thermal resistor) that would lead to an equivalent amount of power per unit frequency being detected by the antenna/telescope (after a lot of corrections for the antenna efficiency)

$$T_A = \frac{P_\nu}{k_B} \quad (2.20)$$

Antenna temperature is a convenient quantity because:

1. 1 K of antenna temperature is a conveniently small power per unit bandwidth. $T_A = 1 \text{ K}$ corresponds to $P_\nu = kT_A = 1.38 \times 10^{-23} \text{ J K}^{-1} \times 1 \text{ K} = 1.38 \times 10^{-23} \text{ W Hz}^{-1}$.
2. It can be calibrated by a direct comparison with matched resistors connected to the antenna receiver input.
3. The units of system noise are also K, so comparing the signal in K with the receiver noise in K makes it easy to compare the signal and noise powers.

(definitions taken from <https://www.cv.nrao.edu/~sransom/web/Ch3.html>)

Chapter 3

Continuum Emission Mechanisms

3.1 Thermal Plasma

3.1.1 Free-Free Emissivity and Absorptivity

Also known as Bremsstrahlung or electron braking radiation, free-free emission is a type of continuum emission (e.g., emission that is not quantized into discrete allowed energies). A result of the conservation of energy, photons are emitted by electrons that scatter off of ions, in the process decelerating and losing kinetic energy. As this is emission from a thermal plasma, the energies involved range from very small changes in the overall energy (long-wavelength radio emission) up to energies on the order kT , where a typical thermal plasma temperature is 10^4 K.

We can write the free-free emissivity of a thermal plasma (the power that is radiated by this process per unit frequency per unit volume per steradian) as:

$$j_{ff,\nu} = \frac{8}{3} \left(\frac{2\pi}{3} \right)^{1/2} g_{ff} \frac{e^6}{m_e^2 c^3} \left(\frac{m_e}{kT_e} \right)^{1/2} e^{-hv/kT_e} n_e Z_i^2 n_i \quad (3.1)$$

Some points of interest to note here: in the Rayleigh-Jeans limit (the radio regime), $hv \ll kT_e$ and $e^{-hv/kT_e} \sim 1$. If for some wild reason you want to evaluate this by hand, don't forget that this formulation assumes cgs units (e.g., Statcoulomb units of charge). n_e is the electron density and n_i is the ion density; for a pure-hydrogen nebula, $n_i = n_p$ or the proton density. Z_i is the ion charge number; again, for a pure-hydrogen nebula, this is just 1. Finally, g_{ff} is the Gaunt factor.

We are not going to dig too deeply into the Gaunt factor, as Draine is itself not particularly chatty on this subject, merely ascribing the Gaunt factor to "quantum effects" and giving some analytic approximations to a quantity that typically must be evaluated numerically. Classically, $g_{ff} = 1$, and we will note that if g_{ff} is a constant, then the emissivity would not vary as a function of frequency, and the spectral energy distribution of free-free emission would be essentially flat as a function of frequency.

We can evaluate two limiting conditions for estimates of the Gaunt factor: the "plasma frequency" $\nu_p = 8.98(n_e/\text{cm}^{-3})^{0.5}$ kHz and the frequency corresponding to the thermal plasma energy $\nu_{th} = kT_e/h$. For a typical electron density of $n_e \sim 10^3 \text{ cm}^{-3}$ and electron temperature of $T = 10^4$ K, $\nu_p = 300$ kHz and $\nu_{th} = 220$ THz. For $\nu n u_p \ll \nu \ll \nu_{th}$, which is well-satisfied by typical radio/mm

emission with frequencies 1-100 GHz, the Gaunt factor (and thus the emissivity and the overall shape of the free-free spectrum) is proportional to $\nu^{-0.118}$.

Remembering the definition of the source function $\frac{j_\nu}{\kappa_\nu} = S_\nu$, and noting that for a thermal plasma we can say $S_\nu = B_\nu$, we can use our definition of the free-free emissivity to determine the free-free absorption coefficient. For simplicity, we will stick to the Rayleigh-Jeans or radio regime, where $B_\nu = \frac{2\nu^2 k_B T}{c^2}$. Then we can define

$$\kappa_{ff,\nu} = \frac{4}{3} \left(\frac{2\pi}{3} \right)^{1/2} g_{ff} \frac{e^6}{m_e^2 c \nu^2} \left(\frac{m_e}{kT_e} \right)^{1/2} g_{ff} n_e Z_i^2 n_i \quad (3.2)$$

3.1.2 Emission Measure

The emission measure (*EM*) of a thermal plasma is both a quantity that is relatively easy to determine from observations, and a quantity that describes several useful properties of the plasma.

Draine introduces the emission measure starting with the equation of radiative transfer for a uniform-temperature source:

$$I_\nu = I_{\nu,0} e^{-\tau} + \left[\frac{j_{ff,\nu}}{\kappa_{ff,\nu}} \right] (1 - e^{-\tau}) \quad (3.3)$$

(note that if there is no background radiation source for which are worried about the propagation of radiation through the nebula, the $I_{\nu,0}$ term can be ignored.)

We can rewrite the equation of radiative transfer by incorporating the emission measure:

$$I_\nu = I_{\nu,0} e^{-\tau} + \left[\frac{j_{ff,\nu}}{n_e n_p} \right] \frac{(1 - e^{-\tau})}{\tau} EM \quad (3.4)$$

Or, ignoring the background radiation term,

$$I_\nu = \left[\frac{j_{ff,\nu}}{n_e n_p} \right] \frac{(1 - e^{-\tau})}{\tau} EM \quad (3.5)$$

Then, by definition,

$$EM \equiv \int n_e n_p ds = \left[\frac{n_e n_p}{\kappa_{ff,\nu}} \right] \tau \quad (3.6)$$

For a pure-hydrogen nebula, $n_e = n_p$ and we can then just write

$$EM = \int n_e^2 ds \quad (3.7)$$

Since $\kappa_{ff,\nu}$ is itself proportional to $n_e n_i \sim n_e^2$ this means the emission measure is basically proportional to the optical depth of the nebula. Note that the typical units of the emission measure are *very* non-standard: they are pc cm⁻⁶.

For a spherical, uniform-density HII region you can skip the integral and define the peak emission measure as

$$EM_{peak} = n_e^2 D_{HII} \quad (3.8)$$

where D_{HII} is the diameter of the HII region.

Example: Draine 10.4

Consider an HII region with $n(\text{H}^+) = n_e = 10^3 \text{ cm}^{-3}$, $T = 8000 \text{ K}$, and radius $R = 1 \text{ pc}$. Estimate the radio frequency at which the optical depth across the diameter of the HII region is $\tau = 1$. To make this estimate you may assume that the Gaunt factor $g_{ff} \sim 6$.

We will use two equations:

(1) Draine 10.14:

$$\frac{\kappa_{ff,\nu}}{n_e n_i} \approx \frac{4}{3} \left(\frac{2\pi}{3} \right)^{1/2} \frac{e^6}{(m_e k_B T)^{3/2} c \nu^2} Z_i^2 g_{ff}$$

(2) Draine 7.14

$$\tau_\nu = \int_s \kappa_\nu ds$$

For $s = D_{HII}$

$$\tau_\nu = \kappa_\nu D_{HII}$$

Putting these together (solving for $\tau = 1$)

$$\nu^2 = \frac{4}{3} \left(\frac{2\pi}{3} \right)^{1/2} \frac{e^6}{c (m_e k_B T)^{3/2}} Z_i^2 g_{ff} n_e n_i D_{HII}$$

3.2 Continuum Radiation

Draine really discusses continuum radiation mechanisms only in the context of background radiation fields. This is probably a correct and complete description of continuum radiation for a purely theoretical consideration of the ISM, however it misses some important nuances for observers. The strength and frequency-dependence of the ambient (background) radiation field is indeed extremely important to consider when considering ionization, absorption and stimulated emission processes in the ISM gas. Background radiation from starlight, dust-reprocessed starlight, X-ray emission from plasma can all be expected to be present and impacting the ISM conditions.

In general, my issue with Draine here is that it does not consider the fundamentally inhomogeneous nature of the ISM. Focusing on the solar neighborhood and defining typical intensities of the background radiation in this part of our galaxy is important, but is also not representative of many environments interesting to study (the centers of galaxies, regions in close proximity to massive stars and star clusters, and the interiors of dense clouds). In particular, inside of dense clouds

the background stellar radiation field is unimportant, though you can still expect limited penetration from X-rays. Here some of the most important continuum processes to consider (apart from the CMB background) are not really discussed: thermal dust radiation, and the ionizing radiation from cosmic ray interactions.

Direct observations of the continuum radiation itself (particularly the free-free and dust thermal emission) can yield valuable insight into the properties of the ISM. However, when studying discrete sources of continuum emission one very often has to contend with the fact that these continuum emission sources are not present in isolation. The very nature of the cycle of star formation (particularly massive star formation) means that you can have dust emission from dense clouds, free-free radiation from the HII regions surrounding massive stars, and synchrotron emission from young supernova remnants, all in close proximity. Here, I supplement Draine with a brief overview of the observational properties of three main types of continuum emission that you would expect to observe at wavelengths from the radio to the infrared.

3.2.1 Synchrotron Radiation

Although it is intrinsically low-frequency radiation, synchrotron is associated with high-energy processes. In fact, if one looks at multiwavelength images of the Milky Way, one sees many more similarities between synchrotron emission and the x-ray and gamma-ray emission than with other radio/IR emission sources like ionized gas, neutral gas, and dust.

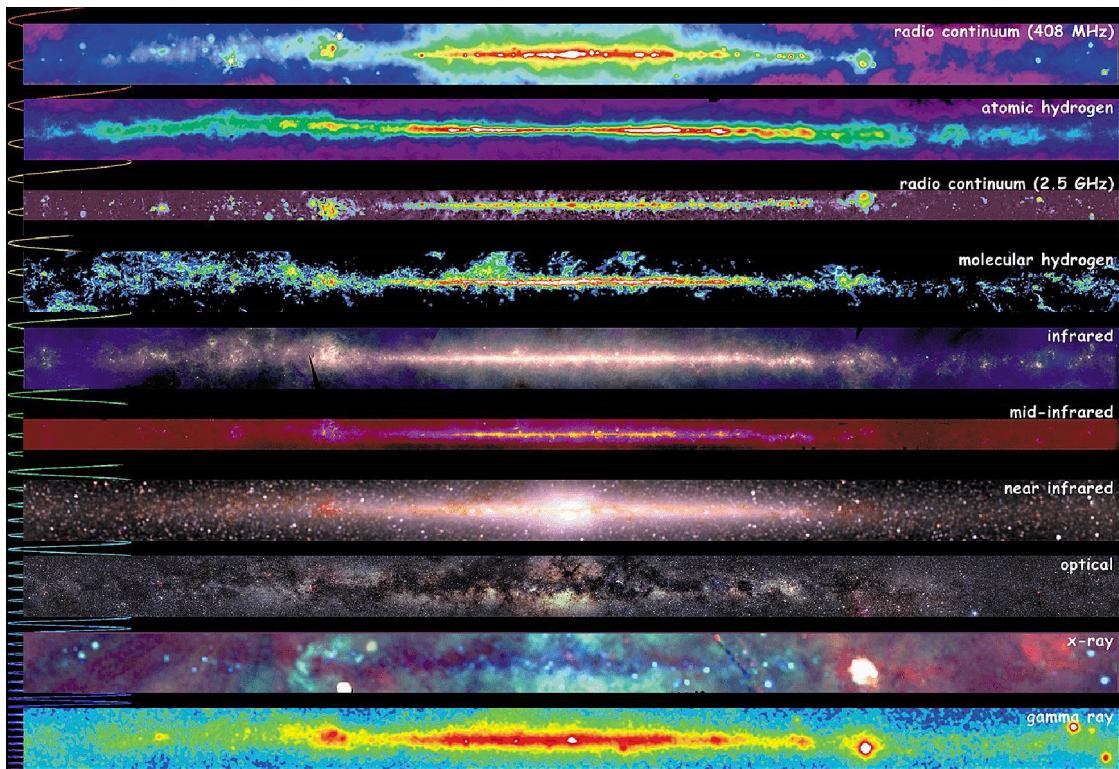


Figure 3.1: MW_plane.jpeg

Indeed, synchrotron and gamma-ray emission processes can be directly linked, as the same relativistic electrons that give off synchrotron photons are also responsible for inverse Compton scattering.

tering of photons to gamma-ray energies/frequencies.

The frequency-dependence of synchrotron emission, resulting in the observed spectrum of this radiation, is dependent on several variables, including the energy distribution of the radiating electrons. An important consideration is that the lower-frequency photons are more likely to undergo scattering, which hinders (or eventually prevents) these photons from leaving the plasma. This is called synchrotron self-absorption, and is analogous to the situation where a gas becomes optically-thick. Taking self-absorption into account, we get a synchrotron spectrum that has 3 distinct components: an optically-thick (self-absorbed) increasing spectrum at low frequencies, a turnover where self-absorption becomes unimportant, and an optically thin decreasing spectrum at higher frequencies.(this is the most commonly-observed portion of the spectrum at radio wavelengths). The synchrotron spectrum of an idealized, homogeneous cylindrical source is shown below.

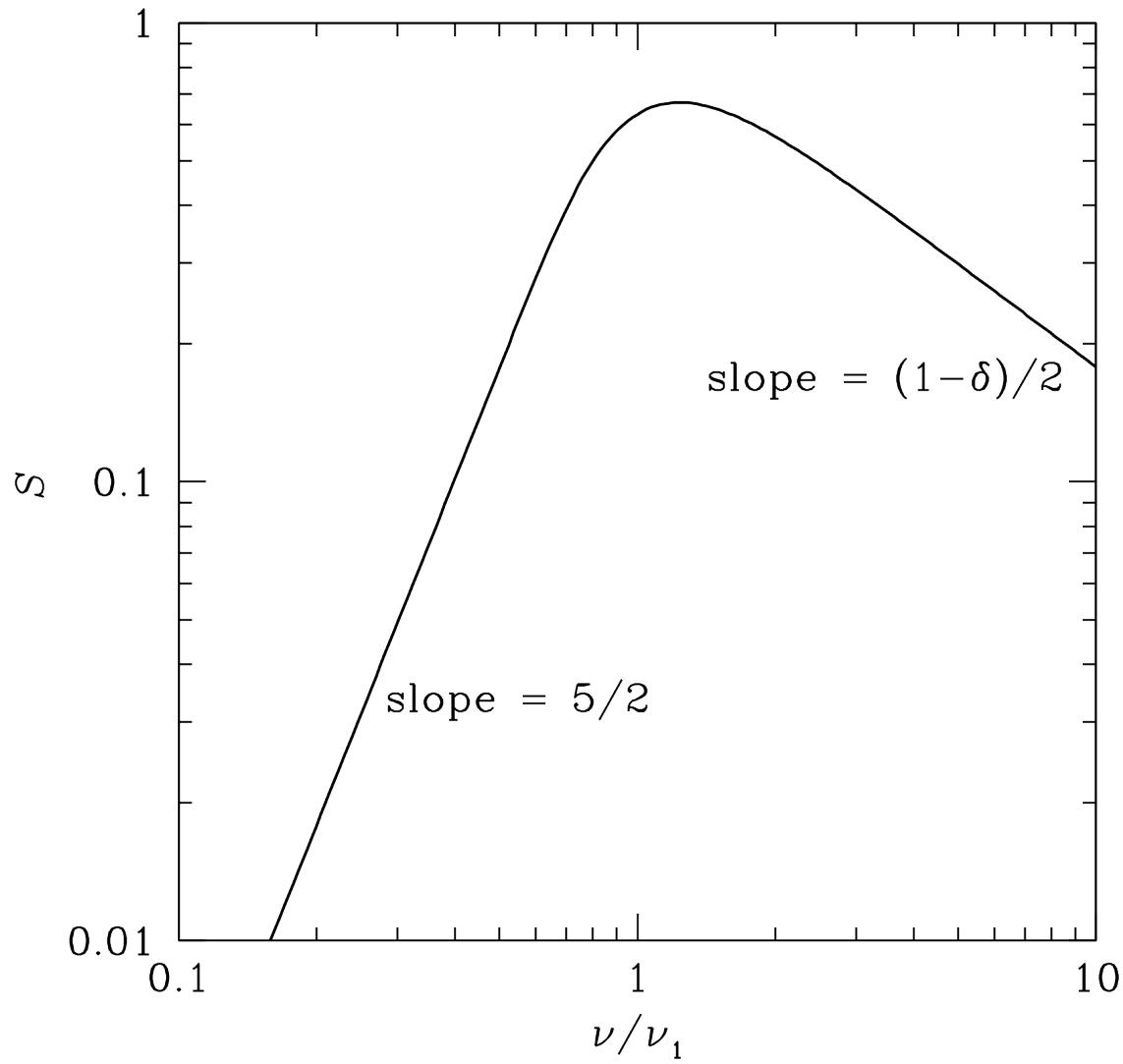


Figure 3.2: *synchrotron.png*

The two roughly linear portions of the spectrum are generally described as power laws with a spectral index α :

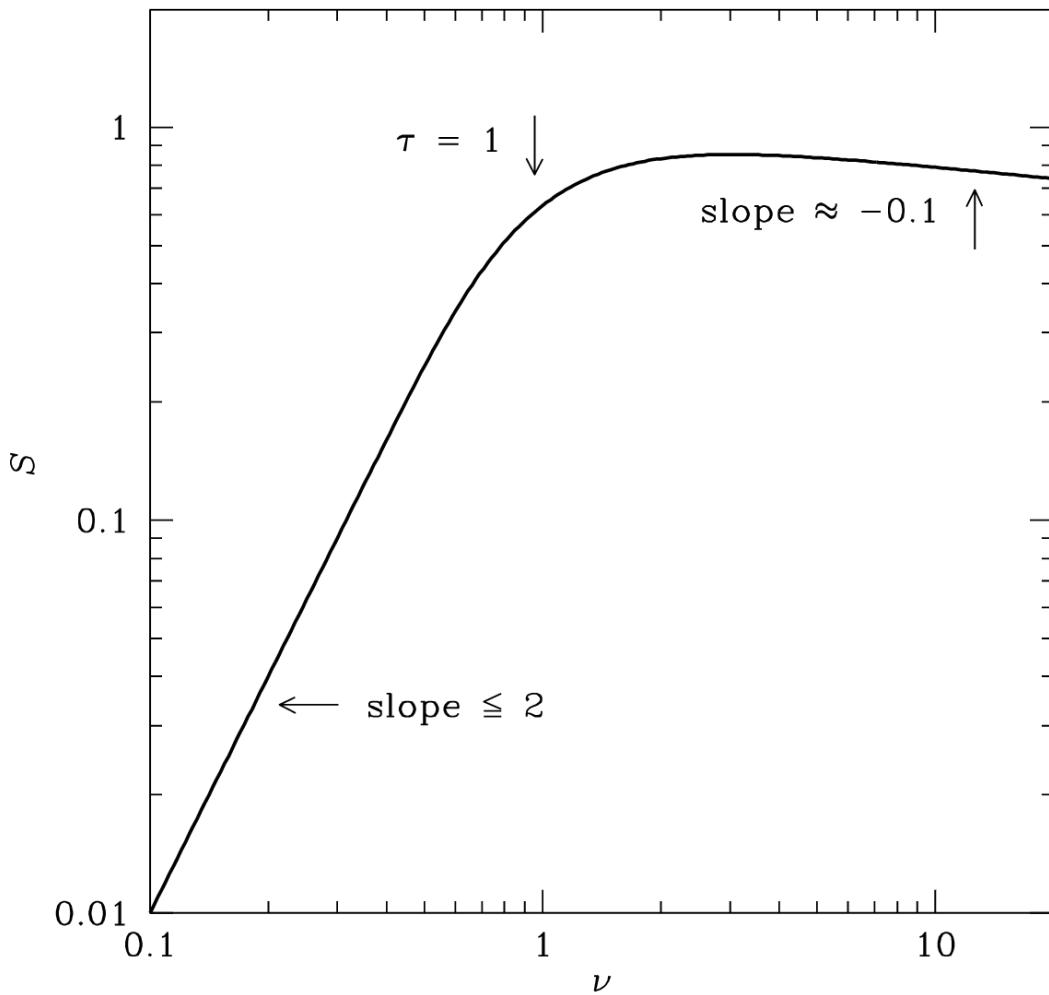
$$\alpha = -\frac{d \log S_\nu}{d \log \nu}. \quad (3.9)$$

The self-absorbed portion of the spectrum has an increasing power-law spectral index of $\alpha < 5/2$. The optically-thin portion of the spectrum has a decreasing power law spectral index that depends on the amount of relativistic beaming, characterized by the Doppler boosting factor δ . The power law index of this portion of the spectrum varies widely, with typically-observed values ranging from a flat spectrum ($\alpha \sim 0$) for AGN, to $\alpha \sim -0.7$ for radio galaxies and $\alpha \sim -2$ for pulsars. Typical turnover frequencies are too low to be detected, but some sources like quasars and pulsars can have turnover frequencies of a few hundred MHz - 1 GHz.

3.2.2 Free-Free Radiation

This was discussed in good detail in **Draine Ch. 10**, and we summarized the theoretical expectation for a thermal plasma already in this chapter, but I want to also highlight some of the key observational properties implied by all those equations.

Similar to what we just discussed for synchrotron radiation, a typical free-free spectrum is characterized by both an optically-thick (steeply rising) and an optically-thin (flat or slightly decreasing) component of the spectral energy distribution. For sources with large electron densities, the turnover frequency can be pushed to frequencies of tens of GHz. Below, I summarize physical and observational properties of different types of HII regions

**Figure 3.3:** freefree.png

Giant HII regions Relatively short-lived nebulae (ages up to a few million years), surrounding OB associations or young massive clusters. The Orion Nebula and the Tarantula nebula (surrounding 30 Doradus) are classic examples of giant HII regions. These are typically the HII regions that you see in optical images of other galaxies.

Diameters: 10's - 100's of parsecs

Average n_e : $1 - 100 \text{ cm}^{-3}$

Emission measure: $100 - 10^5 \text{ pc cm}^{-6}$

Turnover frequency: $\lesssim 1 \text{ GHz}$

Compact HII regions Relatively short-lived nebulae (ages up to \sim a million years) that can be excited by a single O-type star.

Diameters: $\sim 0.1 - 1 \text{ pc}$

Average n_e : $\sim 10^3 \text{ cm}^{-3}$

Emission measure: $\sim 10^6 \text{ pc cm}^{-6}$

Turnover frequency: $\sim 10 \text{ GHz}$

Ultra/Hypercompact HII regions Generally the youngest sources (ages $< 10^5$ years), corresponding to very early stages of massive star formation.

Diameters: $\lesssim 0.05\text{-}0.1 \text{ pc}$

Average n_e : $\geq 10^4 - 10^5 \text{ cm}^{-3}$

Emission measure: $\geq 10^7 - 10^8 \text{ pc cm}^{-6}$

Turnover frequency: $\sim 15 - 50 \text{ GHz}$

3.2.3 Dust Emission

In fairness, Draine devotes MANY chapters to the microphysics of dust (we'll get there), and is one of the foremost experts on this subject. Here, I just want to briefly summarize the continuum radiation properties of a cloud of dust at some temperature T . Dust emission is generally described as a 'modified blackbody', where the main modification is a result of the fact that dust grains are small, and so are inefficient radiators at wavelengths larger than their sizes. This effectively steepens the Rayleigh-Jeans portion of the Planck distribution.

As a reminder, in the Rayleigh-Jeans regime, $I_\nu \propto \nu^\alpha$ (Equation 2.18). Optically-thin dust has a typical spectral index that is the Rayleigh-Jeans spectral index (2) plus the emissivity β (which depends on dust grain size and composition) and typically ranges between 1.5-2:

$$\alpha = 2 + \beta \sim 3.5 - 4$$

For optically-thick dust, the spectral index just approaches the Rayleigh-Jeans value of $\alpha = 2$.

The peak of the dust spectral energy distribution typically occurs around far-infrared wavelengths or THz frequencies, and the location of the peak can be used to determine the dust temperature. Dust emission generally becomes detectable at millimeter wavelengths, starting at frequencies as low as 30-40 GHz, and dominating other continuum emission processes at frequencies larger than 200 GHz.

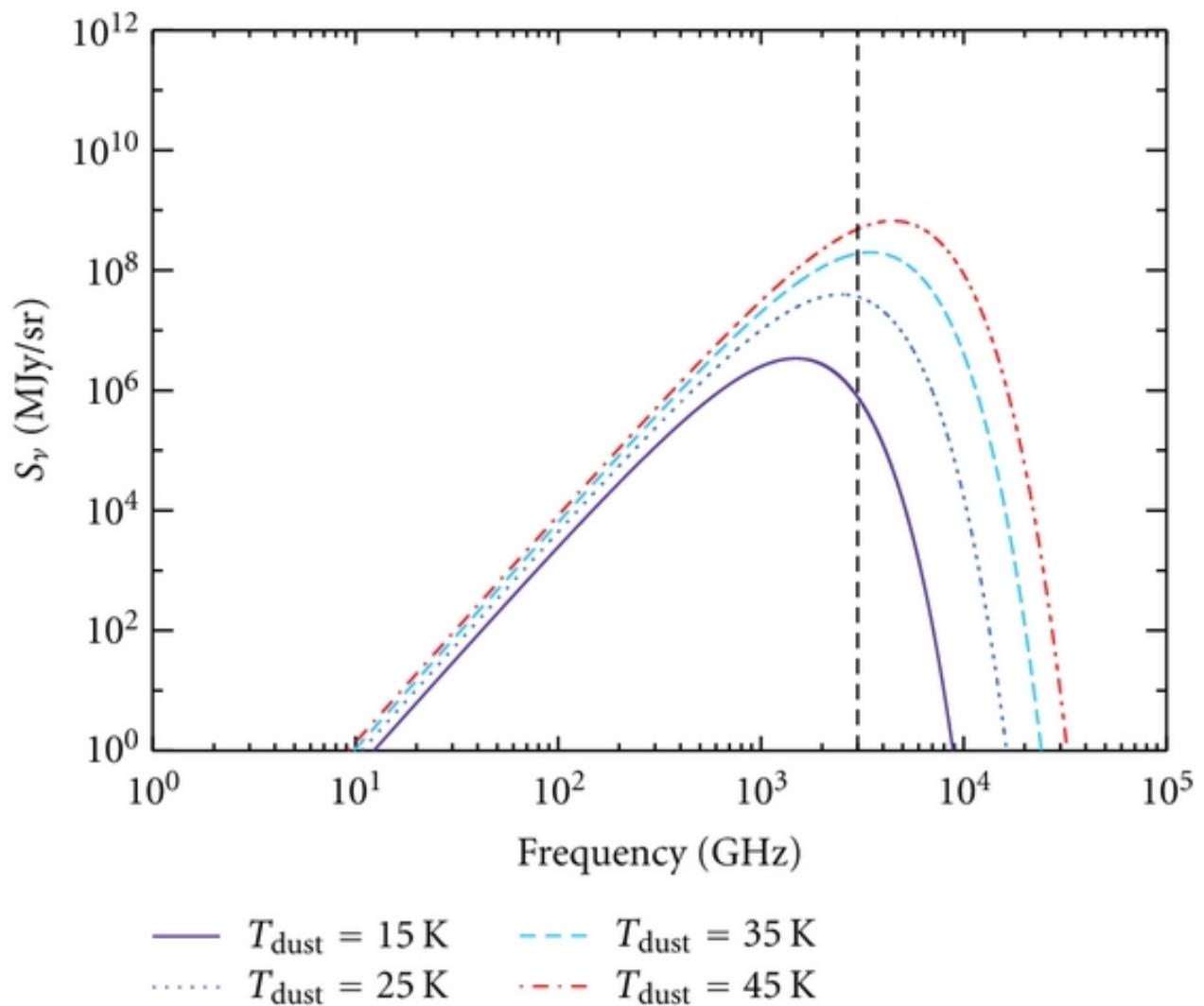


Figure 3.4: *dust.png*

Chapter 4

Atomic Structure

4.1 Electron Orbital Structure

4.1.1 Fermi Statistics

Atomic transitions will all involve the different quantum energy states of the electrons in an atom.

An important consideration in determining the arrangement of these states is that electrons are Fermions (particles having a half-integer spin, in this case, $s=1/2$, for a spin angular momentum of $\hbar/2$). Electrons then obey Fermi-Dirac statistics, which here basically just means they follow the Pauli Exclusion principle: no two electrons can occupy the same energy state of an atom.

4.1.2 Energy States

A typical way to visualize energy states (which we will see again when we talk later about rotational and vibrational energy states) is a diagram in which the y axis represents the energy (this can be shown schematically, with energy increasing with increasing y, or quantitatively, where the y axis has values in terms of eV or K). The x axis does not have a quantitative meaning, but instead is generally used to separate energy levels corresponding to different quantum numbers.

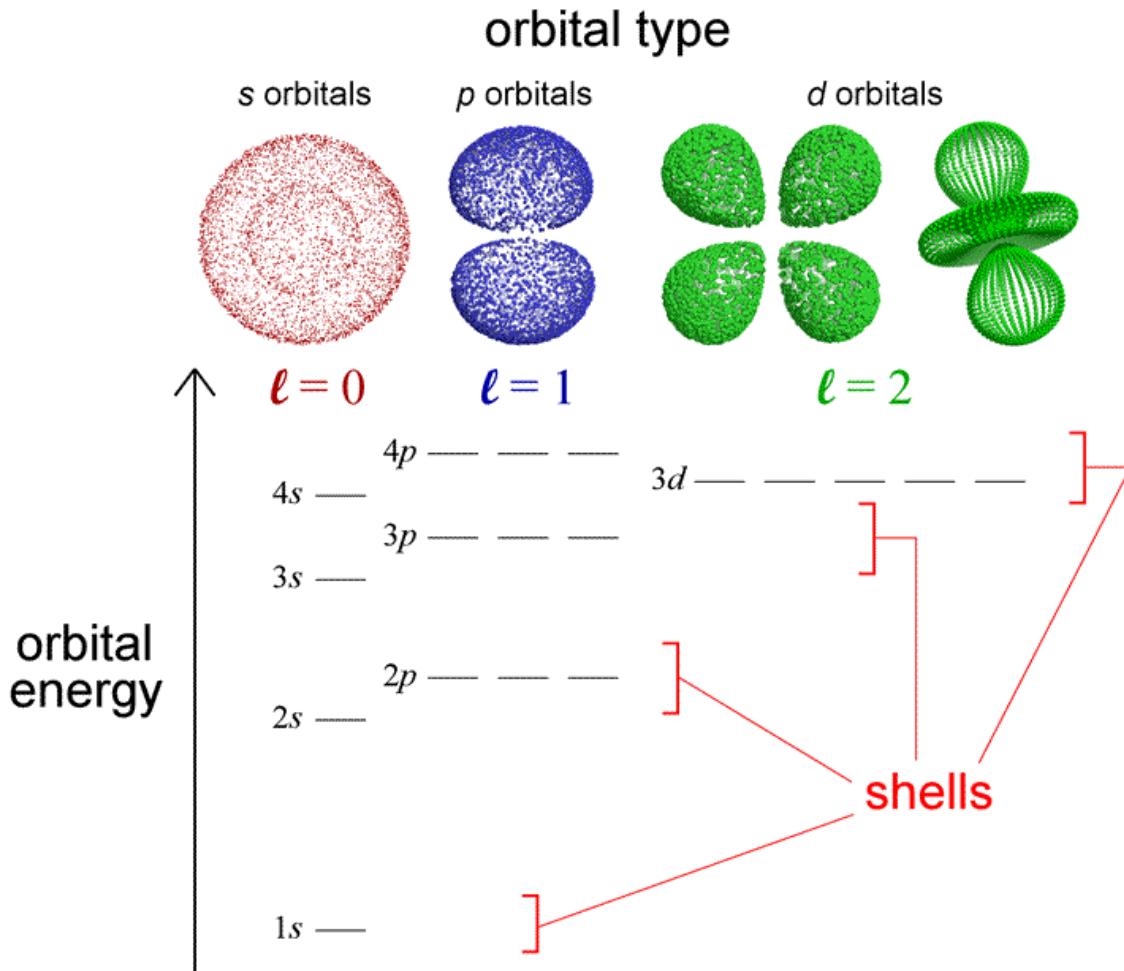


Figure 4.1: orbital_energy_and_type_diagram.png

In this figure, you can more clearly see the ordering of the electron orbitals by increasing energy, as given in the text: $1s$, $2s$, $2p$, $3s$, $3p$, $4s$, $4d$ and so on. In this diagram, each dashed line segment represents an electron pair: the s-orbitals can hold 2 electrons, the p-orbitals can hold 6 electrons, and the d-orbitals can hold 10 electrons.

4.1.3 Quantum Numbers

The orbitals in the figure above are organized by two of the principal quantum numbers we will deal with for atoms. Increasing up the y-axis we have n , the principal quantum number. n is determined by the electron wave function, and corresponds generally to the energy of the electron (bottom diagram) and its most likely distance from the nucleus (e.g., the size of the shell in the top image). The lowest value of n is 1, and it can (in theory though not in practice) be infinitely large.

Increasing to the right on the x-axis we have ℓ which is the orbital angular momentum quantum number. The orbital angular momentum of an electron is $l\hbar$ and this is different from the intrinsic spin angular momentum (which is also quantized, with a spin magnetic quantum number m_s , which can only take two values: $\pm \frac{1}{2}$, corresponding to the orientation of the electron spin relative to the spin of the nucleus). Note that s , p , and d are not actually quantum numbers, but are simply

ways of referring to the $l = 0$, $l = 1$, and $l = 2$ states. So, the $4p$ orbital has quantum numbers $n = 4$ and $l = 1$.

l can have values ranging from 0 to $n - 1$ (which is why, for example, there is no $1p$ state). l determines the number of subshells in an orbital (e.g., for $n = 1$, the only possible l value is 0, so there is just one subshell: $1s$. For $n = 2$, l can now take on the values of 0 and 1, so there are two subshells: $2s$ and $2p$).

The final quantum number, which sets the total number of independent quantum states in a subshell and thus the number of electrons in that subshell, corresponds to the projection of the angular momentum of the electron along the z-axis of the atom. This quantum number (denoted m_l , the magnetic quantum number) then has to do with the orientation of the electron relative to the atom— or basically, which suborbital the electron is located within. , which is usually most important when we are discussing a situation in which the atom is in a magnetic field. m_l takes values of $\pm l$.

Taking all of these quantum numbers together, we can finally define the number of electrons in a given subshell l : $2(2l + 1)$. Checking this: for $l = 0$, there is only one possible value of m_l . The two possible states in which an electron can exist are then set by its two possible intrinsic spin values, $m_s = \pm \frac{1}{2}$. We also see that the number of electrons in a subshell does not depend on n , so the number of electrons in any s , p , or d subshell is always the same.

4.2 Spectroscopic Terms

Because the relative orientation of the intrinsic and orbital electron spins is important (vectors!), it isn't always enough to just specify an energy state by these quantum numbers. To deal with this, we introduce the additional concept of 'Terms', which are (quantum-mechanically-allowed) combinations of the total spin and orbital angular momentum of the whole-atom system.

The total spin angular momentum of an atom is the vector sum of the individual spin angular momenta of the electrons (each one having $\pm \frac{1}{2}\hbar$). This total value is denoted as $S\hbar$. Similarly, the total orbital angular momentum is given as $L\hbar$, the sum of all the $\pm l\hbar$ values for all electrons (as with the spin, the plus/minus convention corresponds to the direction of the orbit).

A term (L, S) is written as $^{2S+1}\mathcal{L}$.

Note that like we did for l , we denote specific values of the total orbital angular momentum quantum number with (now capital) letters: $\mathcal{L} = S, P, D$ corresponds to $L = 0, 1, 2$

You may notice that a horrible thing happens here in Draine: the total spin quantum number S is incredibly and unfortunately similar to the $L = 0$ total angular momentum state, S . Such is our lot.

To illustrate how terms are constructed, let's consider the Hydrogen atom, a one-electron system. Depending on its energy, this single electron can occupy any of the orbitals we have defined so far. The lowest-energy state that it can occupy is the $n = 1$, $1s$ orbital. The total orbital angular momentum L is 0, and the total spin angular momentum S is just the spin of a single electron ($S = 1/2$). We can then write the term for this state as 2S . Note that the term by itself contains no information on the excitation of the atom (the n quantum number of the electron). As there is one term, this state is called a singlet (although actually, as we will find, this state has hyperfine

structure due to slight differences in the energies between different configurations of the spins of the electron and the proton).

4.3 Fine Structure (Spin-Orbit Coupling)

Of course things get a lot fancier than the ground state of a single-electron atom.

Once we are considering both spin and orbital angular momentum, the total electronic angular momentum of an atom (J) is useful to define. J is the vector sum of L (all the orbital angular momentums) and S (all the spin angular momentums). This information can be added to our notation for a term, which is now written as $^{2S+1}\mathcal{L}_J$

We can use this to rewrite the term for the ground state of hydrogen. The orbital angular momentum of the electron is zero, so the net angular momentum J for the whole atom is just the total spin: $J = S = 1/2$. The full ground state of the hydrogen atom is then sometimes written as $^2S_{1/2}$.

A much more illustrative case of the role of J is the next-highest excited state of the Hydrogen atom is $n = 2$. We find that considering the placement of just a single electron in this orbital results in two different possibilities for the total angular momentum J . For an electron in the $2s$ orbital, we again have no orbital angular momentum and a net angular momentum $J = 1/2$. Like the ground state, the term for this state would be written as $^2S_{1/2}$. However, the electron could also be in the $2p$ orbital, in which case it would have $L = 1$. Depending on the orientation of its spin relative to its orbit, the total angular momentum could be $J = L + S = 3/2$ or $J = L - S = 1/2$. We thus find that, in total, there are three different terms for the $n = 2$ state: $^2S_{1/2}$, $^2P_{1/2}$ and $^2P_{3/2}$. As there are only two distinct values of J among these terms, this state is called a doublet.

This doublet represents not just two distinct angular momentum states of the hydrogen atom, but due to spin-orbit coupling it also represents two distinct energy states. This can be shown, with a bunch of tedious quantum mechanics, to be the result of the different interactions of the magnetic dipole of the electron (due to its intrinsic spin) as it orbits in the electric field of the nucleus. As a result, the $n = 2$ orbital energy level is split into two slightly different energies. The difference in wavelength of a photon emitted as an electron transitions between one or the other of these energy states can be observed with sufficiently high resolution spectroscopy.

The fine structure splitting of the $n = 2$ state of the hydrogen atom is shown in this figure. On the left are the energy levels one would expect with a simple Bohr ‘shell’ model of an atom (in which the electron does not have an assumed intrinsic spin). On the right is the observed splitting in a ‘Dirac’ model that includes the spin of the electron as an additional quantum number. The transition from $n = 2$ to $n = 1$ is known as the “Lyman α ” transition.

Note that in most (but not all!) astronomical situations, the Bohr approximation is sufficient to define the observed transitions: because the line splitting is so small, the fine structure or hyperfine structure is much less than the observed doppler broadening or pressure broadening of the line. We often think of a spectral line as arising from just a transition between different n orbitals.

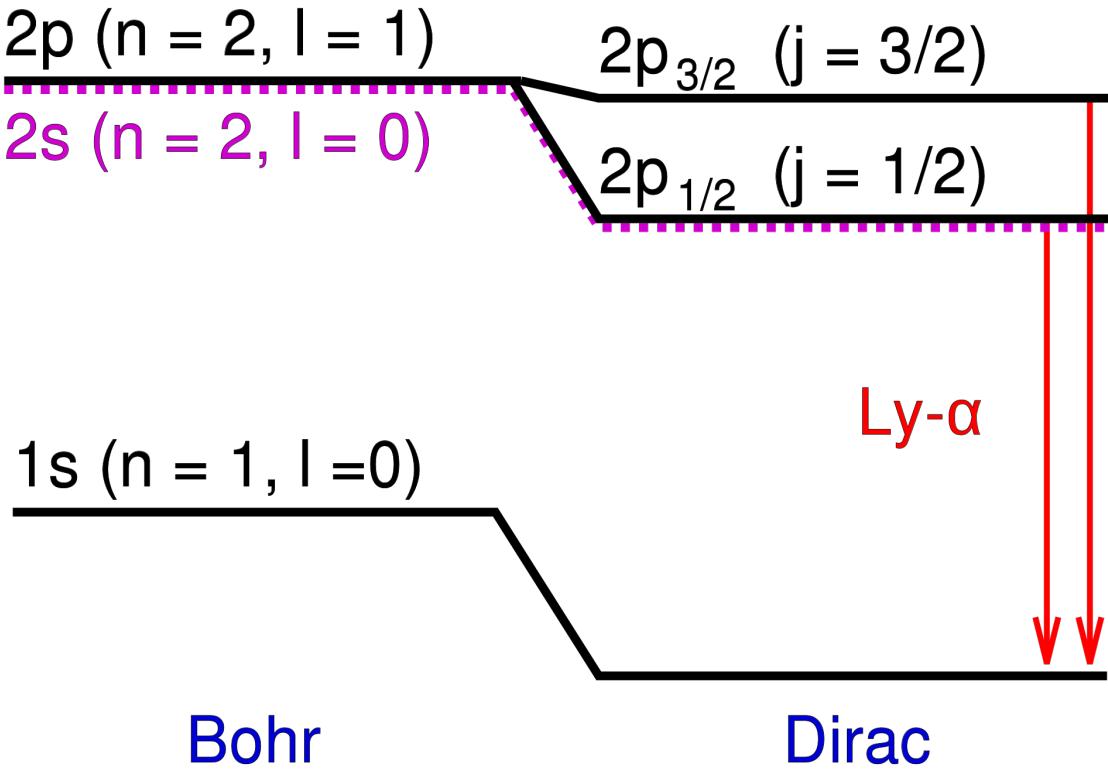


Figure 4.2: Hydrogen_fine_structure.png

4.4 Parity, Degeneracy, and Multiplicity (oh my)

Now that we have a way to describe the full angular momentum properties of a given configuration of electrons in an atom (which corresponds to an energy level or an excitation state) as a set of terms, we can turn to some different ways to classify these terms.

The parity of an energy level (which is determined by the symmetry of the wave function: whether or not it changes sign when reflected about the origin, see figure below) can be determined to be odd or even based on the sum of all of the l quantum numbers.

The parity is even if $\sum_i l_i$ is even

The parity is odd if $\sum_i l_i$ is odd

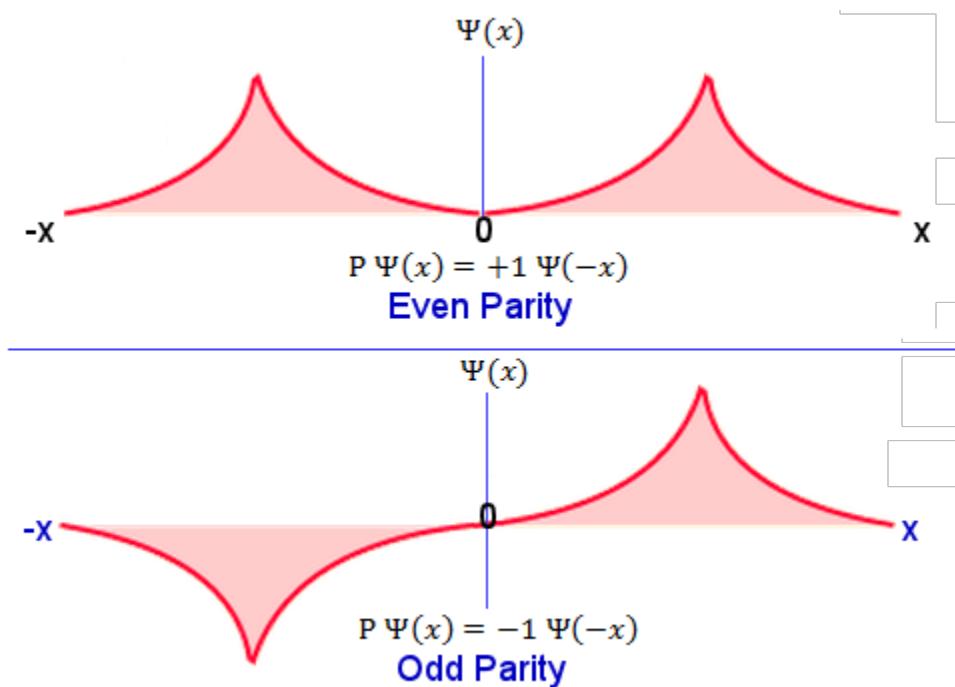


Figure 4.3: parity.png

Given all of the different possible combinations of L and S , especially for many-electron atoms, we also have to consider which of these are actually distinct energy levels— or conversely, for a given energy level, how many different quantum angular momentum states it encompasses.

The multiplicity describes the total number of different angular momentum states in an energy level. This is given by $g = (2S + 1) \times (2L + 1)$. For the 2P term of the $n = 2$ hydrogen orbital, with $S = 1/2$ and $l = 1$ we would calculate a multiplicity of 6.

The degeneracy, while a very similar concept, refers to the number of different angular momentum states which correspond to unique values of the total angular momentum ($J = 1/2$ and $J = 3/2$). This is basically the remaining multiplicity once you account for whether a line is a singlet, doublet, triplet, etc. due to the fine structure splitting resulting from spin-orbit coupling. Mathematically, the degeneracy of these states is given by $g = 2J + 1$.

For the fine structure splitting of the 2P term of the $n = 2$ hydrogen orbital, the degeneracies of the two doublet states are 2 (for $J = 1/2$) and 4 (for $J = 3/2$).

4.5 Electric Dipole Selection Rules and Forbidden Transitions

What is the likelihood that we can observe the transition of an electron from one energy state to another? Here, we are going to focus on radiative transitions, and specifically on the emission of a photon resulting from an electron transitioning from a higher energy state to a lower energy state. Note that there are other possible ways for the electron to transition between energy states (for example, increasing its energy due to the absorption of a photon of the correct energy, or moving up or down in energy due to the exchange of energy from the collision of an atom with another particles). We will discuss all of these in Section 6.2.

For now, it turns out there are a lot of rules that govern how likely (if at all) an electron is to move from one orbital to a lower one by way of emitting a photon.

Fundamentally, to understand these rules it helps to think of a state transition as the electron physically moving in the electromagnetic potential of the atom. There are then several relevant terms that can be defined, and we will have to dig a bit into quantum mechanics to really get to the meat of what is going on.

First, we need to characterize the distribution of charges in the atom. To do this, we define the dipole moment μ as $\mu = \sum_j q_j r_j$ where the q_j are all the charges in the system (electrons and nucleus) and the r_j are the position vectors, representing their relative location in the atom. The dipole moment describes the electrical polarity of the system. Note however that it is just an approximation: the dipole term is the lowest order term to consider: there are also higher-order perturbations (the quadrupole and even octopole moments).

Now, by moving the electron relative the the electric (magnetic) field of the atom, the total energy of the system changes. We can define the energy for a given state as

$$E = -\mu \cdot \epsilon \quad (4.1)$$

where ϵ is the electric field of the atom.

What we really want to characterize is the transition between two of these energy states: what is the dipole moment of this change? (we call this the transition dipole moment or μ_T). Because this is a quantum-mechanical and so probabilistic process, what we need to find is the expectation value of this quantity. We can start with the expectation value for the energy that it takes to transition between these two states, which we denote with an initial (ψ_i^*) and final (ψ_f) wave function:

$$\langle E \rangle = \int \psi_i^*(-\mu \cdot \epsilon) \psi_f dV \quad (4.2)$$

We will assume the electric field is constant on the size scale of the atom, which leaves us to evaluate

$$\mu_T = \langle \mu \rangle_T = \int \psi_i^*(\hat{\mu}) \psi_f dV \quad (4.3)$$

Long story short (a few matrix evaluations later) – if μ_T is nonzero, then the transition has an electric dipole moment, and is an "allowed" transition. However, if μ_T is zero, there is no electric dipole moment for the transition, and the transition is "forbidden" or "semiforbidden". This does not mean that no transition is possible! It just means that you would have to go to higher-order terms (the transition may have a nonzero magnetic dipole moment or electric quadrupole moment).

In general however such transitions are much less likely, and so statistically will take longer to occur. In terrestrial environments, it is likely that a collision wil de-excite an atom waiting to make this transition before it gets the chance to decay via radiating a photon, hence the initial surprise at seeing "forbidden" lines in the ISM: they only become viable in the low-density environment of space. In general, a good takeaway from all of this is that electric dipole transitions are the

strongest and most likely type of transitions to see (and we will return to this theme when we talk about rotational transitions in molecules).

All that wave function math can be a bit tedious, so we can use some patterns in the results to define a set of selection rules that tell us when electric dipole transitions can and cannot happen.

- (1) The parity of the initial and final state must be different
- (2) $\Delta L = 0$, or ± 1
- (3) $\Delta J = 0$, or ± 1 but $J = 0 - 0$ is not allowed
- (4) $\Delta l = \pm 1$ and only one single-electron wave function nl can change
- (5) $\Delta S = 0$: Spin does not change

Transitions that satisfy all of these selection rules are allowed, and are written with no brackets around the species name, e.g., NII ${}^3P_0 - {}^3D_1$ at 108.4 nm

A transition that only violates (5) is known as semi-forbidden (or intersystem). Such transitions are written with a single brackett, e.g., NII] ${}^3P_2 - {}^5S_2$ at 214.3 nm

Forbidden transitions violate at least one of rules (1)-(4), typically occurring via a magnetic dipole interaction (e.g., [NII] ${}^3P_1 - {}^1D_2$ at 654.99 nm) or electric quadrupole interaction (e.g., [NII] ${}^1D_2 - {}^1S_0$ at 575.62 nm).

(Consider also as an example of parity selection: the Lyman alpha doublet shown in the figure in an above cell. As a strong line, it must come from a transition between the $2p$ and $1s$ orbitals. Even though the $2s$ orbital has the same energy as one of the $2p$ terms, the transition $2s$ - $1s$ would not result in a change in parity, and so is not allowed by electron dipole selection rules).

4.6 Hyperfine Structure

If the nucleus of an atom has nonzero spin, then it has a magnetic moment. This results in additional splitting of those fine structure levels which also have a nonzero total spin. Essentially, the interaction of the magnetic moments is slightly different, depending on the directionality of the electron spin (whether it is parallel or antiparallel to the spin of the nucleus). The difference in orientation results in an (extremely small!) energy difference, and the single “fine structure” energy level splits into several distinct “hyperfine” energy levels.

4.7 Applications

Why bother learning all of this? It turns out that atomic transitions are ubiquitous in the ISM, and radiation from these transitions is used for many important diagnostics of gas conditions. Many atomic lines like those from hydrogen ($H\alpha$) occur within the visible range of wavelengths, and have become key tracers of the ISM by virtue of their brightness and ease of observation. Atomic line radiation is also one of the most important mechanisms for radiatively cooling the ISM (especially the hot ISM). Finally, the quantum structure of atoms is also important when considering molecules. We will revisit many of these concepts (e.g., hyperfine structure) in a molecular context. We also will have to consider the population of the outermost electron orbital (valence electrons) when determining chemical reactivity. An example of this unique to the ISM is a class of molecules called radicals, which have unpaired valence electrons. These molecules

are extremely reactive, and in the high-density conditions of terrestrial chemistry, are extremely short-lived. However, in the low-density environment of interstellar space, these molecules can survive for long periods of times, and radicals like OH are commonly observed.

Chapter 5

Molecular Structure

5.1 Molecular Quantum Numbers

Unlike atoms, molecules have many more types of quantized energy states to consider. These are generally treated with classical approximations to the underlying quantum physics of the transitions between these states. Such a large number of different energy states also means we must regularly deal with combinations of excitations of different states (e.g., rovibrational transitions involving changes in the rotation energy of a molecule with nonzero vibrational energy) and sometimes we must deal with more complicated interactions (for example, resonances between states with similar overall energies).

5.1.1 Electronic Transitions

Like atoms, molecules also have different electronic energy states. These can be quite involved, as molecules have complex electron configurations, some of which are involved in the molecular bonds themselves. (The most common type of bond you will encounter in the ISM is a covalent bond, in which the two atoms involved share valence electrons).

When considering electronic transitions, a useful approximation is to assume that the nuclei are fixed and the electrons move around in the Coulomb potential of the molecule (The Born-Oppenheimer approximation). Unlike in atoms, this potential is not spherically symmetric, but in a diatomic molecule it is at least symmetric when rotated about the axis that runs through the centers of the two atoms (known as the intranuclear axis)

Two representative quantum numbers for the electronic energy states of molecules are

1. The projection of the total electron spin on the intranuclear axis divided by \hbar : $S_{e,z}$ ($\Sigma = |S_{e,z}|$)
2. The projection of the total electron angular momentum (spin and orbital) on the intranuclear axis, also divided by \hbar : $J_{e,z}$ ($\Lambda = |J_{e,z}|$)

Like atoms, molecules can also have fine and hyperfine structure in their energy levels, with this small energy splitting being due to (1) the presence of multiple angular momentum Λ -states and (2) interactions between the nuclear magnetic moment and the magnetic field of the electrons.

5.1.2 Rotational Transitions

As unlike atoms, molecules are not spherically symmetric, we can also meaningfully describe rotation about different axes of a molecule, which will be quantized in terms of its angular momentum.

Rotational energies and angular momenta are generally approximated as a classical rigid rotor, in which the intranuclear separation is assumed to be constant. In fact of course, this approximation breaks down at high energies, when the centrifugal forces of rotation stretch the bond, and when the rotation can also be coupled with vibrational modes, in which the bond length is changing.

However, this approximation is generally good for the molecular ISM ($T_{kin} <$ a few hundred K).

The kinetic energy of a rigid rotor is given as

$$E_{rot} = \frac{J(J+1)\hbar^2}{2I} \quad (5.1)$$

where J is the rotational angular momentum and I is the rotational inertia:

$$I = m_r r_0^2 \quad (5.2)$$

where m_r is the reduced mass:

$$\left(\frac{m_1 m_2}{m_1 + m_2} \right) \quad (5.3)$$

and r_0 is the intranuclear separation.

This is often parametrized using a quantity B_0 known as the rotational constant for a molecule:

$$B_0 = \frac{\hbar^2}{2m_r r_0^2} \quad (5.4)$$

so that the rotation energy for a linear diatomic or polyatomic molecule can be written as:

$$E_{rot} = B_0 J(J+1) \quad (5.5)$$

(Note that B_0 is technically the rotational constant for the $v = 0$ vibrational state, and will change based on changes in the mean intranuclear separation for higher vibrational states).

5.1.3 Vibrational Transitions

The simplest vibrational transition is a stretching mode, in which the intranuclear separation varies over values of r_n like a classical harmonic oscillator with a spring constant k that is related to the overall bond strength.

Classically, we can approximate this as movement in an electric potential V_q where

$$V_q(r_n) \equiv E_q^{(e)}(r_n) + Z_1 Z_2 \frac{e^2}{r_n} \quad (5.6)$$

where $E_q^{(e)}(r_n)$ are the eigenergies.

Expanding this potential around a minimum at $r_n = r_0$ we can write

$$V_q(r_n) \sim V_q(r_0) + (1/2)k(r_n - r_0)^2 \quad (5.7)$$

For small excursions r_n the solution to motion in this potential is a harmonic oscillator with angular frequency

$$\omega_0 = \left(\frac{k}{m_r} \right)^{1/2} \quad (5.8)$$

Putting this all together, we can then write the vibrational energy for the harmonic oscillator approximation for a diatomic molecule, where the energy is quantized based on the number of nodes v in the vibrational wave function:

$$E_{vib} = h\nu_0 \left(v + \frac{1}{2} \right) \quad (5.9)$$

where $\nu_0 = \omega_0 / 2\pi$

As with the rotational modes, the vibrational modes step up in energy from $v = 0$ (the fundamental vibrational mode) to higher “overtone” modes with $v > 0$.

Note that while k is related to the bond strength, variations in the vibrational frequency are primarily driven by differences in the reduced mass, with lighterweight molecules having a higher frequency for their fundamental vibrational mode.

Other types of vibrational transitions occur for more complex molecules; for a linear polyatomic molecule, you also have bending modes, in which the atoms of the molecule undergo out-of-plane motions.

5.1.4 Spectroscopic Terms

Just treating the major parts of the spectroscopic terms for diatomic molecules, these are written very similarly to what we saw in Section 4.2 for the spectroscopic terms of atomic energy levels:

$2\Sigma+1\mathcal{L}$

As before, the superscript describes the total electron spin (projected onto the intranuclear axis), and \mathcal{L} describes the total electron angular momentum (also projected onto the intranuclear axis).

Note that a HORRIBLE thing happens with the notation here, which seems to be an all-too-regular occurrence.

As with atoms, \mathcal{L} takes on letter values signifying the different angular momentum states Λ . Thus, $\mathcal{L} = \Sigma, \Pi, \Delta$ for $\Lambda = 0, 1, 2, \dots$

This is NOT the same Σ as the quantum number for the spin!

For homonuclear diatomic molecules (e.g., H₂, N₂, O₂) the term has a subscript that indicates whether the electron configuration is symmetric (gerade, *g*) or antisymmetric (ungerade, *u*) under reflection through the molecule center of mass.

For heteronuclear diatomic molecules (e.g., HD, CO, ¹⁸O¹⁶O) the term has a subscript corresponding to the value of $J_{e,z}$.

Luckily, terms describing the energy states of molecules are less commonly used in molecular spectroscopy. However, there are a number of molecules that have important transitions involving different electronic states. These include OH, CH, and SO.

States with non-zero Λ (note that this includes the ground state of CH and OH) are subject to a phenomenon called “Lambda-doubling”, in which there is a slight difference in the energies for otherwise identical states which differ only in the direction of the projection of the angular momentum along the intranuclear axis.

OH is a good example of the complexity of energy states present just in a diatomic molecule. OH has a $^2\Pi$ ground state with non-zero angular momentum Λ . Depending on the coupling of the spin and orbital angular momentum of the electrons, the ground state has fine-structure splitting between states with $J_{e,z} = 1/2$ and $J_{e,z} = 3/2$. Adding in the spin of the *nucleus* in addition to the electron spin, each of these states actually have multiple different energy values corresponding to *total* angular momentum values of $|J_z| = 1/2, 3/2, 5/2, \dots$ (or $J_z = 3/2, 5/2, 7/2$ etc). Each of these states is doubly degenerate ($\Lambda - \text{doubling}$) and each of those states has hyperfine splitting. The whole mess is shown here:

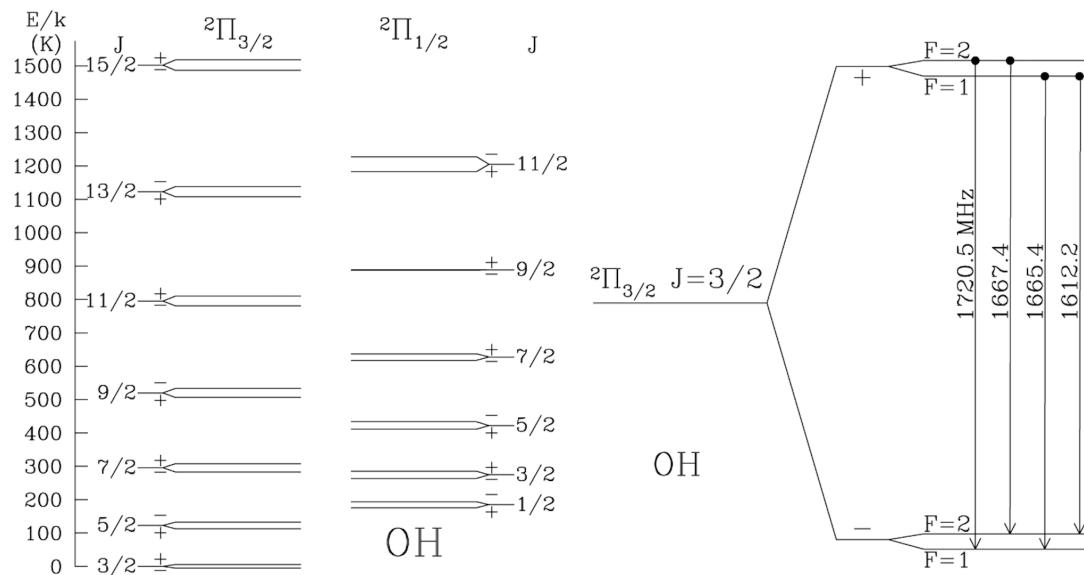


Figure 5.1: Energy levels of OH. Left: the rotational ladders of the $^2\Pi_{3/2}$ and $^2\Pi_{1/2}$. The splitting of the levels due to Λ -doubling has been exaggerated. Hyperfine splitting is not shown. Right: Λ -doubling and hyperfine splitting of the $^2\Pi_{3/2}(J = 3/2)$ state, showing the four 18-cm lines. In a magnetic field, each of these four lines is further split by the Zeeman effect.(Draine figure 5.3)

5.2 Symmetric Top molecules

We are now moving beyond the realm of diatomic and linear molecules into more complicated geometries. Below, I attempt to give as full a description of these topics as I can without actually getting into the quantum mechanical formalism of the Hamiltonians, array diagonalizations, Eigenvalues, etc. As with many topics in this class, there are often several widely-used or accepted conventions for expressing these quantities. I attempt to be as self-consistent as possible, but note that you may encounter different notations in different sources (particularly when comparing chemistry vs. physics sources).

When considering rotation, we stick to our rigid rotor approximation. Of course this is only an approximation: for accurate scientific applications, one has to include a term for the centrifugal distortion of the molecule. Ignoring this, we define three primary axes of rotation: \hat{A} , \hat{B} , and \hat{C} . Using this notation, we can redefine the rotational kinetic energy of the molecule as:

$$E_{rot} = \frac{J_A^2 \hbar^2}{2I_A} + \frac{J_B^2 \hbar^2}{2I_B} + \frac{J_C^2 \hbar^2}{2I_C} \quad (5.10)$$

Here, $J_A \hbar$, $J_B \hbar$, and $J_C \hbar$ represent the projection of the total (quantized) angular momentum $J\hbar$ onto each of the axes \hat{A} , \hat{B} , and \hat{C} . Classically, $J^2 = J_A^2 + J_B^2 + J_C^2$, but quantum mechanically, we substitute $J(J+1)$ for J^2 .

We can then define three rotation constants:

$$A_0 = \frac{\hbar^2}{2I_A} \quad B_0 = \frac{\hbar^2}{2I_B} \quad C_0 = \frac{\hbar^2}{2I_C} \quad (5.11)$$

So that

$$E_{rot} = J_A^2 A_0 + J_B^2 B_0 + J_C^2 C_0 \quad (5.12)$$

The general convention is to assign these axes to a molecule so that $I_A < I_B < I_C$ and $A_0 > B_0 > C_0$, (or so that A_0 is largest/ C_0 is smallest in cases where there are degenerate values of I), though this varies somewhat in the literature (Townes and Schawlow appear to follow the opposite convention, and Draine does not adhere to this convention at all.).

For linear molecules, there are two degenerate rotational axes (\hat{B} and \hat{C}), and a single axis along which rotation is not distinguishable (\hat{A} ; $I_A = 0$). Hence, we defined a single rotational moment of inertia I or I_B and a single rotational constant B_0 . With only one value for the moment of inertia, there is only one rotational quantum number, which is the total angular momentum quantum number J .

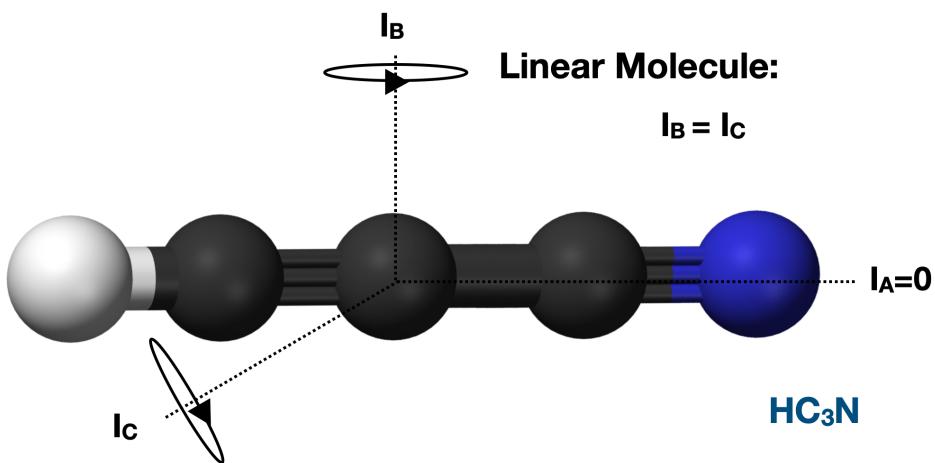


Figure 5.2: Rotational axes of a linear molecule.

The first step up in complexity is symmetric top molecules. These have a unique rotational axis and two degenerate rotational axes with identical rotational moments of inertia. To get the configuration of a symmetric top, you generally need at least 4 atoms in your molecule, however there is an astrophysically-important example of a 3-atom symmetric top: H_3^+ . Symmetric top molecules come in two flavors: oblate and prolate. You can think of these as ‘frisbees’ and ‘footballs’. Prolate molecules are much more common than oblate molecules.

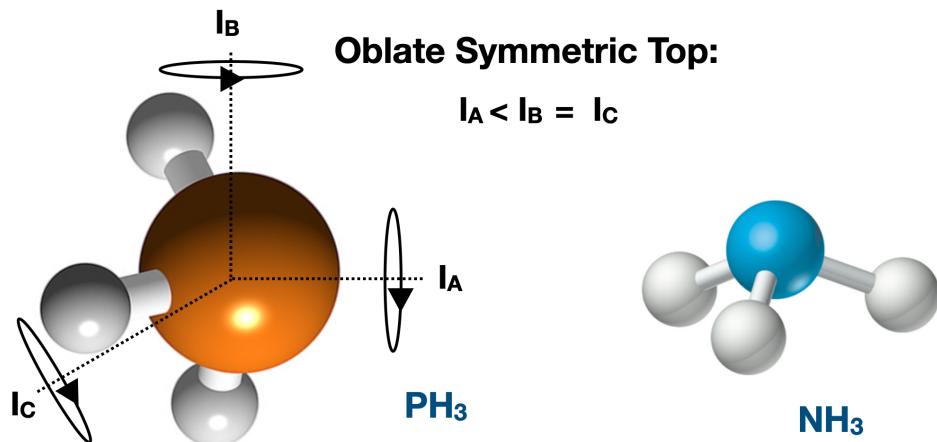


Figure 5.3: Rotational axes of an oblate symmetric top molecule.

For oblate molecules, the moment of rotational inertia of the axis with a unique value of I (defined as \hat{A}) is less than the moment of inertia of the other two axes (which have identical moments of inertia). For prolate molecules, the moment of rotational inertia of the single non-degenerate axis (now defined as \hat{C}) is larger than the two degenerate rotation axes. Examples of astrophysically-important oblate symmetric tops are ammonia (NH_3) and phosphine (PH_3). Examples of astrophysically-important prolate symmetric tops are methyl cyanide (CH_3CN) and methyl acetylene (CH_3CCH).

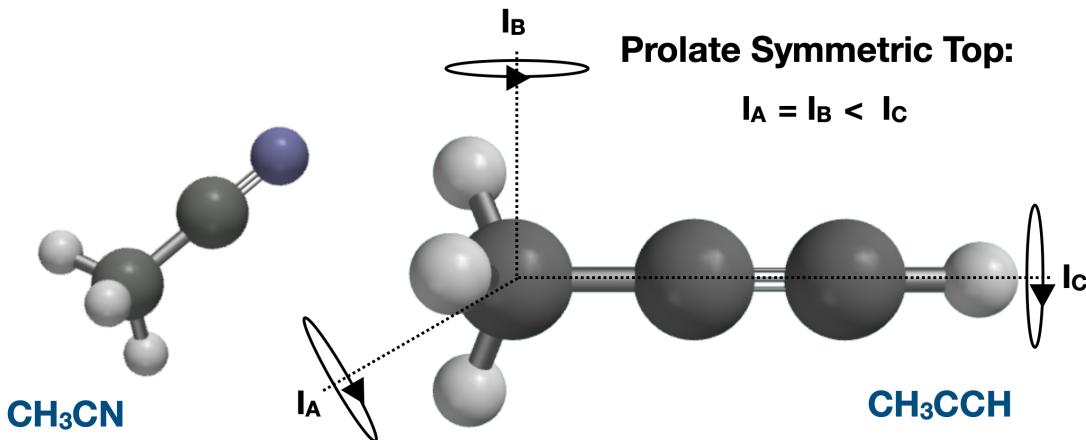


Figure 5.4: Rotational axes of a prolate symmetric top molecule.

For symmetric tops, the symmetry axis (the axis with a nondegenerate angular momentum) is either \hat{A} or \hat{C} , depending on whether the molecule is oblate or prolate. The angular momentum along the other two (degenerate) axes can then just be expressed as the difference between J^2 – again, actually $J(J+1)$ in a quantum mechanical formalism– and the square of its projection along the symmetry axis (we will call this variable $J_{A/C}$, to encompass both oblate and prolate symmetric tops):

$$E_{rot} = \frac{J_{A/C}^2 \hbar^2}{2I_{A/C}} + \frac{J(J+1)\hbar^2 - J_{A/C}^2 \hbar^2}{2I_B} \quad (5.13)$$

As there are two independent values of I for the symmetric top molecule, we then describe the rotational energy with two separate quantum numbers: J is still the total angular momentum quantum number, but now we also define $K = J_{A/C}$, where K can be physically understood as the projection of the total angular momentum onto the symmetry axis:

$$E_{rot} = \frac{J(J+1)\hbar^2}{2I_B} + K^2 \left(\frac{\hbar^2}{2I_{A/C}} - \frac{\hbar^2}{2I_B} \right) \quad (5.14)$$

For oblate symmetric tops:

$$E_{rot} = J(J+1)B_0 + K^2 (A_0 - B_0) \quad (5.15)$$

For prolate symmetric tops:

$$E_{rot} = J(J+1)B_0 + K^2 (C_0 - B_0) \quad (5.16)$$

Since by definition $A_0 > B_0 > C_0$ this means that for an oblate symmetric top, energy increases with increasing K , while for a prolate symmetric top, energy increases with *decreasing* K .

For integer values of J , K can take any value from 0 to J . That is to say, there are $J+1$ different energy levels that have the same value of J . The quantum numbers specifying the rotational states of symmetric top molecules are generally expressed in one of several ways: either (J, K) , $J(K)$, or

J_K . Thus, you may see an observed rotational transition written as $(4, 3) - (3, 3)$, $4(3) - 3(3)$ or as $4_3 - 3_3$.

Note that when $J = K$, the rotation axis of the molecule is aligned with its symmetry axis. We will return to this point below when we discuss permanent electric dipole moments of molecules.

The general selection rules for radiative (dipole) transitions between different rotational energy levels of symmetric tops are:

For $K \neq 0$: * $\Delta J = 0, \pm 1$ * $\Delta K = 0$

For $K = 0$: * $\Delta J = \pm 1$ * $\Delta K = 0$

This is not actually the entirety of rotation-related quantum numbers that can be present: if there is an external magnetic field with direction \hat{M} , one can define a quantum number M which is the projection of J along the magnetic axis.

5.3 Ammonia (NH_3)

5.3.1 Inversion Transitions

In many ways, NH_3 is a prototypical (oblate) symmetric top. However, NH_3 also has some interesting quirks that are not typical of most symmetric tops. One of these is the existence of an unusual type of vibrational mode. This “Umbrella”-type vibration (Figure 5.5), in which the structure of the molecule inverts, is not itself too unusual. Essentially, the nitrogen atom changes position, going from above the plane of the hydrogen atoms to below, and vice versa.

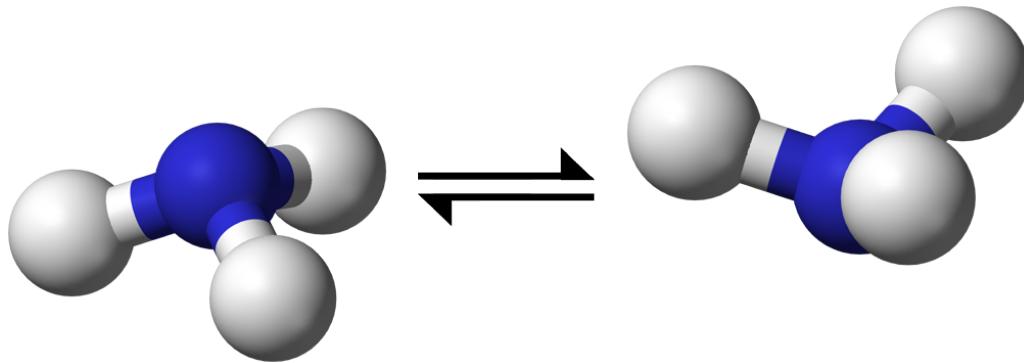


Figure 5.5: The ammonia inversion transition, in which the nitrogen atom moves through the plane of the hydrogen atoms as the pyramidal structure of the molecule inverts.

Where NH_3 is special (though not unique!) is that the energy required for the nitrogen atom to be in the plane of the three hydrogen atoms is larger than the energy of all but the highest vibrational modes (see Figure 5.6). Classically, the nitrogen could not cross this energy barrier, but quantum mechanically, there is a probability that the nitrogen atom will tunnel through this barrier.

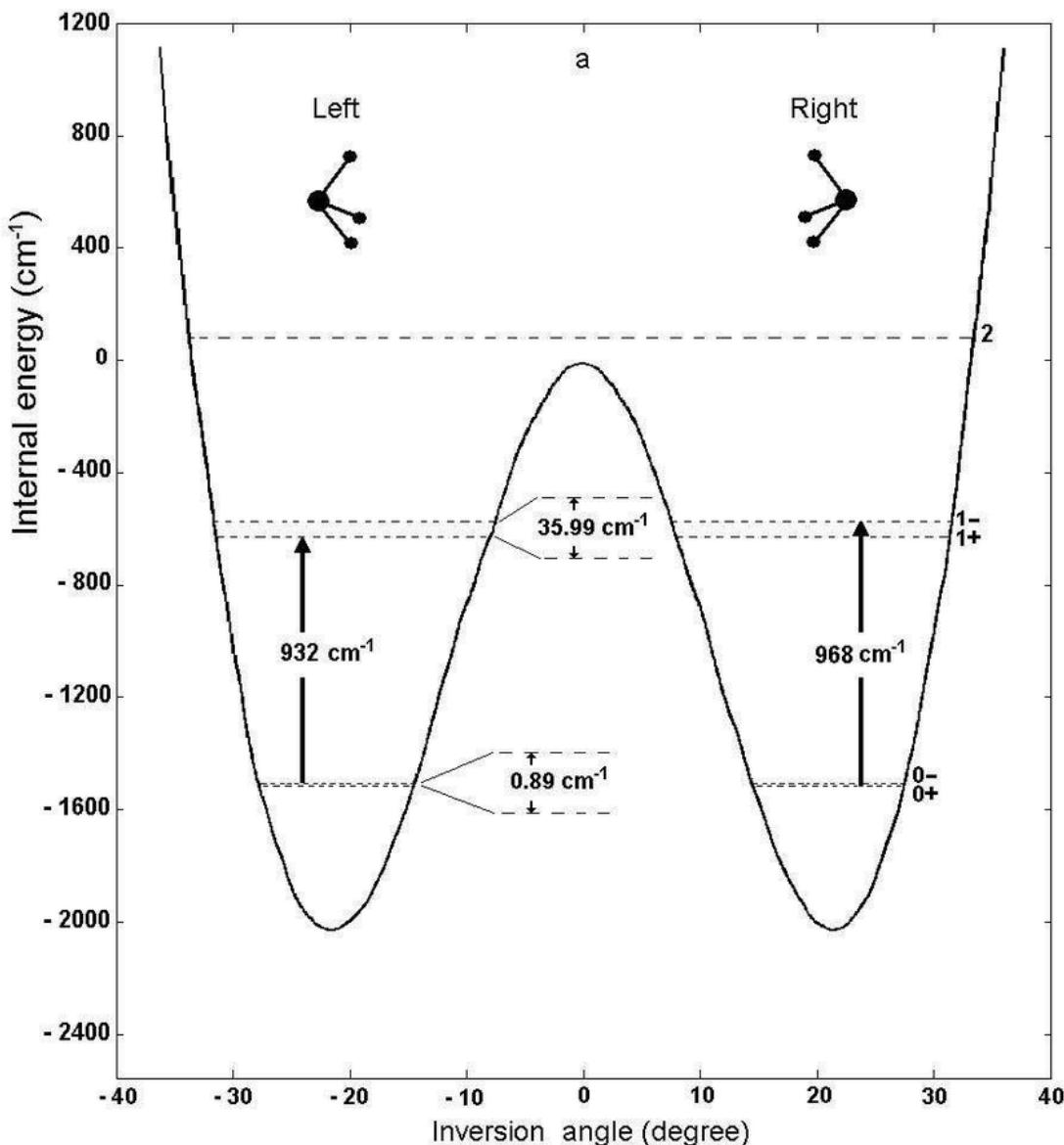


Figure 5.6: An internal energy diagram for ammonia, illustrating the potential energy barrier of the inversion transition.

The requirement that the nitrogen atom must tunnel through the plane of the hydrogen atoms in order for the molecule to fully invert effectively slows the vibrational frequency. While classically-allowed vibrations occur in the infrared (and indeed there are vibrational NH_3 transitions in the IR where the nitrogen changes its position slightly but the molecule does not fully invert) these inversion transitions of NH_3 occur at centimeter wavelengths, in the radio. This phenomenon – vibration slowed by a need to quantum-mechanically tunnel – is known as hindered vibration. It also manifests in some complex molecules like CH_3OH in which part of the molecule is free to rotate with respect to the rest of the molecule, but there are orientations for which there is an energy barrier.

For NH_3 , the observational upshot of this is the existence of inversion splitting in its (J, K) levels. As we said that these transitions occur at radio wavelengths, they have energy separations that are small to be seen in a typical energy-level diagram, though the splitting is often exaggerated in order to emphasize the doublet nature of these levels. These inversion transitions with $\Delta J = 0$ and $\Delta K = 0$ are some of the most commonly-observed transitions of NH_3 , as they can be used to probe the highly-populated $J = K$ states along the “Rotational Backbone” of the energy level diagram.

The $J = K$ levels are often called “metastable” levels due to their long lifetimes. They are also referred to as the “Rotational Backbone” (as seen in the energy level diagram for NH_3) since, as a result of their slow decay, most of the population of a symmetric top molecule like NH_3 will be found in these levels.

For ammonia, the $J > K$ transitions (correspondingly referred to as non-metastable transitions) have typical lifetimes of 10-100 seconds. In contrast, the $J = K$ metastable transitions have typical lifetimes of $10^7 - 10^8$ s, or a few years. A fun fact: for the $J = K$ states of NH_3 that make up the ‘Rotational backbone’, not only are electric dipole transitions forbidden, but electric quadrupole transitions out of these levels are also forbidden. The remaining allowed electric *octopole* transition has a lifetime greater than the age of the universe, and so can be safely ignored!

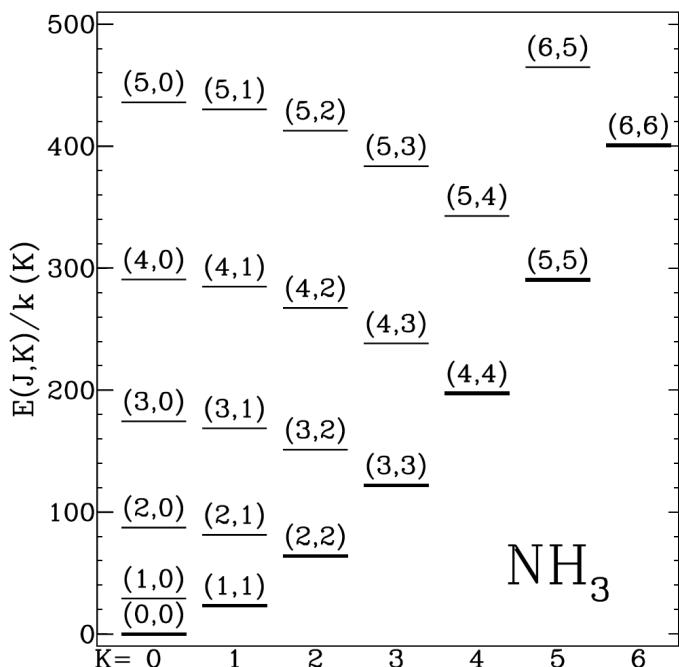


Figure 5.7: The rotational energy diagram of ammonia for $J \leq 6$

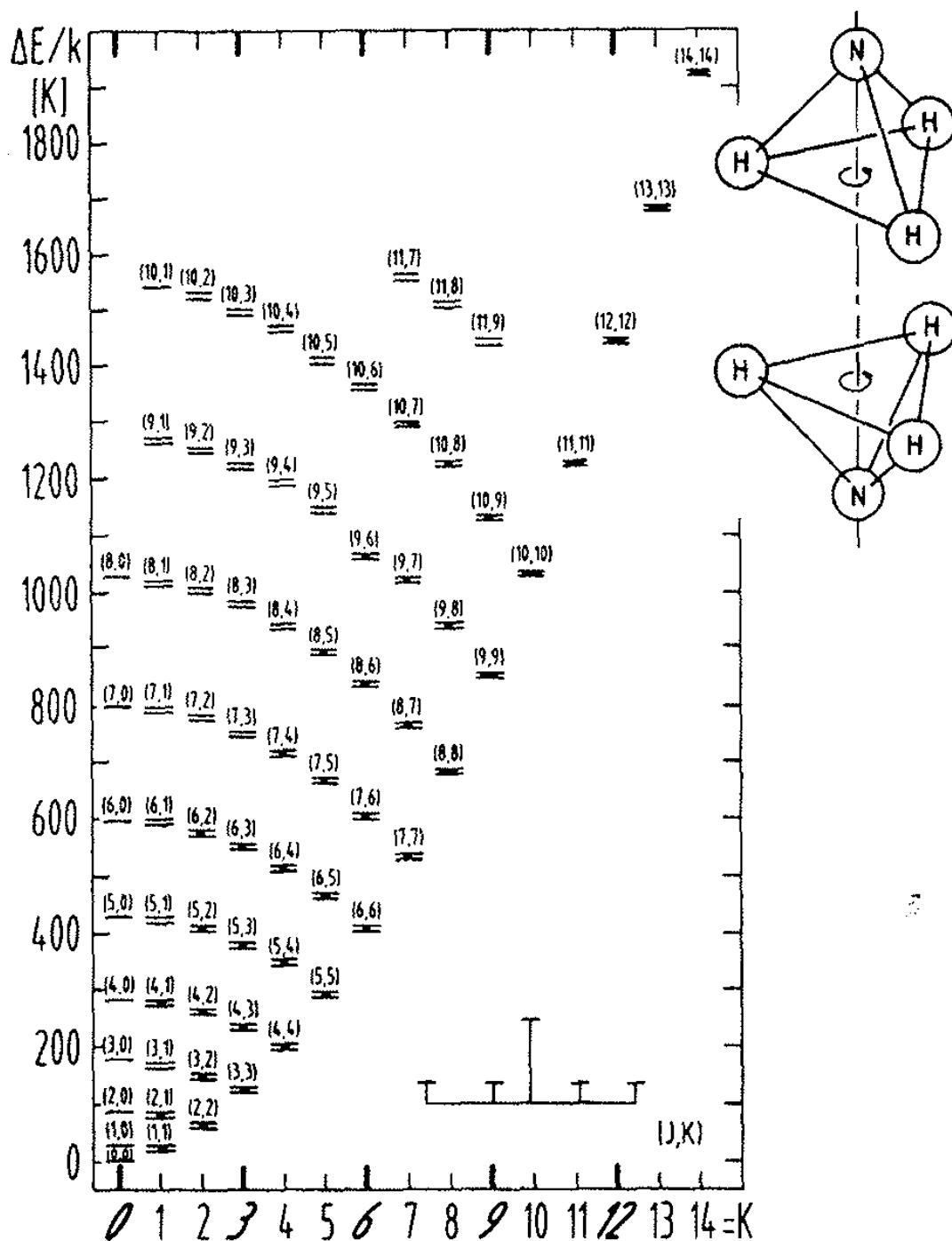


Figure 5.8: A more complete rotational manifold of ammonia, showing energy levels up to $J = 14$, as well as the hyperfine splitting of these levels.

5.3.2 Ortho- and Para- NH₃

Like many molecules (H₂, H₂O, CH₃OH) NH₃ has two forms that differ according to the spin alignment of the identical hydrogen nuclei that make up the structure of the molecule.

Ortho-NH₃ has the spin axes of all three protons aligned to be parallel with each other. Allowed

rotational states for ortho- NH_3 have $K = 3n$ (e.g., $K = 0, 3, 6, \dots$)

Para- NH_3 has the spin axes of two of the protons aligned, while the other is anti-aligned. Allowed rotational states of para- NH_3 have $K = 3n + 1$ or $K = 3n + 2$ (e.g., $K = 1, 2, 4, 5, \dots$)

5.3.3 Level Degeneracies

When considering the different rotational J -levels of NH_3 , the same rotational degeneracy applies as for diatomic and linear molecules:

$$g_J = 2J + 1 \quad (5.17)$$

There is an additional degeneracy of energy states that must be considered due to ortho- or para-symmetry of the molecule, as there are degenerate states due to the different (but indistinguishable) configuration of the anti-aligned proton spins.

This leads to a so-called ortho-para degeneracy where

$$g_{op} = 1 \quad (K = 3n)$$

$$g_{op} = 2 \quad (K = 3n + 1, 3n + 2)$$

5.3.4 Selection Rules

The general electric dipole radiation selection rules for symmetric tops are:

$$\Delta J = 0, \pm 1$$

$$\Delta K = 0$$

There are also additional election rules for collisional transitions of NH_3 due to the existence of effectively separate ortho/para species:

Transitions between all energy levels generally must preserve the ortho/para state of the molecule. While there are ways for a molecule to interchange its ortho/para states through magnetic interactions that precess the nuclear spin, these require excitation to highly-excited electronic/vibration states via interaction with a UV or X-ray photon or a cosmic ray. These interactions are quite unlikely to occur, especially in dense gas clouds. It is generally thought that the only efficient way to change the spin alignment of the protons is to destroy the molecule and re-form it.

This means that all transitions between energy levels—both through radiation and collisions—must be between levels with $K = 3n$ (ortho- NH_3) or $K = 3n + 1, 3n + 2$ (para- NH_3).

As was mentioned before, the only allowed transitions between ortho- or para- $J = K$ levels in the “rotational backbone” are prohibitively slow (electric octopole). Thus collisions are the primary means of redistributing the population between the different “K-ladders” of NH_3 . Because the relative populations of the different K states are almost entirely set by collisions, this makes ammonia an excellent probe of the kinetic gas temperatures.

5.3.5 The Ammonia Thermometer

NH_3 then has many properties that make it an optimal probe of gas temperatures.

As a simple hydride, it is relatively abundant in the gas phase. While many molecules have significant populations across a range of complex energy states (requiring nasty things like a partition function to characterize this distribution), the bulk of the population of a symmetric top like NH₃ lies in the long-lived $J = K$ metastable states along the rotational backbone. This means that in general, one does not need to consider how the $J \neq K$ states are populated in order to measure the total number of NH₃ molecules with a given energy. And as NH₃ gets its energy primarily through collisions with other particles in the gas (mainly H₂), the distribution of NH₃ energies can be assumed to be representative of the typical energy distribution of the H₂ molecules. Thus, the excitation temperature of NH₃ is an excellent proxy for the true gas kinetic temperature.

Additional practical considerations that are just icing on the cake:

- The range of energies of the levels are a good match to typical gas temperatures
- The inversion transitions occur at radio wavelengths—easily observed from the ground (and actually lie in a protected frequency range)
- The inversion transitions are relatively closely-spaced, making it easy to observe multiple transitions simultaneously.
- Hyperfine structure in the inversion lines (due to magnetic dipole interactions) also allows direct measurement of optical depth.

5.4 Asymmetric Top Molecules

Asymmetric top molecules (sometimes just called asymmetric rotors) have three unique and non-degenerate axes of rotation ($I_A \neq I_B \neq I_C$).

The classical expression $J^2 = J_A^2 + J_B^2 + J_C^2$ remains true. Further, because there are three independent moments of inertia, the rotation of the molecule can be described by three quantum numbers. However, apart from the total angular momentum quantum number J , the other two quantum numbers used to specify the different rotational energy levels are not so simply defined (as Townes and Schawlow puts it, ‘There is no set of convenient quantum numbers which can specify the state and also have a simple physical meaning’). K is no longer a “good” quantum number, as the angular momentum is not constant along any direction for a rotating asymmetric molecule.

Instead, we define new quantum numbers that correspond to the $2J + 1$ different energy levels that exist for each J . These other two quantum numbers are commonly referred to as K_{-1} , and K_{+1} , though you may also sometimes see them written as K_C and K_A . Both notations reflect the so-called “limiting cases” for these two quantum numbers, which correspond to the projection of the angular momentum J along the \hat{C} axis (e.g., the case of a prolate top) and the projection of the angular momentum J along the \hat{A} axis (e.g., the case of an oblate top). These states are usually written with the K_n as subscripts next to J , e.g., 3_{03} . An alternative way of specifying the energy states is to write them as J_τ where τ takes values of $-J, 1 - J, \dots, 0, \dots, J - 1, J$. The two notations are related by $\tau = K_{-1} - K_{+1}$. The relationships between the energy levels of symmetric and asymmetric tops is shown in Figure 5.9.

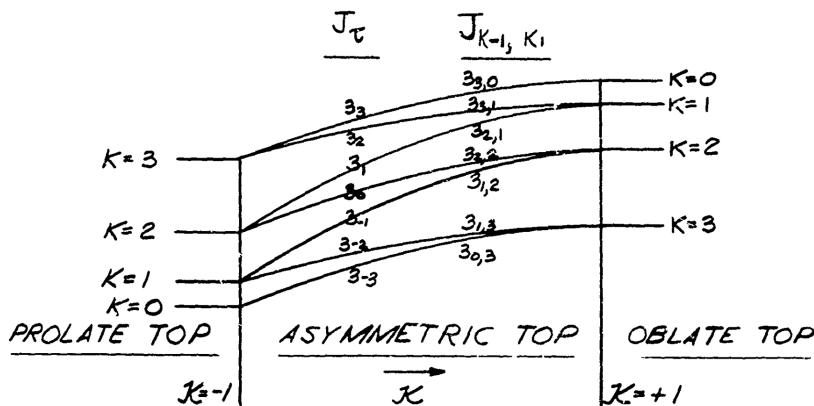


Fig. 1. Energy Levels of Symmetric and Asymmetric Top Molecules for $J = 3$.

Figure 5.9: The $J = 3$ energy levels for prolate, oblate, and asymmetric tops. This diagram is taken (believe it or not!) from a since-declassified 1962 report on the rotational energy levels of asymmetric top molecules, made by the research department of the Grumman aircraft engineering corporation.

Overall, asymmetric top molecules tend to have extremely complicated spectra, and there is no analytical expression for the different energy levels of the rotational states for these molecules. An approximation for these levels is sometimes written as:

$$E_{\text{rot}} = \frac{1}{2}J(J+1)(A_0 + C_0) + \frac{1}{2}(A_0 - C_0)E_\tau \quad (5.18)$$

Here E_τ is called the ‘reduced energy’, and is a function of both J and a parameterization of the asymmetry of the molecule κ :

$$\kappa = \frac{2B_0 - A_0 - C_0}{A_0 - C_0} \quad (5.19)$$

(Note that for a prolate symmetric top $\kappa = -1$ and for an oblate symmetric top $\kappa = 1$, hence the notation of the quantum numbers K_{-1} and K_{+1}).

However, E_τ is not a simple analytical function of these variables. Calculating values of E_τ must be done computationally, and can be extremely difficult. The easiest asymmetric top molecules to understand are those having symmetries that closely approximate symmetric tops (‘near-symmetric tops’).

For these, $I_A < I_B \simeq I_C$ (near-oblate tops) or $I_A \simeq I_B < I_C$ (near-prolate tops).

Astrophysically-important examples of these molecules are water (H_2O) and formaldehyde (H_2CO), both of which are near-prolate tops.

5.5 Permanent Electric Dipole Moment

We first discussed the concept of a dipole moment for atoms in Section 4.5 when we were quantifying the difference between allowed and forbidden transitions.

For both symmetric tops and linear molecules, the permanent electric dipole moment of the molecule is parallel to the symmetry axis.

Purely symmetric molecules (e.g., H₂) do not have a permanent electric dipole moment, and as a result they do not have allowed rotational dipole transitions. This includes so-called “spherical top” molecules like CH₄ that have equal rotational moments of inertia along all three orthogonal rotation axes.

The permanent electric dipole moment is generally written as μ and has units of Debye: this is a cgs unit equal to 10⁻¹⁸ statcoulomb-centimeters.

Examples of molecules with small electric dipole moments include

- CO ($\mu = 0.11$ D)
- CN ($\mu = 0.56$ D)

Examples of molecules with moderate electric dipole moments include

- NH₃ ($\mu = 1.46$ D)
- OH ($\mu = 1.67$ D)
- H₂O ($\mu = 1.85$ D)

Examples of molecules with large electric dipole moments include

- HCN ($\mu = 2.98$ D)
- HC₃N ($\mu = 3.72$ D)
- CH₃CN ($\mu = 3.9$ D)

In order to emit (dipole) radiation, the electric dipole must vary with time as a result of the molecule’s rotation. This naturally occurs for the rotation of linear molecules, and for $J > K$ states of symmetric top molecules, the dipole moment will precess about the axis of the fixed total angular momentum J . However, for the $J = K$ levels of symmetric tops, the rotation is occurring about the symmetry axis of the molecule. Electric dipole transitions out of these levels are then forbidden. In practice of course, they occur very slowly through magnetic or electric quadrupole transitions. These $J = K$ levels are often called “metastable” levels due to their long lifetimes. Overall, they are often referred to as the “Rotational Backbone” (as seen in the below energy level diagram for NH₃) since, as a result of their slow decay, most of the population of a symmetric top molecule like NH₃ will be found in these levels.

Chapter 6

Excitation Rates

6.1 Rate Coefficients

6.1.1 Radiative Transition Rates

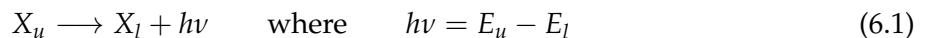
When we talk about changes between internal energy states, we will begin with those involving changes which we can directly observe: radiative transitions that result in the emission (or absorption) of a photon.

Einstein A Coefficient

For emission transitions, the likelihood that such a radiative transition actually occurs and we can observe it is given by the radiative transition probability, commonly referred to as the "Einstein A coefficient" for spontaneous emission.

In [Chapter 17](#), Draine writes this coefficient as A_{if} . As with our collisional rate coefficient, this is a coefficient that is specific to a single transition between two energy levels (going from E_i to E_f) of a particular species. However, as there is the additional requirement that $E_i > E_f$, you will often see this written as A_{ul} : a reminder that this is a transition from an upper to a lower energy level (See [Draine 6.1](#)). The units of Einstein A reflect that this is simply a characteristic rate (per time) at which a photon will be emitted (e.g., s^{-1}). Because of these units, we can define a characteristic timescale for spontaneously emitting a photon as $t_{em} = \frac{1}{A_{ul}}$. As an example, for ammonia, the $J = K$ metastable transitions have typical lifetimes of $10^7 - 10^8$ s, or a few years. In contrast, the $J > K$ transitions (correspondingly referred to as non-metastable transitions) have typical lifetimes of 10-100 seconds.

This probability of spontaneous emission depends on the strength of the permanent electric dipole moment μ . If there is no permanent electric dipole moment, there is no dipole radiation (and you will end up with a very large value of the Einstein A that corresponds to the much smaller strength of the magnetic or electric quadrupole moment of the molecule). We can write out the "reaction" of spontaneously emitting a photon for a species X as:



We can also use the Einstein A to start writing an expression (a differential equation) for the rate at

which the population X in the lower level E_l (or the upper level E_u) changes as a function of time. Solely due to spontaneous emission, we have:

$$\left(\frac{dn_l}{dt} \right) = - \left(\frac{dn_u}{dt} \right) = n_u A_{ul} \quad (6.2)$$

For allowed rotational dipole transitions of linear molecules, we can write an expression for the Einstein A that is a function of μ , J , and the rotational constant B_0 (For reference, a quick way to determine B_0 for a linear molecule is to use $E_{rot} = J(J+1)B_0$ and $E = h\nu$. B_0/h will be half the frequency of the $J = 1 - 0$ transition, which is generally an easy quantity to find in a database.):

$$A_{J \rightarrow J-1} = \frac{128\pi^3}{3\hbar} \left(\frac{B_0}{hc} \right)^3 \mu^2 \frac{J^4}{J + 1/2} \quad (6.3)$$

Note that this does assume a rigid rotor approximation for the rotation, so will become more inaccurate for higher values of J . However, we can still use this to calculate some good approximate Einstein A values for some rotational transitions of CO and HCN.

Example 1: CO

- $B_0/h = 57.64$ GHz
- $\mu = 0.11$ D

$$A_{1-0} = \frac{128\pi^3}{3(1.05 \times 10^{-27} \text{ dyne cm s})} \left(\frac{5.764 \times 10^{10} \text{ s}^{-1}}{2.998 \times 10^{10} \text{ cm s}^{-1}} \right)^3 (0.11 \times 10^{-18} \text{ dyne}^{1/2} \text{ cm}^2)^2 \left(\frac{1}{1.5} \right)$$

$$A_{1-0} = 6.88 \times 10^{-8} \text{ s}^{-1}$$

$$A_{4-3} = \frac{128\pi^3}{3(1.05 \times 10^{-27} \text{ dyne cm s})} \left(\frac{5.764 \times 10^{10} \text{ s}^{-1}}{2.998 \times 10^{10} \text{ cm s}^{-1}} \right)^3 (0.11 \times 10^{-18} \text{ dyne}^{1/2} \text{ cm}^2)^2 \left(\frac{256}{4.5} \right)$$

$$A_{4-3} = 5.87 \times 10^{-6} \text{ s}^{-1}$$

Example 2: HCN

- $B_0/h = 88.63$ GHz
- $\mu = 2.98$ D

$$A_{1-0} = \frac{128\pi^3}{3(1.05 \times 10^{-27} \text{ dyne cm s})} \left(\frac{8.863 \times 10^{10} \text{ s}^{-1}}{2.998 \times 10^{10} \text{ cm s}^{-1}} \right)^3 (2.98 \times 10^{-18} \text{ dyne}^{1/2} \text{ cm}^2)^2 \left(\frac{1}{1.5} \right)$$

$$A_{1-0} = 1.92 \times 10^{-4} \text{ s}^{-1}$$

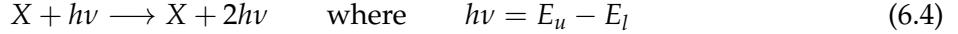
$$A_{4-3} = \frac{128\pi^3}{3(1.05 \times 10^{-27} \text{ dyne cm s})} \left(\frac{8.863 \times 10^{10} \text{ s}^{-1}}{2.998 \times 10^{10} \text{ cm s}^{-1}} \right)^3 (2.98 \times 10^{-18} \text{ dyne}^{1/2} \text{ cm}^2)^2 \left(\frac{256}{4.5} \right)$$

$$A_{4-3} = 1.64 \times 10^{-2} \text{ s}^{-1}$$

Einstein B Coefficients

There are two other relevant rates when considering radiative transitions (See **Draine 6.1**):

The first is the **stimulated emission rate** (the Einstein B stimulated emission coefficient, B_{ul}). Stimulated emission occurs if there are already a bunch of photons present with the exact wavelength that separates the upper and lower levels of X , and not just the same wavelength, but also the same polarization and direction of wave propagation. This causes the following reaction:



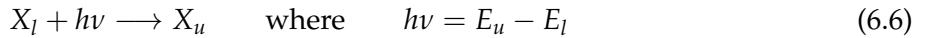
You might notice that this reaction 'amplifies' the amount of light emitted, and indeed stimulated emission is a key component of maser and laser transitions. Note that unlike A_{ul} , the stimulated emission rate depends on the density of appropriate photons. The units of B_{ul} are Volume/ Energy / time² (e.g., m³ J⁻¹ s⁻²) and B_{ul} thus represents the probability of emission per unit time per unit spectral radiance of the radiation field (page back to Chapter 2 for the full radiative transfer rundown and the definition of spectral radiance).

To fold B_{ul} into our equation for the rate of change of our level populations, we can then write the full rate equation for the *downward* radiative transitions as

$$\left(\frac{dn_l}{dt} \right)_{u \rightarrow l} = - \left(\frac{dn_u}{dt} \right)_{u \rightarrow l} = n_u (A_{ul} + B_{ul} u_\nu) \quad (6.5)$$

where u_ν is the radiation energy density per unit frequency (and u becomes yet another example of a symbol that has too many different contextually-dependent meanings.)

The second is the **absorption rate** (the Einstein B absorption coefficient, B_{lu}). This corresponds to the reaction



Once again, we have a rate that is dependent on the radiation field, specifically the density of photons with energy $h\nu$, so we can write the rate of change of the level populations due to upward transitions as:

$$\left(\frac{dn_u}{dt} \right)_{l \rightarrow u} = - \left(\frac{dn_l}{dt} \right)_{l \rightarrow u} = n_l B_{lu} u_\nu \quad (6.7)$$

Putting the upward and downward radiative transitions together, we have:

$$\frac{dn_u}{dt} = \left(\frac{dn_u}{dt} \right)_{l \rightarrow u} + \left(\frac{dn_u}{dt} \right)_{u \rightarrow l} = n_l B_{lu} u_v - n_u (A_{ul} + B_{ul} u_v) \quad (6.8)$$

For the case of a radiation field described by a blackbody (thermal radiation) we can write

$$u_v = \frac{4\pi}{c} B_v = \frac{8\pi h\nu^3}{c^3} \frac{1}{\exp(h\nu/kT) - 1} \quad (6.9)$$

Since this must be true in both the limits $T \rightarrow \infty$ and $T \rightarrow 0$, we can define some relationships between the Einstein A's and B's:

$$B_{ul} = \frac{c^3}{8\pi h\nu^3} A_{ul} \quad (6.10)$$

$$B_{lu} = \frac{g_u}{g_l} B_{ul} \quad (6.11)$$

6.1.2 Collisional Rate Coefficients

We now return to the concept of collisional rate coefficients. Yes, we have seen this before (see Section 1.5).

We will define collisional excitation as a collision that results in a change of the internal energy states of the species of interest. Note that as observers, we can only “see” evidence of collisional excitation through the subsequent radiative decay of the atom/molecule out of the excited level.

The change in internal energy due to collisional excitation (or de-excitation) can be described as a transition from energy state E_i to energy state E_f . Note that we are the more generic terms “initial” and “final” because this can be either a transition from an upper to a lower energy level, or vice versa. The collisional rate coefficient for the process is written k_{if} where

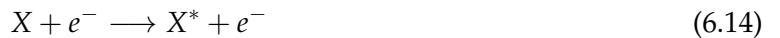
$$k_{if} \equiv \langle \sigma v \rangle_{i \rightarrow f} \quad (6.12)$$

As before, the units of a collisional rate coefficient are Volume / time (or for example, $\text{cm}^3 \text{s}^{-1}$).

A collisional excitation reaction for a given species X with some partner (for example, an electron) can be written as



or just



where the * indicates an arbitrary excited state.

We can express the rate at which the population of level f changes with time due to both collisional excitation and collisional de-excitation as

$$\frac{dn_f}{dt} = n_c(n_i k_{if} - n_f k_{fi}) \quad (6.15)$$

where n_c is the number density of the collision partner.

Unlike the Einstein A values, typical values of the collisional coefficients k_{ul} are largely the same for most molecules within order of magnitude, though they depend on temperature. For a temperature of 20 K, we can take computed values of the collisional coefficients from a database, finding

Example 1: CO

- $k_{1-0} = 3.249 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$
- $k_{4-3} = 5.091 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$

Example 2: HCN

- $k_{1-0} = 1.92 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$
- $k_{4-3} = 4.71 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$

6.2 The Two-Level Atom

6.2.1 Full Rate Equation

Incorporating collisional (de)excitation, we can make a full expression for the rate at which the level populations of a two-level system (atom or molecule) will change.

$$\frac{dn_u}{dt} = n_l(B_{lu}u_v + k_{lu}n_c) - n_u(A_{ul} + B_{ul}u_v + k_{ul}n_c) \quad (6.16)$$

Note that for symmetry with the Einstein A's and B's, you will sometimes see collisional (de)excitation rates written as C_{lu} and C_{ul} where

$$C_{lu} = k_{lu}n_c \quad \text{and} \quad C_{ul} = k_{ul}n_c \quad (6.17)$$

This then makes the units of C_{ul} and C_{lu} consistent with those of A_{ul} :

$$\frac{dn_u}{dt} = n_l(B_{lu}u_v + C_{lu}) - n_u(A_{ul} + B_{ul}u_v + C_{ul}) \quad (6.18)$$

Using this notation, we can visually express the rates in a two-level atom in the following plot.

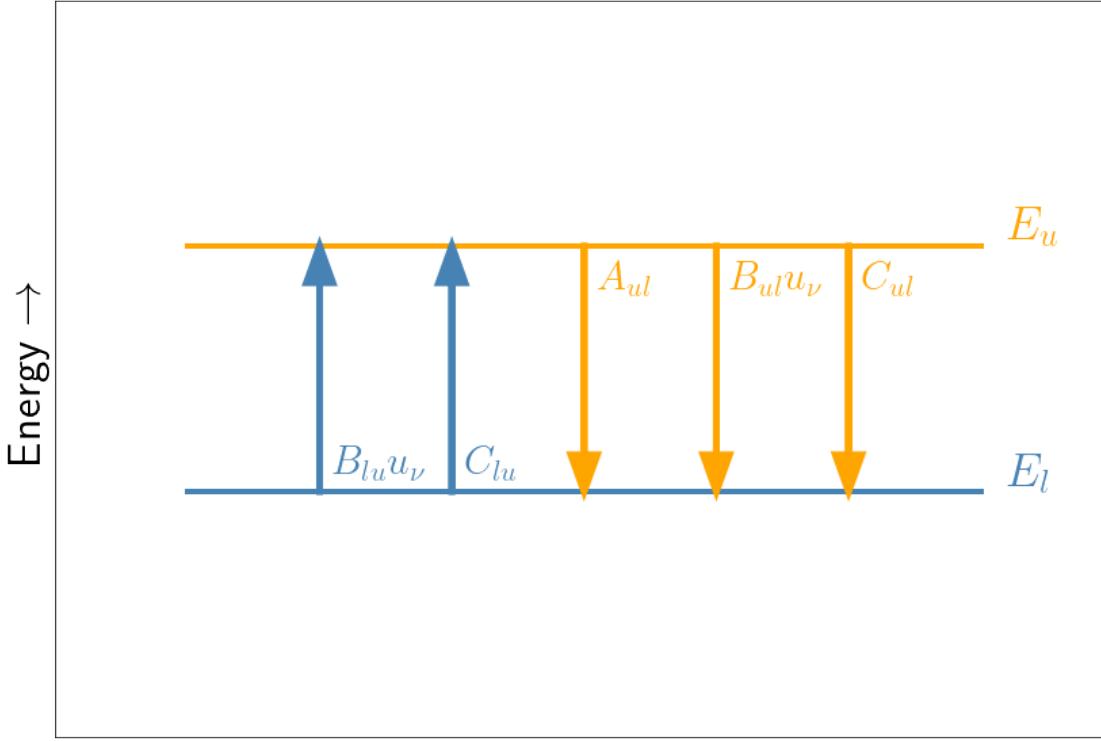


Figure 6.1: Schematic of the rates governing level populations in a two-level atom

6.2.2 Steady-State Solution

The condition for balance is:

$$\frac{dn_u}{dt} = 0 \quad (6.19)$$

$$n_l(B_{lu}u_\nu + k_{lu}n_c) = n_u(A_{ul} + B_{ul}u_\nu + k_{ul}n_c) \quad (6.20)$$

$$\frac{n_u}{n_l} = \frac{B_{lu}u_\nu + k_{lu}n_c}{A_{ul} + B_{ul}u_\nu + k_{ul}n_c} \quad (6.21)$$

Note that in **Draine Ch. 17** you will see this written in terms of just \$A_{ul}\$, \$k_{ul}\$, and \$k_{lu}\$, using the relationships derived between the Einstein A's and Einstein B's, as well as a quantity called the photon occupation number \$\bar{n}_\gamma\$ (averaged over all angles and polarizations), which collects some of the terms from the Blackbody equation:

$$\bar{n}_\gamma = \frac{c^3}{8\pi h\nu^3} u_\nu \quad (6.22)$$

6.3 Critical Densities and Thermalization

Together A_{ul} and k_{ul} determine the critical density of a molecule. The critical density n_{crit} is defined as the density of collision partners n_c for which the rate of spontaneous emission is equal to the rate of collisional de-excitation. For a two-level system, and ignoring for now the external radiation field, we can then express the critical density as

$$n_{crit} = \frac{A_{ul}}{k_{ul}} \quad (6.23)$$

You can verify that this indeed yields units of Volume⁻¹.

What does it mean

Given the relationship between the upward and downward collisional rate coefficients, which comes from the Boltzmann equation:

$$k_{ul} = \frac{g_u}{g_l} k_{lu} \exp\left(\frac{E_l - E_u}{kT_{kin}}\right) \quad (6.24)$$

the critical density for a given gas temperature is often interpreted by observers as the ballpark gas density required to observe emission from a given spectral line. This is because this is roughly where the rate of upward collisional excitation exceeds the rate of downward radiative excitation, which should populate the upper energy level sufficiently to have some reliable and detectable number of photons emitted through spontaneous emission.

In practice, this is a fraught interpretation of critical density that should not be used as a proxy for measuring gas densities, as it can overestimate the true gas densities by an order of magnitude or more.

But what does it really mean

Instead, the critical density is better understood as a measure of when species X approaches local thermodynamic equilibrium or “thermalization” with its collision partner. Below this density, the de-excitation of an atom or molecule from a given excited energy level is dominated by spontaneous emission of a photon. As a result, energy is removed from the gas, keeping species X at a temperature less than the kinetic temperature T_{kin} which is defined by the collision partners (usually a much more abundant species like H, H₂, or e⁻). In reality, even this is not always a reliable metric of gas thermalization, whether due to the presence of a significant radiation field or the simplistic assumption of a two-level atom (when in fact there are transitions between many other energy levels that must be considered).

As an example, **Draine Figure 17.2** shows that the H spin temperature only approaches the gas kinetic temperature for $n_H \gg n_{crit}$. In contrast, **Draine Figure 17.4** shows that the fine structure transition of C⁺ is much more easily thermalized at a density just above n_{crit} .

Calculating the critical densities for these transitions:

Example 1: CO

$$n_{crit,1-0} = \frac{6.88 \times 10^{-8}}{3.249 \times 10^{-11}} = 2.1 \times 10^3 \text{ cm}^{-3}$$

$$n_{crit,4-3} = \frac{5.87 \times 10^{-6}}{5.091 \times 10^{-11}} = 1.2 \times 10^5 \text{ cm}^{-3}$$

Example 2: HCN

$$n_{crit,1-0} = \frac{1.92 \times 10^{-4}}{1.92 \times 10^{-11}} = 1 \times 10^7 \text{ cm}^{-3}$$

$$n_{crit,4-3} = \frac{1.64 \times 10^{-2}}{4.71 \times 10^{-11}} = 3 \times 10^8 \text{ cm}^{-3}$$

A good rule of thumb is that molecules with large dipole moments will have large critical densities. Because they are very efficient radiators, such molecules tend to be slow to reach a thermal equilibrium with the bulk of the gas. They tend instead to be sub-thermally excited, and thus good probes of the overall gas density for densities $n \lesssim n_{crit}$.

6.4 Boltzmann Plots

For a species that is thermalized, its level populations will be described with Boltzmann statistics:

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} \exp\left(\frac{E_l - E_u}{kT_{kin}}\right) \quad (6.25)$$

Note that care must be taken to recognize the assumption that goes into this equation: what is actually observed is an excitation temperature T_{ex} , and only if the molecule is perfectly in LTE will its excitation temperature T_{ex} be equivalent to the kinetic temperature T_{kin} of the dominant collision partner. In general, a safe assumption is $T_{ex} < T_{kin}$. We will revisit this again when we discuss molecules, and discuss the properties that lead to a specific species being a good (easily thermalized) or bad (typically subthermally excited) thermometer for the ISM.

Observations of spectral lines also generally do not directly measure the volume density, but instead are sensitive to a quantity known as the column density N : the number of atoms or molecules per area element on the sky (e.g., atoms per cm^2). We can then re-write the observer's Boltzmann equation as:

$$\frac{N_u}{N_l} = \frac{g_u}{g_l} \exp\left(\frac{E_l - E_u}{kT_{ex}}\right) \quad (6.26)$$

Solving the Boltzmann equation for temperature yields

$$T_{ex} = -\Delta E \ln\left(\frac{N_l g_u}{N_u g_l}\right) \quad (6.27)$$

where the level energies E are expressed in units of K.

This is often visually represented in a Boltzmann plot of the degeneracy-corrected column densities of multiple different energy transitions of a species. A constant slope is an indication that

the species is near LTE, with the slope giving the temperature. Often one sees evidence of multiple temperature components in a Boltzmann plot, as the higher-excitation lines are progressively more sensitive to the presence of small amounts of hot gas.

MILLS & MORRIS

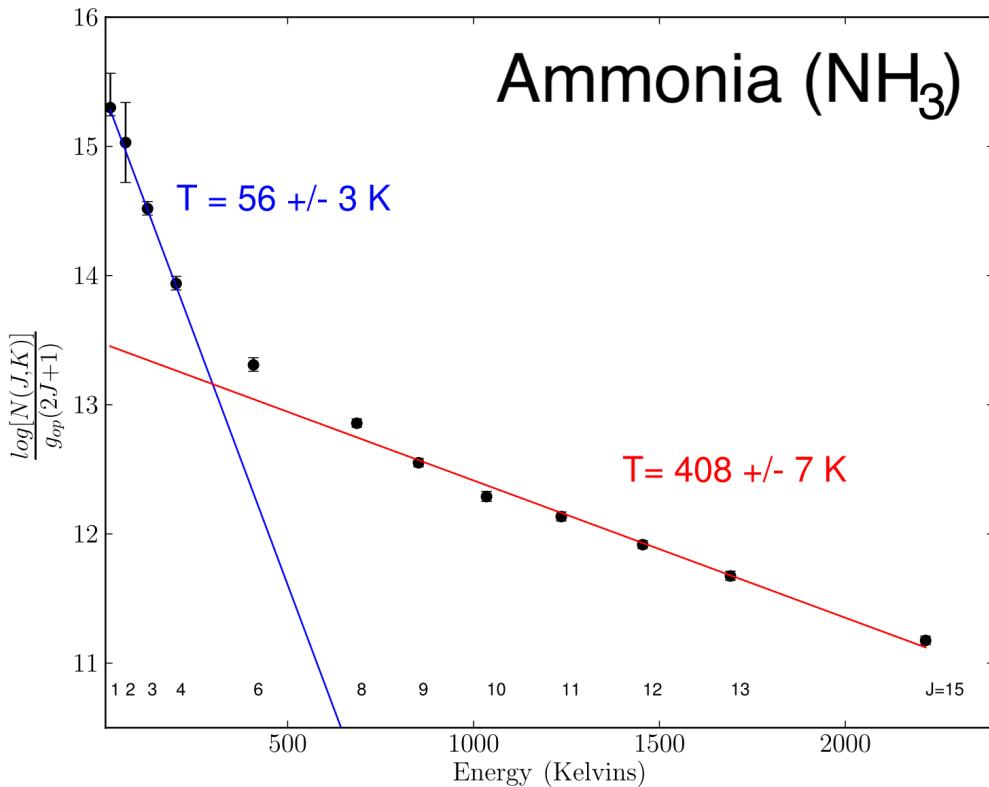


Figure 6.2: An example of a Boltzmann plot for a thermalized molecule (NH_3) which is used to measure two independent temperature components in the gas.

6.5 The Three-Level Atom

I won't say much about this other than that this is getting closer to the reality that must often be considered to accurately measure excitation properties (especially for molecules, where energy states are relatively more closely-spaced). For a multi-level system, the effects of non-LTE excitation must generally be considered computationally. This is also the case when we consider phenomena like radiative trapping (when we must consider not only the background radiation field, but also the likelihood that an emitted photon will be reabsorbed by a neighboring atom or molecule). We will talk more about these in Chapter 8.

I will also note that in general, a three-level system is often used to describe maser and laser emission. These usually occur via (radiative or collisional) "pumping" of a high-energy state which quickly decays down to an intermediate state. This results in a larger population of the intermediate state relative to a lower state, leading to a non-thermal relative population of these states (for thermally populated states, the lower state should have the larger population). This

population inversion is favorable to stimulated emission, and laser (or maser) emission then arises in the transition connecting the intermediate and lower energy states.

Chapter 7

Spectral Lines

7.1 Line Profiles

7.1.1 Natural Broadening (Lorentzian profile)

When we begin moving from a discussion of energy transitions in atoms to observed spectral lines, we must consider the different mechanisms that lead to the observed shapes (widths) of an absorption or emission feature. Classically, we might imagine that a spectral line resembles a delta function: an infinitesimally-narrow frequency range, for which photons of the precise energy corresponding to the difference between the lower and upper energy states of the transition are either emitted or are removed from some ambient radiation field.

However, we are dealing with quantum systems, and so we must take into account that the specification of the energy of these states is inherently uncertain: $\Delta E \Delta t \geq \hbar$

In this case, Δt corresponds to the “lifetime” of the energy state, or more specifically, to the typical lifetime of any given electron in that state. We have already defined the terms that affect these lifetimes: the Einstein A’s and B’s and the collisional C’s: the rates at which an energy state will be depopulated due to radiation or collision, via movement to a higher or lower energy state. Typically one rate will dominate (and often in the ISM, absent a strong background of continuum emission, this is the Einstein A). In this case, to determine the lifetime of level u given this population, one sums over all the Einstein A values for all of the permitted transitions out of this energy level: $\gamma_u = \sum_{E_j < E_u} A_{uj}$, and so the lifetime of the state can be characterized as $1/\gamma_u$. This would yield an “energy width” for this state of $\Delta E = \hbar\gamma_u$. For transitions, which involve two energy levels, we sum the gammas for both levels:

$$\gamma_{ul} = \gamma_{lu} = \sum_{E_j < E_u} A_{uj} + \sum_{E_j < E_l} A_{lj}.$$

The line shape that is created due to this intrinsic uncertainty in the energy states involved in a spectral line transition is described by a Lorentzian profile (note this is not exact, but is an excellent approximation):

$$\phi(\nu) = \frac{4 \gamma_{ul}}{16 \pi^2 (\nu - \nu_{ul})^2 + \gamma_{ul}^2}$$

7.1.2 Collisional (Pressure) Broadening

Pressure broadening is not typically important in the low-density regime of the ISM, however it is very important in environments like stellar atmospheres. The description of pressure broadening follows very closely to our definition of natural line broadening above. The difference now is that the frequency of collisions has increased so that collisional de-excitation is more likely than radiative excitation. The typical lifetime of an energy state is now not set by γ_{ul}/A_{ul} , but instead by the collisional rate coefficients. The resulting line profile is still Lorentzian in shape, but it is now broader, because a smaller level lifetime leads to a larger energy level width.

7.1.3 Doppler Broadening (Gaussian profile)

In most astrophysical environments, the emitting particles will be moving at large enough speeds that the Doppler shifts from the component of their motion along the line of sight will lead to a measurable broadening of the spectral profile. If the motions are thermal, then the resulting line profile is described by a Gaussian:

$$\phi(\nu) = \frac{1}{\sigma_\nu \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(\nu - \nu_{ul})^2}{\sigma_\nu^2}\right) \text{ where } \sigma_\nu^2 \text{ is the variance of this distribution, and is related to the full-width at half-maximum (FWHM) by}$$

$$\sigma_{\nu,\text{FWHM}} = 2\sqrt{2\ln(2)}\sigma_\nu$$

7.1.4 Combined effects (Voight profile)

Most correctly, observed line profiles must be described as a convolution of both the effects of Doppler broadening and natural broadening. This is expressed as a Voight profile, which typically cannot be solved analytically:

$$\phi(\nu) = \frac{1}{\sqrt{2\pi}} \int \frac{1}{\sigma_\nu} \exp\left(-\frac{\nu^2}{2\sigma_\nu^2}\right) \frac{4\gamma_{ul}}{16\pi^2(\nu - (1 - v/c)\nu_{ul})^2 + \gamma_{ul}^2} dv$$

While this more complicated superposition is necessary to consider for atomic lines of typically ionized gas, this is not the rule for all of the ISM. In practice, for many situations like the molecular gas in the ISM, the large (and as we will discuss, non-thermally turbulent) line widths mean that it suffices to fit lines with simple Gaussian profiles.

7.1.5 Damping Wings

7.2 Absorption

7.2.1 Oscillator Strength

The first thing we want to know for absorption lines is the relative strength of a line, which we will define using the oscillator strength f_{lu} . This is an analog to the Einstein A, which characterized the strength of an emission line through the statistical likelihood that an energy state will decay through emission of a photon. In this case, the likelihood of absorption can be understood by going back to our discussion of dipole moments (See Section 5.5). Again, we visualize an atomic transition as an electron oscillating at the frequency of the transition. The oscillator strength is then the degree to which the actual quantum transition rate resembles the rate of (classical) absorption of a photon by the oscillating electron. Generally the quantum transition it is much less efficient

than this classical approximation, and $f_{lu} \ll 1$ —however for multi-electron atoms, you will sometimes see $f_{lu} > 1$ for a strong transition (see Tables in **Draine Chapter 9**)

The oscillator strength is defined as

$$f_{lu} = \frac{m_e c}{\pi e^2} \int \sigma_{lu}(\nu) d\nu$$

Here, σ_{lu} is the effective cross section of the atom for absorption, which geometrically we can understand as being related to the radius within which a photon of an appropriate frequency must pass to be absorbed.

Draine relates f_{lu} to other variables we have encountered (A_{ul} and B_{ul}) in a somewhat roundabout way.

The rate of absorptions ($\frac{dn_u}{dt}$) can be written as a product of the density of atoms in the lower energy state waiting to be excited upward (n_l), the rate at which photons pass by the atom (c), the number density of those photons in the right frequency range ($u_\nu/h\nu$), and the absorption cross section (σ_{lu}), which is wavelength-dependent and must be integrated over the full frequency range of photons that are present:

$$\frac{dn_u}{dt} = n_l c \frac{u_\nu}{h\nu} \int \sigma_{lu}(\nu) d\nu$$

$$\text{Recalling that } \frac{dn_u}{dt} = B_{lu} n_l u_\nu \text{ we can say } B_{lu} = \frac{c}{h\nu} \int \sigma_{lu}(\nu) d\nu$$

$$\text{or } B_{lu} = f_{lu} \frac{\pi e^2}{m_e h\nu}$$

(see Section 6.1.1 for the definition of the Einstein B absorption coefficient)

As we recall that $B_{lu} = \frac{g_u}{g_l} \frac{c^3}{8\pi h\nu^3} A_{ul}$ we can finally put all of this together to relate f_{lu} and A_{ul} :

$$f_{lu} = \frac{g_u}{g_l} \frac{m_e c^3}{8\pi^2 e^2 \nu_{lu}^2} A_{ul}$$

7.2.2 Optical Depth

We first talked about absorption in the context of the equation of radiative transfer:

$$\frac{dI_\nu}{ds} = -I_\nu \kappa_\nu + j_\nu$$

For a source with initial intensity I propagating through some medium, the final intensity will be changed due to both the absorptivity (κ_ν) and emissivity (j_ν) of the intervening medium.

Here we are going to focus specifically on (frequency-dependent) absorption due to (atomic) transitions.

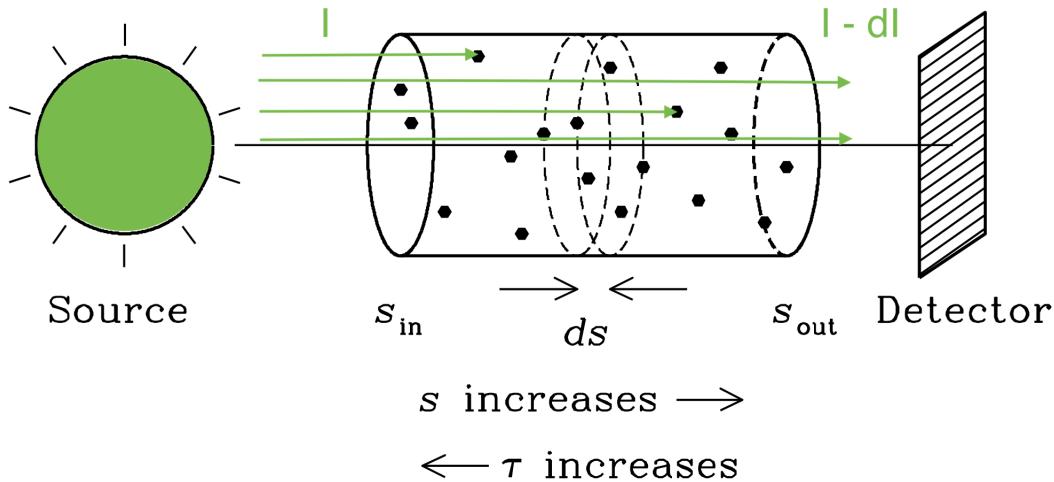


Figure 7.1: *Tau.png*

We will take an aside here to redefine the linear absorption coefficient κ_ν in a probability framework. If we imagine a photon of frequency ν emitted by our source, then we can define the probability that it will be absorbed (p_ν) within some slab of the intervening material with thickness ds as $dp_\nu = \kappa_\nu ds$. The probability of absorption increases proportionally with the thickness of the slab. As probability is a unitless quantity, note that κ_ν has units of inverse length.

We can then rearrange the equation of radiative transfer (keeping only the absorption terms) to write:

$$\frac{dI_\nu}{I_\nu} = -\kappa_\nu ds = -dp_\nu$$

This equation then says that the fraction of intensity lost in traveling through a slab of width ds is just the probability that a photon will be absorbed. If we take a κ_ν value of 1 m^{-1} , and a thin slab with width 10^{-3} m , we can estimate that 10^{-3} of the incident photons will be absorbed: one in a thousand.

The equation of radiative transfer can also be re-cast in terms of a dimensionless quantity known as the optical depth τ , where $d\tau_\nu = \kappa_\nu ds$, or in integral form, $\tau_\nu \equiv \int_s \kappa_\nu ds$. In this case, we can rewrite the equation as:

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu \text{ where } S_\nu, \text{ the source function, is just } \frac{j_\nu}{\kappa_\nu}.$$

For pure absorption, we need only consider $\frac{dI_\nu}{d\tau_\nu} = -I_\nu$

The solution to this equation is $I_\nu = I_{\nu,0} e^{-\tau_\nu}$

Note that, as can be seen in the figure above, s and τ_ν are integrated in opposite directions: the path length s increases moving away from the source ($s = 0$ at the source), but τ_ν is defined to increase as you move from the detector to the source ($\tau_\nu = 0$ at the location of the detector or observer).

For frequency-dependent line absorption, we can write τ_ν in terms of the oscillator strength:

$$\tau_\nu = \frac{\pi e^2}{m_e c} f_{lu} N_l \phi_\nu \text{ (ignoring stimulated emission)}$$

Here N_l is the column density ($N \equiv \int n_l ds$) or the number of absorbers with $E = E_l$ in a pencil beam of solid angle Ω .

As the above expression gives the opacity as a function of frequency across the entire line profile, we can also define a quantity called τ_0 , or the optical depth at the center of the line. This is where the optical depth will be highest, as (for a thermal distribution) the majority of absorbers will have velocities around zero.

We can write:

$$\tau_0 = \sqrt{\pi} \frac{e^2}{m_e c} \frac{f_{lu} N_l \lambda_{lu}}{\sqrt{2} \sigma_v} \text{ where } \sigma_v \text{ is the velocity dispersion (width) of the line.}$$

7.2.3 Equivalent Width and Column Density

The amount of absorption that is present then tells you about the column density of absorbers along your line of sight. Fundamentally, the “amount” of absorption is the amount by which I_ν has decreased. This is most simply parametrized as a quantity called the dimensionless Equivalent width: the fractional decrease in I_ν over the frequency (or wavelength) range of interest. This can be written as:

$$W \equiv \int \frac{1}{\nu_0} \left(1 - \frac{I_\nu}{I_{\nu,0}} \right) d\nu = \int \frac{1}{\nu_0} (1 - e^{-\tau_\nu}) d\nu$$

For the observers among you, you will likely much more commonly encounter a quantity with units of length (hence, the origin of the name ‘equivalent width’).

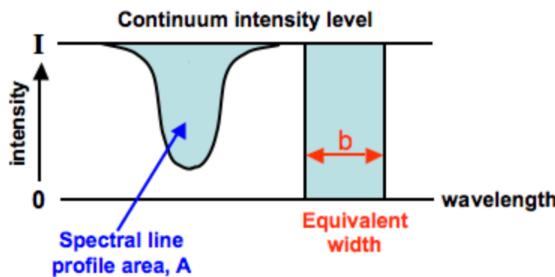


Figure 7.2: eqw.png

This is defined as

$$W_\lambda \equiv \int \left(1 - \frac{I_\lambda}{I_{\lambda,0}} \right) d\lambda$$

where $W_\lambda \approx \lambda_0 W$

Note that such a simplistic definition of the amount of absorption given by the equivalent width is extremely practical for many observations at optical and infrared wavelengths, where instrumental limitations mean that you cannot typically fully resolve spectral line profiles. In the radio and millimeter regime however, the spectral resolution is not a limiting constraint, and so the concept

of equivalent width is rarely used. Instead, one can simply fit, e.g., a Gaussian profile to the line to recover the amount of absorption (the area of the Gaussian) relative to the background continuum emission.

7.2.4 Optically-thin vs. Saturated Lines

For the optically-thin case ($\tau_\nu \ll 1$) we can write the column density in terms of the equivalent width as:

$$N_l = \frac{m_e c^2}{\pi e^2} \frac{W}{f_{lu} \lambda_{lu}}$$

But— what happens when τ_ν becomes large? In this case, a vanishingly small amount of the incident intensity $I_{\nu,0}$ can be observed— and at the line center, none of the intensity is getting through: the observed intensity drops to zero, and the line profile flattens— you can no longer directly relate the amount of absorption to the number of absorbing particles, because so many of the particles near to you didn't see any intensity to absorb. However, as the opacity is lower away from the line center, the outer parts of the line are less affected, and so the wings of the line will continue to grow, and for very large τ_0 the line will appear to broaden appreciably.

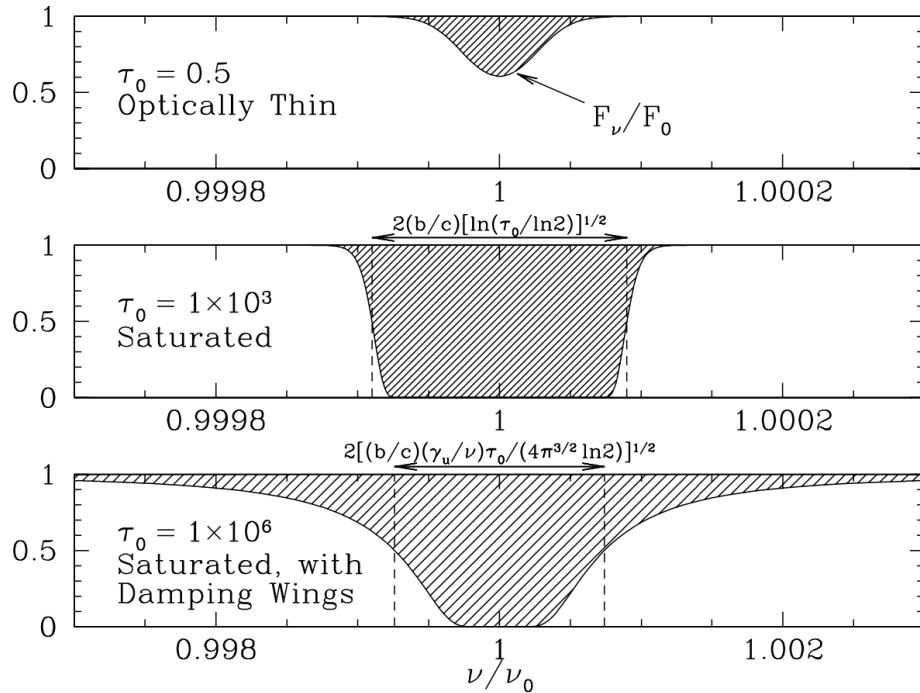


Figure 7.3: *opacity_width.png*

Here we encounter a big problem for observations with limited spectral resolving power: if you cannot spectrally resolve your lines, you cannot tell if the shape indicates that the line has become saturated. This leads to a diagnostic tool called the curve of growth.

7.2.5 Curve of Growth

Honestly, I basically never think about this, because at the wavelengths I work at, we almost always have high enough resolution that we can resolve the relevant line profiles and clearly see if a line is optically thick. But if you CAN'T resolve the line and you KNOW that your unresolved line may be optically thick, this is a way to use the fact that you still detect emission from the optically-thin line wings to (somewhat) accurately estimate the column density.

The curve of growth is basically just a relationship between N and W . When the line is optically thin, there is a well-defined linear relationship, as we have already seen. When the line opacity is large and the line profile is saturated(flat line bottom without yet having prominent damping wings) the curve of growth is “flat” and W is very insensitive to N – this is a bad region for using W to determine an accurate column density. However, once you get to *extremely* large optical depths, one enters a regime where the (Lorentzian) line wings dominate and we can again usefully relate W to N – this is called the “square root” portion of the curve of growth.

Chapter 8

Radiative Trapping

So far, in an effort to fully describe the excitation of molecular gas, we have considered:

- the internal structure of the molecule (Einstein A values and other probabilities for transitions between energy levels set by quantum considerations related to the molecular geometry)
- the transfer of energy due to the rate of physical collisions with other particles (including cosmic rays)
- the transfer of energy from a background continuum radiation field

However, there is one other key consideration needed to make a complete characterization of molecular excitation. This is the transfer of energy from molecule A (as it emits a photon to go from a state with E_u to E_l) to molecule B, which may be able to use that photon itself to go from E_l to E_u .

This process is known as radiative trapping, and it becomes important when the gas becomes optically thick to line emission (e.g., there is a very short mean free path for a photon of a given energy to be absorbed by a molecule in a low-energy state that can absorb that photon and become excited).

Putting this together with all of the previous considerations we have discussed we finally arrive at the real deal: this is the basic physics that underlies observational ISM research which takes observed line intensities and uses them to determine the physical properties of the gas: its density, temperature, and column density.

This is also not a simple problem, as it is non-local: one cannot consider a single molecule at a time, but has to consider the impact of molecules in one part of a cloud on the excitation of molecules in another part of the cloud. As we will see, it thus depends on the geometry of the gas cloud as well as its velocity structure, as a sufficient Doppler shifting of the emitted photons will change their frequency enough to avoid reabsorption.

8.1 Escape Probability Approximation

As is often the case, we will start by backing away from the full complexity of this problem, and making an approximation that greatly simplifies the math.

Here, we begin by assuming that in many cases the problem is *not* as non-local as it seems, and we can begin to effectively decouple the excitation of a single molecule from the impact of molecules elsewhere in the cloud. We do this by determining, based on the geometry and velocity structure of the cloud, the likelihood β_ν that a photon emitted at a frequency ν and a given location r will be able to escape the cloud. This will be a function of its optical depth τ_ν along all the possible paths with directions \hat{n} that it could take:

$$\beta_\nu(r) = \int \frac{d\Omega}{4\pi} e^{-\tau_\nu(r, \hat{n})}$$

This is typically averaged over the normalized line profile ϕ_ν (note that many basic radiative transfer codes just adopt a simplified rectangular line profile, while more sophisticated codes include more complicated line shapes):

$$\langle \beta_\nu(r) \rangle = \int \phi_\nu \beta_\nu(r) d\nu$$

We then construct a binary scenario which assumes that, based on this escape probability, either a photon is effectively absorbed by a molecule so close by that it can be assumed to be absorbed exactly where it was emitted (this is called the 'on the spot' approximation), or it leaves the cloud entirely.

To further eliminate the importance of the location within the cloud on a molecule's excitation, we will also assume that the excitation is uniform across the entire cloud. In this case, we can write the local equation of radiative transfer at a point r as:

$$I_\nu = I_\nu(0)e^{-\tau_\nu} + B_\nu(T_{ex})(1 - e^{-\tau_\nu})$$

where as we recall, B_ν is the source function j_ν / κ_ν , and per our assumptions above is the same at any point in the cloud.

To determine the level populations and so the overall excitation state of the gas, we return to our formula for the rate of change of the population of a given energy level E_u :

$$\frac{dn_u}{dt} = n_l(B_{lu}u_\nu + k_{lu}n_c) - n_u(A_{ul} + B_{ul}u_\nu + k_{ul}n_c)$$

Here the k 's are the collisional rate coefficients, u_ν is the energy density of the radiation field, and n_c is the number density of the collision partner (typically H₂ for molecular gas). Draine rewrites this in terms of the photon occupation number n_γ , removing both the dependence on the Einstein B and on u_ν and adding a dependence on the level degeneracy g to get:

$$\frac{dn_u}{dt} = n_l \left(n_c k_{lu} + n_\gamma \frac{g_u}{g_l} A_{ul} \right) - n_u(n_c k_{ul} + A_{ul} + n_\gamma A_{ul})$$

We now apply the definition of n_γ , noting that $e^{-\tau_\nu}$ is just the escape probability β_ν :

$$n_\gamma(\nu) = \langle \beta \rangle n_\gamma^{(0)} + \frac{1 - \langle \beta \rangle}{(n_l g_u / n_u g_l) - 1}$$

where

$$n_\gamma^{(0)} \equiv \frac{c^2}{2h\nu^3} I_\nu(0)$$

Putting all of these pieces back together and collecting terms we end up with:

$$\frac{dn_u}{dt} = n_c k_{lu} n_l - n_c k_{ul} n_u - \langle \beta \rangle A_{ul} n_u + n_l \frac{g_u}{g_l} \langle \beta \rangle A_{ul} n_\gamma^{(0)} \left(1 - \frac{n_u g_l}{n_l g_u} \right)$$

And this monster is the escape probability approximation. In order, these terms correspond to (1) collisional excitation (2) collisional de-excitation (3) spontaneous emission and (4) absorption and stimulated emission.

As long as we have an expression for β , we can use this equation to solve for the level populations that result from a full consideration of all collisional processes *and* all radiative processes (absorption, stimulated and spontaneous emission) that occur due to photons from the background radiation field and photons that originate within the cloud itself.

We can now also write a more accurate expression for the critical density of a transition (the density required for thermalization) that considers radiative trapping:

$$n_{crit,u} = \sum_c \left(\frac{\sum_{l < u} \langle \beta_{ul} \rangle A_{ul}}{\sum_{l < u} k_{ul}(c)} \right)$$

This is a complete expression for this quantity that includes summing over all collision partners c , and all levels $l < u$ that have transitions connecting to u .

8.2 Geometry: Spherical cloud

We now consider how to determine a geometry-specific approximation for β . As is so often the case in physics, this takes the form of a spherical cow (or rather, cloud).

If we consider a uniform-density static spherical cloud:

$$\beta_{cloud} = \frac{3}{4\pi R^3} \int_0^R \langle \beta(r) \rangle 4\pi r^2 dr$$

Other commonly assumed geometries that one may encounter include a homogeneous slab, a static shell, and expanding shells and spheres.

Using the equation for optical depth:

$$\tau_v = \frac{g_u}{g_l} \frac{A_{ul}}{8\pi} \lambda_{ul}^2 \phi_v \int n_l \left(1 - \frac{n_u g_l}{n_l g_u} \right) ds$$

We define the optical depth at line center going from the center of the cloud ($s = 0$) to its surface ($s = R$) to be:

$$\tau_0 \equiv \frac{g_u}{g_l} \frac{A_{ul} \lambda_{ul}^3}{4(2\pi)^{3/2} \sigma_V} n_l R \left(1 - \frac{n_u g_l}{n_l g_u} \right)$$

Here σ_V is the 1D velocity dispersion, assuming a Gaussian line profile. (I am not going through the full maths here, but note the not-unexpected similarity to the last term of the equation describing the escape probability approximation, which corresponds to absorption and stimulated emission, e.g., processes that would prevent an emitted photon from escaping the cloud.)

β_{cloud} can be well-fit by the function

$$\beta_{cloud} \simeq \frac{1}{1 + 0.5\tau_0}$$

We can understand this result (this is the fraction of escaping photons averaged over the whole cloud) as corresponding to the fraction of the cloud's mass (e.g., the number fraction of equal-mass molecules in the cloud) that lie within a distance from the surface corresponding to $\tau \sim 2/3$ – or the “photosphere” of the cloud.

Ultimately, the determination of a β is an iterative problem: guess a β , solve for the level populations, use these to calculate τ and a new value of β , repeat until a consistent result is arrived at.

So yes, guess and check is a real thing that real scientists do.

Generally, one has some constraints on the level populations directly from observations of the escaping photons that originate from a transition between some upper and lower energy levels. However, one generally cannot observe every allowed transition between every possible energy level. So, to determine the approximate fraction of molecules in a given level of interest compared to all molecules in all levels, we turn to our old friend the partition function.

The relative population in a rotational level J for a linear molecule is given by:

$$\frac{n(J)}{n} = \frac{(2J+1) \exp(-B_0 J(J+1)/k_B T_{ex})}{\sum_J (2J+1) \exp(-B_0 J(J+1)/k_B T_{ex})}$$

The denominator of this expression is the partition function:

$$Z \equiv \sum_J (2J+1) \exp(-B_0 J(J+1)/k_B T_{ex}) \approx \left[1 + \left(\frac{k_B T_{ex}}{B_0} \right)^2 \right]^{0.5}$$

8.3 Velocity Field: Large Velocity Gradient

The simple escape fraction approximation assumed that all of the molecules in the cloud are at rest (though of course they are still undergoing collisions). A more realistic assumption is that the particles are actually moving, either with some uniform bulk motion, or randomly due to turbulence. As mentioned before, this has an important impact on the likelihood that a photon will be reabsorbed, as molecules in the cloud will “see” each other (and photons emitted by other molecules) as having some red- or blue-shift.

This will change the optical depth τ as well as the resulting escape probability β .

Two important cases are that of a uniform ('Hubble') flow like that in an expanding sphere, and that of a uniformly turbulent medium.

The first is typically referred to as the “large velocity gradient” or LVG approximation. For a sphere expanding at every point with a rate $|dv/dr|$, the escape probability can be written as

$$\langle \beta \rangle = \frac{1 - e^{-\tau_{LVG}}}{\tau_{LVG}}$$

Here, the optical depth for the LVG approximation is just

$$\tau_{LVG} \equiv \frac{\lambda_{ul} \int \kappa_\nu d\nu}{|dv/dr|}$$

which can be understood as the total optical depth across the cloud for a photon at any frequency that could be absorbed by a nearby molecule:

$$\nu \text{ in } \nu_0[1 \pm (R/c)(dv/dr)]$$

For turbulent clouds, one just adapts the earlier expressions for β and τ for a uniform, static cloud to effectively encompass a larger line width corresponding to the turbulent broadening.

One can take the specific case of line broadening in a self-gravitating cloud, defined by the virial theorem. Here,

$$\sigma_V^2 = \frac{GM}{5R}$$

This value can be plugged in to the expression we previously saw for τ_0 :

$$\tau_0 \equiv \frac{g_u}{g_l} \frac{A_{ul}\lambda_{ul}^3}{8\pi} \left(\frac{5}{2\pi G} \right)^{0.5} \frac{n_l R^{3/2}}{M^{1/2}} \left(1 - \frac{n_u g_l}{n_l g_u} \right)$$

Finally, relating this back to observable quantities, we can express the total cloud luminosity in a given spectral line as:

$$L_{ul} = \int 4\pi r^2 dr n_u A_{ul} h\nu_{ul} \langle \beta \rangle_{cloud} \simeq \frac{4\pi}{3} R^3 n_u A_{ul} h\nu_{ul} \frac{1}{1 + 0.5\tau_0}$$

This then relates the observed brightness of a spectral line to the number of molecules in a level with energy E_u , allowing us to go from an observed line intensity to a determination of the level populations n_u .

While this exact expression describes a (turbulent) sphere, one can also write expressions for other common geometries (e.g., shells, cylinders, or uniform slabs).

Chapter 9

Astrochemistry

9.1 Astrochemistry 101

Chemical reactions in space:

- Are more difficult than reactions on earth (due to low temperatures and extreme low densities). However, they can occur over extremely long timescales (many thousands of years), which does allow complex gas chemistry to arise despite these challenges. Additionally, the low densities also means that some species that don't last on earth b/c they are so highly reactive (e.g., radicals) can be present in some quantity in the ISM.
- Do not have high (or any) activation energy/ activation barrier. This is because most rate coefficients (k) have some exponential term $\exp(-E/kT)$: remember what happens to such a term if $E \gg kT$!
- Favor reactions with large (Coulomb-aided) cross sections (e.g., Ion-molecule and dissociative recombination) because these tend to be faster

Astrochemistry reactions are generally divided into two broad categories: gas-phase and grain-surface reactions.

Gas Phase

In the gas phase, chemical reactions proceed via collisions, with a typical collisional rate coefficient on the order of $k=10^{-10} \text{ cm}^3 \text{ s}^{-1}$.

For a cold ($T_{kin} = 10 \text{ K}$) and dense ($n = 10^4 \text{ cm}^{-3}$) molecular cloud, we can then calculate:

- Collision rate per particle of $kn = 10^{-6} \text{ s}^{-1}$
- Timescale = 1/Collision rate = 10^6 s or 2 weeks
- Mean free path = Timescale * mean velocity = 10^5 km

(Note that in order to determine the mean free path we made an assumption that the velocity of the gas was equal to the thermal velocity at 10 K)

Though this seems limiting, it turns out that you can still make a lot of (rather complex!) molecules in ISM gas. HOWEVER, saturated organic species require high temperatures and/or 3-body reac-

tions, and so are not easily formed in the gas phase (the same goes for H₂, as we discuss more in Section 9.2.3.)

Grain Surfaces

Dust grain surfaces then play a key role in moderating ISM chemistry: as a meeting place to get atoms and molecules together, as a third body to stabilize highly exothermic reactions (and sometimes a shield against UV light, if reactions take place within and under ice layers).

We discussed how to generally form molecules on dust grains in the **October 29** lecture. Once molecules form on grain surfaces:

- If the dust grain is cold enough, the molecules will stick around, and will tend to form ices on the grain surface (Some common ice constituents are H₂O, CH₃OH, CO₂, CH₄, NH₃)
- If it is VERY cold, then molecules already in the gas phase will also adhere to the dust grains (this causes depletion of gas-phase species like CO)
- Molecules can be driven (back) off of dust grains generally through two main processes: shocks and heating (and combinations of the two). Shock sputtering can even erode and destroy the dust grains themselves, releasing components like Si (molecules like SiO are commonly seen to be tracers of shocks).

9.1.1 Important environments for astrochemistry

- **Diffuse Clouds** $n < 10^2 \text{ cm}^{-3}$, $T = 20\text{--}100 \text{ K}$: Despite low densities and correspondingly long collisional time scales, a surprising wealth of molecules is observed in these environments. The presence of simple diatomic molecules in the mostly atomic gas is at least in part due to the role played by UV radiation in ionizing atoms and molecules (which then promotes efficient ion-molecule reactions), though UV radiation will also play an important role in destroying molecules. Such rapid photodissociation means that the chemistry observed in these environments is often very close to an equilibrium between formation and destruction. The formation of more complex molecules observed in these environments is still an active area of research, but may be due to higher-density substructures or the passage of shocks through this gas. Diffuse cloud chemistry is most commonly probed through absorption spectroscopy toward background continuum sources.
- **Dense Clouds** ($n = 10^2 - 10^5 \text{ cm}^{-3}$, $T = 10 - 20 \text{ K}$): These environments have low temperatures and low ionization fractions but high densities, which increases the collisional rate coefficients, and makes grain-surface reactions productive. As UV light cannot penetrate to the centers of these clouds, the molecules that form here are not subject to photodissociation. The densest substructures in these clouds (for example, cores that are collapsing to form low-mass stars) is also often characterized by a high fraction of deuterated molecules. A classic example of dense cloud chemistry is the Taurus molecular cloud, a quiescent region with no ongoing high-mass star formation. Chemistry in this type of environment (sometimes called dark cloud chemistry) is characterized by the formation of long carbon chains, like HC₁₁N, as well as the depletion of some molecules from the gas phase onto the dust grains.
- **Massive Protostellar Envelopes** ($n = 10^4 - 10^{10} \text{ cm}^{-3}$, $T = 25 - 2000 \text{ K}$): These environments have both high densities and high temperatures, which allows formation of some of the most complex molecules observed in the ISM. The high temperatures release molecules

formed on grain surfaces into the gas phase, and in the densest of these environments, three-body reactions can become important. Molecules in these environments are often detected through vibrationally-excited transitions as well as rotationally-excited transitions. The classic example of this chemistry (sometimes called hot core chemistry) is the Sgr B2 star forming region in the Galactic center. Of the ~200 molecules detected in space, more than half were discovered here, and nearly all known ISM molecules have ultimately been observed in this source. A more nearby source with slightly less complex chemistry is Orion-KL.

- **Shocks:** These are a common feature of both protostellar envelopes and turbulent gas in the ISM. Recently shocked gas can reach extremely high temperatures and densities that can overcome activation barriers in reactions, and favor neutral-neutral reactions that would otherwise proceed too slowly to be important. Shocks enhance the abundance of Si-bearing species like SiO, and can be observed in vibrationally-excited H₂ and masers of CH₃OH and H₂O.
- **Evolved Star Envelopes** ($n = 10^8 - 10^{11} \text{ cm}^{-3}$, $T = 600 - 1000 \text{ K}$): Not only are these environments important for forming dust grains, they also (perhaps unsurprisingly) can be extremely chemically rich. Here again we have many important ingredients for facilitating chemical reactions: high temperatures, high densities, and an abundance of dust for grain surface reactions. The chemistry in evolved star envelopes divides into two cases: oxygen-rich (O/C > 1, typical of “normal” cosmic abundances) and carbon-rich. In each case, all of the carbon (or oxygen) is locked up in CO, leaving the other species to form the basis of subsequent chemistry. A classic example of carbon-rich envelope chemistry is IRC +10216.

9.1.2 Some general classes of molecules

- Molecules that predominately form via grain-surface reactions: H₂, NH₃, CH₃OH
- Molecules that are typically thermalized: CO, NH₃ *Used for temperature measurements*
- Molecules with large electric dipole moments: HCN, HC₃N, HCO⁺, CS *Used to trace dense (star-forming) gas*
- Molecules that trace shocks: SiO, CH₃OH, masers of OH, H₂O, and CH₃OH *Used to trace outflows from protostars*
- Carbon-chain molecules: HC_nN, C_nS, C_nH *Indicative of complex, organic chemistry in relatively quiescent environments*
- Ions: HCO⁺, N₂H⁺, H₃⁺, H₃O⁺ *Abundances measure the degree of ionization in molecular gas and kinematics can give insight into the magnetic structure of the gas*
- Radicals: OH, CN *Because radicals are highly reactive (due to their unpaired valence electrons), these are most commonly observed in low-density environments*
- Commonly-substituted isotopes: D/H, ¹²C/¹³C, ¹⁶O/¹⁸O, ¹⁴N/¹⁵N. *Used for abundance studies and avoiding optical depth issues*

It is useful to mention that not all spectral lines that are observed in space have been identified! A large number of unidentified features (e.g., the diffuse interstellar bands) and lines (as high as 10-15% in millimeter surveys of sources like Sgr B2 and Orion-KL) have been observed, and astronomers have so far been unable to assign them to known molecules. How can this happen? For one, the environment of the ISM is sufficiently different from the terrestrial environment that chemistry (and even just spectra under a wide range of energetic conditions) has not been fully probed in the lab. While some of these lines may be from molecules not yet detected in space, many of the lines may simply originate in the rich vibrational spectra of already-identified com-

plex molecules. There may even be a small number of culprits behind large numbers of unidentified lines: many known ISM molecules already have the reputation of being “weeds” (vinyl cyanide, CH_2CHCN , is a notorious example) due to the overwhelming complexity of their spectra!

9.1.3 Types of Reactions

We have already seen a large number of examples of chemical reactions throughout this course. To summarize:

- * Ionization/Recombination : $\text{H} + \text{h}\nu \rightleftharpoons \text{p} + \text{e}^-$
- * Radiative Association/Photodissociation: $\text{H} + \text{H} \rightleftharpoons \text{H}_2 + \text{h}\nu$
- * Collisional Excitation/De-excitation: $\text{H} + \text{H} \rightleftharpoons \text{H}^* + \text{H}$
- * Radiative Excitation/De-excitation: $\text{H} + \text{h}\nu \rightleftharpoons \text{H}^*$

However, some of the most important types of reactions for ISM chemistry are those involving charged species, for which there is a Coulomb attraction that significantly increases the collision/reaction rate. Two important examples of these are:

- * Ion-Neutral Reactions: $\text{H}_2^+ + \text{H}_2 \rightarrow \text{H}_3^+ + \text{H}$
- * Dissociative Recombination: $\text{H}_3^+ + \text{e}^- \rightarrow \text{H}_2 + \text{H}$

There are also neutral-neutral reactions. In addition to radiative association (which can be either neutral-neutral or ion-neutral) we have:

- * Neutral-Neutral Exchange: $\text{OH} + \text{H} \rightarrow \text{O} + \text{H}_2$

Already slower than ion-molecule reactions, these reactions generally need to be exothermic in order to proceed at a reasonable rate at the low temperatures that characterize most of the ISM.

9.1.4 Reaction Rates

As a reminder: reaction rate coefficients k_{AB} (e.g., for a reaction $\text{A}+\text{B} \rightarrow \text{AB} + \text{h}\nu$) have standard units of $\text{cm}^3 \text{s}^{-1}$.

A volumetric reaction rate corresponds to $n_A n_B k_{AB}$ and has units of $\text{cm}^{-3} \text{s}^{-1}$. You will also sometimes encounter a ‘per molecule’ reaction rate which is equivalent to $n_A k_{AB}$ and has units of just s^{-1} .

Whether or not a reaction is likely to be important must be judged against some time scale. Ultimately, it had better not be longer than the age of the universe! Astrophysically-important time scales include the lifetime of a molecular cloud (max 20-30 million years), a protostellar disk (max 10 million years), or an asymptotic giant branch star (max 1 million years).

For some common ISM reaction types:

Reaction Type	Typical rate, k	
Radiative Association		
$[\text{A}^+ + \text{B} \rightarrow \text{AB}^+ + \text{h}\nu]$	diatomic	10^{-17}
	polyatomic	10^{-17} to 10^{-9}

Reaction Type		Typical rate, k
Neutral-Neutral		
$[A + B \rightarrow C + D]$	neither radical one or both radical	10^{-13} 10^{-11} to 10^{-10}
Ion-Molecule		
$[A^+ + B \rightarrow C^+ + D]$		10^{-9}
Dissociative Recombination		
$[A^+ + e^- \rightarrow C + B]$		10^{-7} to 10^{-6}

9.1.5 What's what in the ISM?

So how do scientists figure out what molecules they are looking at in the ISM?

Many references compile data taken from theory/computations, laboratory experiments and astronomical observations which include quantum numbers of the transition, the energies of the states involved, and the central line frequency. However, these databases are HUGE, so even searching a frequency range of just 1 GHz (e.g., 100 to 101 GHz) can yield thousands of possible identifications for a line. However, all is not lost.

First, of all, if you know your source velocity to ± 10 km/s, then you should know the frequency to within .003% (so, for a frequency around 100 GHz, this is to within ± 3 MHz). This helps narrow things down a lot!

Some additional rules of thumb: * Making complex molecules is hard enough that simpler species (diatomic/triatomic) are more likely to be abundant and so responsible for strong emission and absorption lines. * Similarly, H is at least $1000\times$ more abundant than other (reactive) atoms, so most molecules you see will not contain elements beyond the most common (H,C,N,O,S and Si), and the more non-H atoms in a molecule, the less likely it is to be highly abundant. (Values below from Wilson & Rood 1994)

Table 1 Abundances for CNO and heavier elements

Element	Solar System ^a	Orion A ^{b,c}
(C/H)	3.6×10^{-4}	3.4 to $\sim 2.1 \times 10^{-4}$
(O/H)	8.5×10^{-4}	4.0 to 3.8×10^{-4}
(N/H)	1.1×10^{-4}	6.8 to 8.7×10^{-5}
(Ne/H)	1.2×10^{-4}	~ 8 to $\sim 40 \times 10^{-5}$
(Si/H)	3.6×10^{-4}	3.0×10^{-6}
(S/H)	1.8×10^{-5}	8.5 to 13.3×10^{-6}
(Ar/H)	3.6×10^{-6}	4.5 to 2.1×10^{-6}

Figure 9.1: elemental.png

Rare isotopes are also, as their name suggests, rare. ^{13}C is at least 20 times less abundant than ^{12}C , and rare isotopes of O or N can be several hundred times less abundant. (Values below from Wilson & Rood 1994)

Table 4 Ratios for Galactic center, 4 kpc molecular ring, carbon stars, Solar System, local ISM, and galaxies

Isotope	Galactic center	4 kpc molecular ring	Local ISM ^b	Solar System ^c	Carbon stars ^d	Nuclei of galaxies
(¹² C/ ¹³ C)	~ 20	53 ± 4 ^b	77 ± 7 ^b	89	> 30	~ 40 ^h
(¹⁴ N/ ¹⁵ N)	> 600	375 ± 38 ^b	450 ± 22 ^b	270	> 515	...
(¹⁶ O/ ¹⁸ O)	250	327 ± 32 ^b	560 ± 25 ^b	490	320 to 1260 > 2700	~ 200 ⁱ
(¹⁸ O/ ¹⁷ O)	3.2 ± 0.2 ^e	3.2 ± 0.2 ^e	3.2 ± 0.2 ^e	5.5	0.6 to 0.9 < 1	8 ^j
(³² S/ ³⁴ S)	~ 22 ^f	~ 22 ^f	~ 22 ^f	22
(²⁹ Si/ ³⁰ Si)	1.5 ^g	1.5 ^g	1.5 ^g	1.5

Figure 9.2: isotopic.png

Energy levels are important! In most ISM environments, you are more likely to see transitions with low E_{up} (say, substantially less than a few hundred K). This means that in typical environments, most vibrationally-excited lines will be weak or absent.

These rules apply additively, e.g., if you think you see a vibrationally-excited line of NaH¹³CO₃, you should probably question this identification!

Further key considerations: * Has anyone ever seen this molecule in space before? (https://en.wikipedia.org/wiki/List_of_interstellar_and_circumstellar_molecules) * Can you identify other lines of the same molecule? If not, are there any lines (same molecule, similar level energies) that you *should* be seeing in this frequency range if your identification is correct? * How well do you know your source velocity, and are there multiple distinct velocities present in this source? * Are the conditions in the source uniform? (for example, do you expect it to contain any shocks or regions where the temperature and density are extremely different than the average values)

9.2 H₂

9.2.1 Observational Challenges

H₂ is the simplest and most lightweight molecule that can be formed, and it is also the most abundant molecule in the interstellar medium. However, its abundance is pretty much the only thing that it has going for it, in terms of facilitating direct observations of H₂ in cool gas.

The first problem that we run into is that H₂ in its simplicity is also quite symmetric. The upshot of this is that it does not have a permanent electric dipole moment, and electric dipole radiative transitions do not occur. The molecule can then only radiate through much weaker electric quadrupole transitions. This leads to a two-fold issue: first, these transitions are much weaker (occurring less frequently), which means that the lines will be relatively faint.

The second problem with the symmetry of H₂ is that the two protons are identical fermions, which restricts the possible spin states of the nuclei and electrons: the wave function must be antisymmetric under exchange of electron positions. As the nuclear spin-state does not change in these transitions, rovibrational and pure-rotational transitions can only occur for $\Delta J = \pm 2$, e.g., transitions must occur between para (even J , spin states anti-aligned) or ortho (odd J , spin states aligned) states.

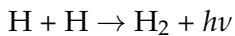
The next problem is that H_2 is quite lightweight, and so the energies of the different levels are also very widely separated: there is a huge jump in energy in its rotational states, compared to a heavier molecule like CO which has level energies $E = 5.5 \text{ K}, 16.6 \text{ K}, 33.2 \text{ K}$ for $J = 1, 2, 3$.

The fact that H_2 is lightweight is also a problem for observers, as it puts the wavelengths of its rotational transitions in the mid-infrared (as opposed to typical rotation transition wavelengths in the millimeter/submillimeter for moderate-weight molecules like CO, and typical rotation transition wavelengths in the centimeter regime for really heavy molecules like HC_3N). This is a problem because mid-infrared wavelengths are not easy to observe from the ground, due to low atmospheric transmission. As a result, the only telescope that can observe the longest-wavelength, lowest- J transitions of H_2 is SOFIA (a telescope that flies in the stratosphere on a 747). Due to limitations on the instruments it has on board, it cannot observe the lowest pure-rotational line: $S(0), J = 2 - 0$, and this line will not be observable again until the James Webb Space Telescope launches.

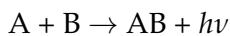
9.2.2 Gas-Phase Formation

Even though H_2 is the simplest molecule, it turns out that it is *not* the easiest molecule to make!

The most basic way you might imagine making it: Slam a couple hydrogen atoms together and call it a day.



This is an example of a more general type of reaction called radiative association:



The trouble with this reaction is not getting two H atoms together: in gas of density $n_H = 10^2 \text{ cm}^{-3}$ and temperature $T = 500 \text{ K}$, the rate of collisions among H atoms will be (assuming a radius of 1 angstrom or 10^{-8} cm):

$$R_{coll,H} = n_H \langle \sigma v \rangle$$

$$R_{coll,H} = \pi(2r)^2 n_H \left(\frac{8k_B T}{\pi m_H} \right)^{1/2}$$

$$R_{coll,H} = \pi(2 \times 10^{-8})^2 (10^2) \left(\frac{8(1.38 \times 10^{-16})(500)}{\pi(1.67 \times 10^{-24})} \right)^{1/2} = 4 \times 10^{-8} \text{ s}^{-1}$$

for a typical collisional time scale of 0.8 years.

When you consider this is the timescale on which any given H atom will be expected to come in contact with another, this would lead to extremely rapid formation of H_2 !!

The problem is that once you do get two H_2 molecules close enough together to form a covalent bond, there is an excess energy of formation (this is an exothermic reaction) that must be gotten rid of for the molecule to be stable (and not immediately vibrate to bits). And as we just saw in Section 9.2.1, H_2 is particularly awful at giving off photons: the chance that it will give off a photon in time to avoid coming apart (e.g., during the duration of the collision) is extremely low. How low?

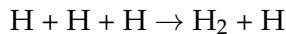
Well, the timescale on which it must give off a photon is 10^{-13} s . An appropriate Einstein A for the required quadrupole radiation is 0.38 s^{-1} .

This means the likelihood of giving off that photon is 3.8×10^{-14} . When you factor this probability into the collision rate to get a formation rate, the formation rate becomes

$R_{2\text{-body}} = 1.5 \times 10^{-21}$: now the time for any given H to undergo a collision under these conditions that would make H₂ is 20 trillion years: much longer than the age of the universe.

There are a few ways to get around this, each with their own challenges.

- (1) The first way to still make H₂ in the gas phase is to require three H atoms to collide, so that the third hydrogen atom can carry away the excess energy from the reaction. This is a three-body interaction, and so requires a very high density to operate efficiently:



How high a density does this need? Let's just say we would settle for being able to turn H into H₂ on the timescale of the age of the universe.

Taking the 3-body collision rate (which would basically equal the formation rate, since there is no radiative barrier to forming H₂ with this reaction):

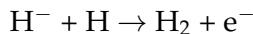
$$R_{3\text{-body}} = \frac{32\pi}{3!} n_H^2 (2r)^5 \left(\frac{8k_B T}{\pi m_H} \right)^{1/2}$$

Setting $\frac{1}{R_{3\text{-body}}}$ to be 10^{10} years,

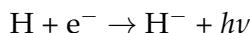
$$3 \times 10^{18} = \frac{32\pi}{3!} n_H^2 (2 \times 10^{-8})^5 \left(\frac{8(1.38 \times 10^{-16})(500)}{\pi(1.67 \times 10^{-24})} \right)^{1/2}$$

and $n_H = 10^{25}$ cm⁻³ which is indeed enormous, and not really of much help here.

- (2) Alternately, we can consider a different two-body reaction pathway:



This is called an exchange reaction and it makes a more promising route for forming H₂ because now the electron carries away the excess energy of formation, so that this reaction can proceed relatively quickly. The tricky thing is that the rate of this process now also depends on the rate of forming H⁻:



This reaction (another radiative association reaction) is slow except at high T , so this reaction is usually the limiting factor in forming H₂ through the above exchange reaction. Still, all together this is the fastest way to form H₂ in the gas phase.

9.2.3 Grain-Surface Formation

We know that actually there is a lot of H₂ around– it is by far (99.99999common molecule! This means we really need some way to make it efficiently. This brings us to forming H₂ on grain surfaces.

The basic procedure to make molecules on dust grains is as follows:

1. An atom (or molecule) hits a dust grain. This is helped by relatively large dust cross sections (though the number density of dust is fairly low).
2. It then has some probability of sticking (the

particles sit in small surface potential wells) and remaining on the grain (either due to electrostatic forces, or by forming a chemical bond with the surface). 3. Large nuclei tend to stay put in their potential well, and the lighter H atoms hop around, interacting with the particles in other wells. 4. Over time, (saturated) molecules then form on the grain surface: H_2 , H_2O , NH_3 , CH_4 , CH_3OH , H_2CO etc. 5. The dust grain lattice acts as a third body to stabilize formation of molecule, allows it to release excess energy of formation.

To estimate the formation rate, we assume that if an H atom collides with a grain it sticks, and that if it sticks it inevitably meets up with another H atom and forms H_2 . Then the formation rate on dust grains can be described as

$$R_{dust} = 0.5v_H\sigma_d n_d$$

We can evaluate this by assuming v_H is just the thermal velocity, taking $n_d = 10^{-12}n_H$ and assuming a typical grain radius to be $r_d = 3 \times 10^{-5}$ cm. Plugging in the same values as before:

$$R_{dust} = 0.5 \left(\frac{8k_B T}{\pi m_H} \right)^{1/2} (2\pi r_d)^2 n_H n_d$$

$$R_{dust} = 2\pi \left(\frac{8(1.38 \times 10^{-16})(500)}{\pi(1.67 \times 10^{-24})} \right)^{1/2} (10^{-5})^2 (10^2)(10^{-12}) = 2 \times 10^{-14} \text{ s}^{-1}$$

for a MUCH more reasonable formation timescale of 1.6 million years.

9.2.4 Lyman and Werner Bands

We know that a pure-rotational transition from $J=1$ to $J=0$ is not allowed for H_2 . This can make it difficult to probe the populations of these levels (which will contain the bulk of the H_2 molecules) as observing emission from these levels requires that levels $J = 2$ and higher are populated, which requires relatively high temperatures ($T > 100$ K).

However, it is possible to directly probe the populations of these levels in colder gas via measurements of absorption transitions due to background radiation, where H_2 goes from an electronic ground state to an electronically-excited state. These transitions have permitted dipole transitions (so much larger oscillator strengths, meaning they will lead to strong absorption features). Transitions from the electronic ground state to the first electronically-excited state are known as Lyman Band transitions, and transitions from the ground electronic ground state to the second electronically-excited state are known as Werner Band transitions. Looking at cold gas along a line of sight to a strong background radiation source (for example, a quasar or a white dwarf) can then be a good way to directly measure H_2 column densities in relatively diffuse gas, and to infer things like the required formation rate of H_2 .

9.2.5 Photodissociation

Note, as these are electronic transitions (and so highly energetic), these absorption lines occur at UV wavelengths. Additionally, because the required photons are so energetic, there is about a 10% chance that absorption through the Lyman or Werner bands will result in the dissociation of H_2 . Here, instead of decaying back to a low rovibrational state in the electronic ground state, it decays into a highly vibrationally-excited bound level of the electronic ground state and then emits a photon at a wavelength between 1200 and 1600 angstroms (the H_2 “vibrational continuum”) as

the molecular bond breaks and H₂ dissociates. This process, known as photodissociation, occurs in only 10⁻¹⁴ s! This is the primary mechanism for the destruction of H₂.

9.2.6 UV Pumping

For the remaining 90% of cases, the H₂ molecule will eventually de-excite by emitting a cascade of radiation as it returns to the electronic ground state, likely ending up in a rotationally-excited state of H₂. The phenomenon in which the absorption of UV radiation leads to an enhanced population in the high-*J* rotational states of H₂ is known as pumping.

(Note that H₂ formation into an excited electronic state is actually another potential way to form H₂ in the gas phase, as there are dipole-allowed transitions out of the excited electronic state that are much faster than the rovibrational transitions)

9.2.7 H₂ Self-Shielding

The presence of strong self-shielding can affect the extent to which UV pumping will influence H₂ level populations: for high column densities, the level populations will be primarily affected by collisions, and so can be used to measure the kinetic temperature of the gas. In general, when attempting to use H₂ rotational lines as a thermometer, the relative populations of the vibrationally-excited levels, the strength of the background radiation field, and the total gas column density must all be closely scrutinized in order to determine whether the rotational level populations are consistent with collisional excitation or whether the population of high-*J* levels are primarily influenced by this cascade down from the electronically-excited states.

9.2.8 H₂ Cooling

Even though H₂ is not a particularly efficient radiator, and H₂ can't effectively cool gas down to the temperatures observed in present-day molecular clouds (*T* = 10 – 50 K).

Recall, to be an effective cooling mechanism, collisional cooling requires

- (1) Frequent collisions that do not require a low-abundance partner
- (2) The excitation energy must be comparable to or less than the thermal kinetic energy of the gas particles. (The lowest *J* level of H₂ that can radiatively decay has *E* ~ 500 K)
- (3) There must be a high probability of emitting a photon after being excited

H₂ fails 2 out of the 3 criterion here, yet H₂ collisional cooling is still important in warm (*T* > 100 K) molecular gas simply due to the sheer number of H₂ molecules that are present in the ISM, which overwhelms the low probability of an individual molecule emitting a photon after excitation (via a quadrupole rather than dipole interaction).

H₂ was also an extremely important species for cooling gas in the very early universe, when making the first stars. At that time, there was no carbon, oxygen, or other metals, and no dust, so the only cooling available is from continuum emission, H recombination lines, and H₂. Of course, even forming the H₂ necessary to cool the gas was difficult then, as dust-mediated formation is out of the question, so one must rely on a slower gas-phase formation route.

Chapter 10

Dust

10.1 Dust Properties

10.1.1 Composition

The makeup of interstellar dust grains is broadly well-constrained, but depends on many inferences. Thus, the specifics of dust grain composition (the actual mix of minerals and how this changes as a function of time and environment) is still an active area of investigation, dominated in many places by theoretical modeling.

Inferred from Gas Depletion

One of the main ways that the composition of dust grains is determined on an atomic level is through the depletion of elements in the gas phase. Here, observations of interstellar gas are compared to stellar/solar abundances, which are assumed to be the benchmark of a ‘typical’ composition. Elements that show particular depletion and are believed to be some of the primary constituents of dust include C, Si, and Fe.

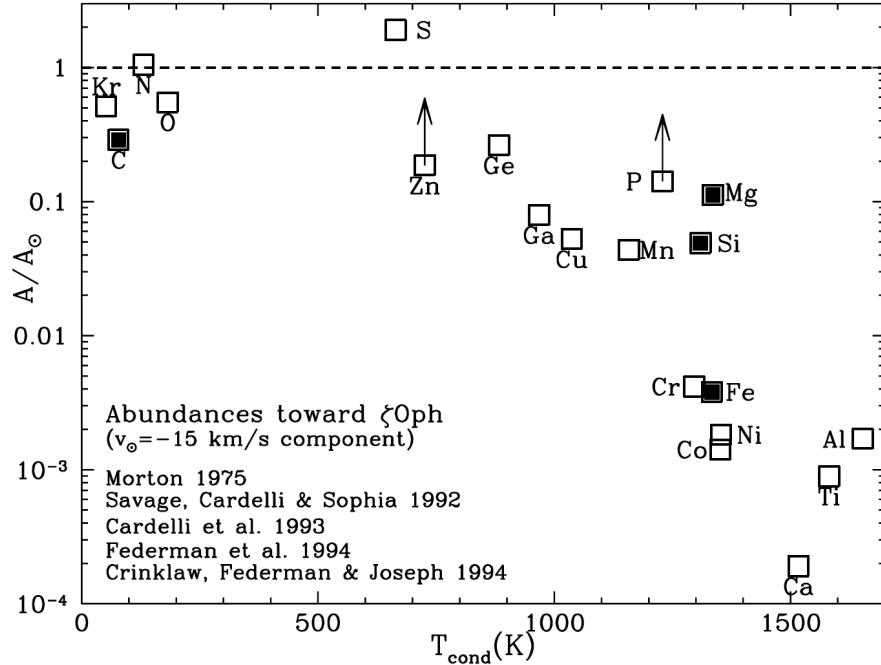
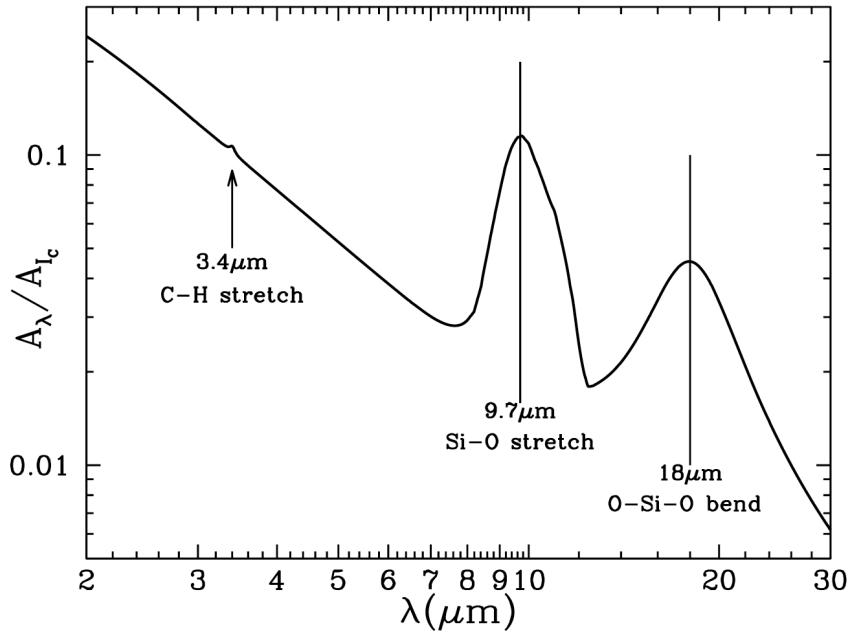


Figure 10.1: *dust_depletion.png*

Inferred from Extinction Features

Another way to get constraints on the composition of dust is through spectral features that are seen (for example) in the amount of dust extinction as a function of wavelength. These are broad features, but they are generally accepted as being due to vibrational features in the complex molecules making up dust grains. These are characteristic of bending and stretching of specific bonds (C-H, Si-O, O-Si-O), where the vibrational frequencies of these modes are modified by the moment of inertia of the molecule as a whole.

Figure 10.2: *dust_features.png*

Determined from Solar System Compositions

The most direct way to measure dust composition, albeit with the limitation of potentially applying to a specific stage of dust evolution, is to study the composition of dust grains in samples of material from the solar system. While this allows for accurate identifications of specific mineralogical compositions and their relative abundances, the complication is that this may not be representative of typical interstellar dust, as some of the dust grains may have undergone additional (thermal) processing in the pre-solar nebula.

Table 23.2 Types and Properties of Major Presolar Materials^{a,b} Identified in Meteorites and IDPs.

Material	Source	Grain Size (μm)	Abundance ^c (ppm) [†]
Amorphous silicates	Circumstellar	0.2–0.5	20–3600
Forsterite (Mg_2SiO_4) Enstatite (MgSiO_3)	Circumstellar	0.2–0.5	10–1800
Diamond		~ 0.002	~ 1400
P3 fraction	Not known		
HL fraction	Circumstellar		
Silicon carbide	Circumstellar	0.1–20	13–14
Graphite	Circumstellar	0.1–10	7–10
Spinel (MgAl_2O_4)	Circumstellar	0.1–3	1.2
Corundum (Al_2O_3)	Circumstellar	0.5–3	0.01
Hibonite ($\text{CaAl}_2\text{O}_{19}$)	Circumstellar	1–2	0.02

Figure 10.3: *dust_composition.png*

10.1.2 Grain Sizes

Constraints on the sizes of dust grains can be determined in several ways. One way (apart from direct size measurements for solar system grains) is to determine dust grain sizes from their emission properties: for example, in the far-infrared, dust grains cease to be efficient radiators at wavelengths larger than their sizes. Other observed features at shorter wavelengths can also give insight into relative compositions (for example, the relative strengths of the C-H and Si-O stretch features). A combination of models and observations can then be used to infer size and composition distributions, for example to measure grain growth in protostellar disks. Making these dust models is an area where Draine is particularly well-known.

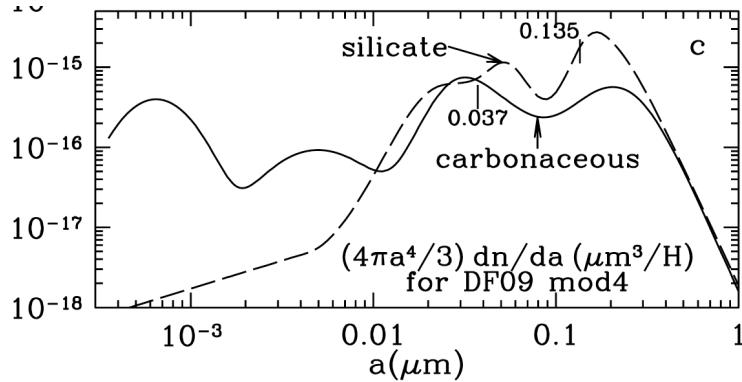


Figure 10.4: *dust_size.png*

There is some hope that in the size range between molecules and the smallest dust grains, in the regime of PAHs, one could directly observe line emission that would identify the overall molecular structure. The existence of PAHs— polycyclic aromatic hydrocarbons— is suggested by the presence of vibrational features in the near- and mid-IR that correspond to bending/flexing of carbon ring structures.

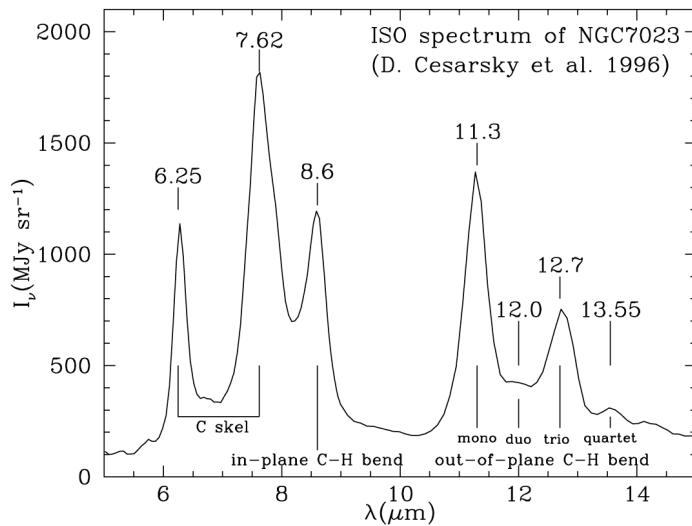


Figure 23.7 The 5 to 15 μm spectrum of the reflection nebula NGC 7023 (Cesarsky et al. 1996).

Figure 10.5: *dust_PAH.png*

The largest such molecule that has been identified in the ISM is C₆₀- Buckminsterfullerene. However, despite some candidates for PAHs, none of the planar PAH molecules have been observationally identified (generally due to a lack of emission due to their symmetry of their structure).

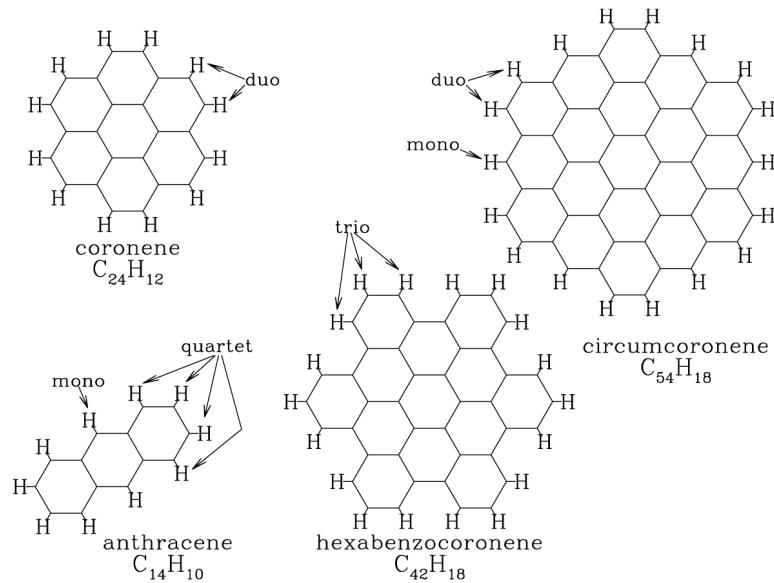


Figure 10.6: *dust_PAH2.png*

10.2 Dust Heating

10.2.1 Radiative Heating

The heating rate of dust grains due to the absorption of light from the background radiation field will be a function of the radiation density u_ν and the dust grain cross section (πa^2 , assuming a spherical dust grain shape) and can be written as:

$$\left(\frac{dE}{dt} \right)_{abs} = \int \frac{u_\nu d\nu}{h\nu} \times c \times h\nu \times Q_{abs}(\nu) \times \pi a^2$$

We can understand this as an analogy of a collisional rate, where the "collision" occurs between dust grains and photons. Here

$$\frac{u_\nu d\nu}{h\nu}$$

is effectively a number density of photons with a given frequency, and we recall that

$$u_\nu = \frac{4\pi}{c} B_\nu(T).$$

c is the speed of the collision, $h\nu$ is the energy imparted, and $Q_{abs}\pi a^2$ is the grain cross section.

Q_{abs} is first defined in **Draine Equation 22.6** as $C_{abs} = Q_{abs}\pi a^2$, where Q_{abs} is an efficiency factor that normalizes the absorption cross-section C_{abs} to the equivalent cross-section of a spherical grain.

We can further normalize Q_{abs} by averaging over the spectrum of the incident radiation:

$$\langle Q_{abs} \rangle_* = \frac{\int Q_{abs} u_{*,\nu} d\nu}{\int u_{*,\nu} d\nu}$$

This just lets us simplify our equation for the radiative heating rate:

$$\left(\frac{dE}{dt} \right)_{abs} = \langle Q_{abs} \rangle_* \pi a^2 u_* c$$

Here * indicates that the radiation field is assumed to be starlight. Note that in diffuse clouds (e.g., the CNM and WNM) the background radiation field is dominated by direct starlight, while in dense clouds (e.g., giant molecular clouds) the radiation field is dominated by re-processed stellar emission via thermal dust radiation in the far-infrared.

10.2.2 Collisional Heating

Collisional heating of dust grains can occur if the dust grains get energy via frequent collisions with atoms/electrons/ions.

As the heating is due to collisions, the heating rate will be proportional to the collisional cross-section of the dust (πa^2), the typical collision speed, and the energy difference between the gas and the dust. Heating only occurs if the gas is hotter than the dust, however at the densities of most of the ISM (e.g., the CNM, WNM, and some of the molecular gas) this is expected. An expression for this heating rate is given by:

$$\left(\frac{dE}{dt} \right) = \sum_i n_i \left(\frac{8k_B T_{gas}}{\pi m_i} \right)^{1/2} \pi a^2 \times \alpha_i \times 2k_B (T_{gas} - T_{dust})$$

where the sum is over all possible collision partners.

In very dense molecular clouds, thermal coupling of gas and dust (due to collisions) becomes important at typical densities of $n \sim 10^4 \text{ cm}^{-3}$. However, because dust grains are so efficient at radiating energy (see discussion of cooling below), the net effect is only a small increase in the dust grain temperature, and a much more significant drop in the gas temperature (see the following figure from Goldsmith 2001)

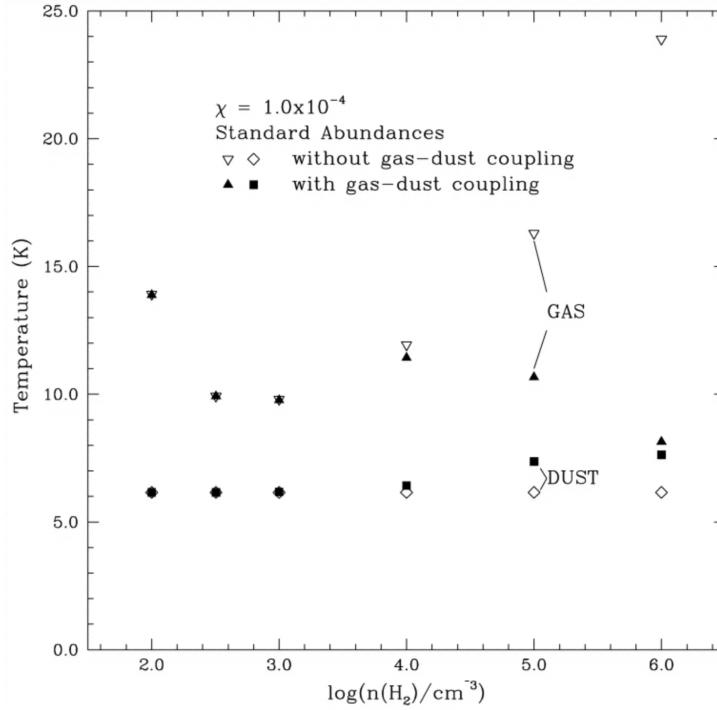


Figure 10.7: Goldsmith_2001.png

10.3 Dust Cooling and Emission

10.3.1 Radiative Cooling

The thermal energy radiated by a dust grain (the radiative cooling rate) can be expressed as:

$$\left(\frac{dE}{dt}\right)_{\text{emiss.}} = 4\pi \int B_\nu(T_{\text{dust}}) C_{\text{abs}}(\nu)$$

Taking this integral over B_ν , and from our definition of C_{abs} , we can write this as

$$\left(\frac{dE}{dt}\right)_{\text{emiss.}} = 4\pi a^2 \langle Q_{\text{abs}} \rangle_{T-\text{dust}} \times \sigma T_{\text{dust}}^4$$

Here $\langle Q_{\text{abs}} \rangle_T$ is

$$\langle Q_{\text{abs}} \rangle_T \equiv \frac{\int B_\nu(T) Q_{\text{abs}}(\nu) d\nu}{\int B_\nu(T) d\nu}$$

Note that it is common to approximate Q_{abs} as a power law:

$$Q_{\text{abs}} = Q_0 \left(\frac{\lambda}{\lambda}\right)^{-\beta}$$

Doing this, $\langle Q_{\text{abs}} \rangle_{T-\text{dust}}$ can be evaluated analytically:

$$\langle Q_{\text{abs}} \rangle_T = \frac{15\pi}{4} \Gamma(4 + \beta) \zeta(4 + \beta) Q_0 \left(\frac{kT}{h\nu_0}\right)^\beta$$

Where Γ is the gamma function and ζ is the Riemann zeta-function.

10.3.2 Dust Emission

The radiative cooling of grains is one of the primary ways that we observe dust, particularly dust in relatively cool, dense regions. The power per unit frequency radiated by a grain is approximately

$$P_\nu = 4\pi B_\nu(T_{dust}) C_{abs}(\nu)$$

where

$$\left(\frac{dE}{dt} \right)_{emiss.} = \int P_\nu d\nu$$

We can re-write this in terms of Q_{abs} :

$$P_\nu = 4\pi^2 a^2 B_\nu(T_{dust}) Q_{abs}$$

10.3.3 Steady-State Temperature

The equilibrium grain temperature (T_{eq}) can be determined by equating the heating and cooling rates. Ignoring collisional heating, we can write:

$$4\pi a^2 \langle Q_{abs} \rangle_{Teq} \times \sigma T_{eq}^4 = \pi a^2 \langle Q_{abs} \rangle_* u_* c$$

Assuming Q_{abs} is a power law in the infrared:

$$T_{eq} = \left(\frac{h\nu_0}{k} \right)^{\beta/(4+\beta)} \left[\frac{\pi^4 \langle Q_{abs} \rangle_* c}{60\Gamma(4+\beta) \zeta(4+\beta) Q_0 \sigma} \right]^{1/(4+\beta)} u_*^{1/(4+\beta)}$$

This yields a model-dependent solution of

$$T_{eq} = T_{const}(\text{K}) \left(\frac{a}{0.1 \mu\text{m}} \right)^{-\alpha} U^{1/6}$$

$\alpha = (1/15)$ for silicates, with a constant of proportionality of $T_{const} = 16.4$ K.

$\alpha = (1/40)$ for graphite, with a constant of proportionality of $T_{const} = 22.3$ K.

10.4 Dust Destruction

Dust forms from the condensation of refractory materials (e.g., like iron, silicates, or carbonaceous materials) – the term refractory here referring to materials with relatively high condensation temperatures (> 1500 K), as opposed to volatiles, which have lower condensation temperatures (< 1000 K). This condensation is especially efficient in environments like the atmospheres of red giant stars, where the dominant chemistry (Oxygen-rich vs. Carbon-rich) determines the typical type of grains that will form (silicates vs. graphites). Recently-formed dust is also seen in ejecta from more massive stars and even in supernova remnants. However, the processes that form the dust in the ISM are also balanced by processes that destroy it.

10.4.1 Sputtering

One of the most common avenues for destroying dust in the diffuse ISM is shocks. For the conditions characteristic of shocked gas (high temperatures, and a sizable difference between typical

dust and gas velocities), dust grains will be bombarded by high-energy atoms and ions. This process is known as sputtering, as such collisions will remove mass from the dust grains, and even at a rate of just one atom lost per collision, this can result in the destruction of dust grains. The amount of mass that is knocked off of a dust grain with each collision is known as the sputtering yield, and is primarily determined through lab experiments. The sputtering rate is expressed as the rate at which dust grain radius decreases, and for $10^5 \text{ K} < T < 10^9 \text{ K}$ can be approximated as:

$$\frac{da}{dt} \sim \frac{10^{-6}}{1 + (T/10^6 \text{ K})^{-3}} \frac{n_H}{\text{cm}^{-3}} \mu\text{m yr}^{-1}$$

Grain lifetimes can then be estimated by comparing this rate to a typical starting grain size. In the typical conditions of a supernova remnant, a grain with a starting radius of $0.1 \mu\text{m}$ would then have a lifetime of only 10^5 years.

It is important to note that the refractory elements (Si, C, Fe) that originally formed the nucleus of these dust grains still remain in the ISM. Some of these are incorporated into molecules that are known as good tracers of the recent passage of strong shocks (like SiO) but most of these will re-condense into dust grains as the gas cools behind the passage of the shock. This is seen in the slight paradox of the role supernovae play in the life cycle of dust: while they launch the shocks that lead to the destruction of dust grains, they can also be dust factories as the post-shock gas in the remnant cools.

10.4.2 Erosion

This is a very similar process to sputtering, but now we are in the regime of much cooler temperatures ($<$ a few hundred K) and much higher densities. This regime is particularly important for the formation of planets, as it applies to conditions in the protostellar disks surrounding stars in the process of formation. Processes of dust grain destruction in these environments are an extremely important topic, as under some conditions they can limit grain growth beyond a certain size, hindering the formation of larger objects like planetesimals.

For erosion, the collisions of interest still involve particles of different sizes, but now they are between very small (“monomer”) dust grains and substantially larger dust grains. If the collision velocities are slow, this just results in coagulation (the small grain will stick to the large grain) and grain growth. However, if collision velocities become too large, these collisions will knock bits off of the larger dust grains, effectively eroding them. If there is a large population of these small grains, erosion can be an important source of dust grain destruction in disks, preventing the efficient formation of larger grains.

10.4.3 Fragmentation

Again, here we concern ourselves with dust-dust collisions at relatively low temperatures and high densities. Now the issue becomes, once dust grains reach a certain size (typically on the order of tens of centimeters, collisions between similar-sized grains are no longer expected to result in the coagulation of the two grains, but rather will be destructive, fragmenting both grains). Fragmentation is believed to be one of the major barriers to grain growth in disks, and much research currently goes into finding processes that can circumvent fragmentation and allow for the formation of large objects like asteroids, planetesimals, and planets that are commonly observed around stars.

10.5 Drag and Radial Drift

10.5.1 Drag and the Meter-Sized Barrier

Another significant impediment to the growth of large dust grains in disks results from the interactions between the dust and gas in the disk. Any time that the dust and gas grain velocities differ, dust grains will experience a drag force that leads to an overall deceleration. In a protostellar disk, such a velocity difference is expected because dust grains and gas particles experience different forces: a freely-moving dust grain will experience only gravitational forces and so will assume a Keplerian orbit, while a gas particle will experience both an inward gravitational force and an outward pressure force, and so will have a speed slightly less than that of Keplerian motion at the same radius. This velocity difference will then lead to deceleration of the dust grain, which will lose angular momentum and spiral inward to smaller radii. A general time scale describing how drag on a dust grain will slow its speed is given as:

$$\tau_{drag} = 1.1 \times 10^5 \text{ yrs} \left(\frac{\rho}{3 \text{ g/cm}^3} \right) \left(\frac{a}{10^{-5} \text{ m}} \right) \left(\frac{30 \text{ cm}^{-3}}{n_H} \right)^2 \frac{T}{100 \text{ K}}$$

The radial drift velocity for grains in a protostellar disk will be largest for meter-sized dust grains, typically $\sim 100 \text{ m/s}$, making the lifetimes of these large grains short, as they will quickly accrete onto the central protostar.

10.5.2 Poynting-Robertson Effect

Less important in protostellar disks, but important in more evolved stellar systems (for example, in debris disks, or for dust around main-sequence or white dwarf stars), is another effect that leads to inward migration of dust grains: Poynting-Robertson drag.

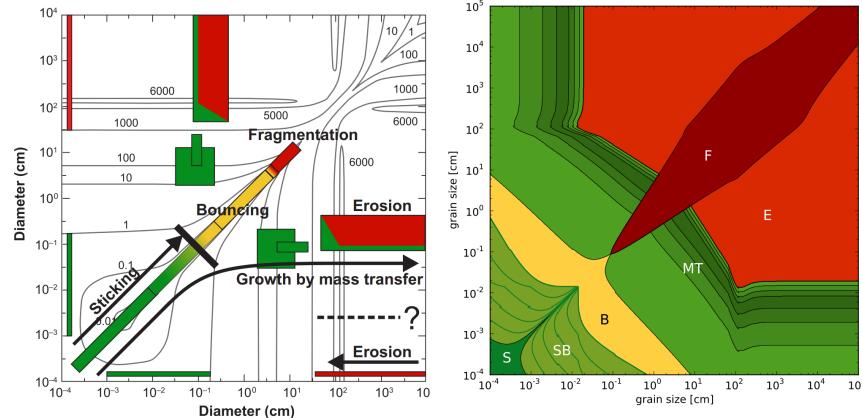
This is the result of a torque on the dust grain from the radiation field of a star, which reduces the orbital angular momentum of the grain, leading its orbit to decay and spiral closer to the star. The time scale on which this occurs is a function of the radiation field, the grain size, and the distance from the star, and can be expressed as:

$$\tau_{PR} = 8.3 \times 10^7 \text{ yrs} \left(\frac{\rho}{3 \text{ g/cm}^3} \right) \left(\frac{a}{\text{cm}} \right) \left(\frac{R}{\text{AU}} \right)^2 \frac{1}{\langle Q_{pr} \rangle} \frac{L_\odot}{L_*}$$

where ρ is the mass density of the dust grain, a is the effective dust grain radius, R is the distance from the star, L_* is the stellar luminosity, and $\langle Q_{pr} \rangle$ is an efficiency factor for the radiation pressure that can generally be assumed to be 1.

10.6 Additional Barriers to Grain Growth

Fragmentation and radial drift are two of the most significant barriers to grain growth and the formation of large dust grains in protoplanetary disks, however they are not the only barriers that are present. This figure, from Testi et al. 2014 (a review chapter from the Protostars and Planets VI conference) illustrates some of these barriers and regimes in which they are important. On the left the colored areas show the results of laboratory experiments, and on the right these can be compared with the results of a numerical model:

Figure 10.8: *grain_growth.png*

10.6.1 Bouncing Barrier

For interactions between similar-sized dust grains there is a regime in-between coagulation and fragmentation, where typical collisions are nearly elastic. The grains do not fragment, but they also do not stick together, and so do not grow larger. One possible way around this barrier (and likely the fragmentation barrier as well) is for grain growth to proceed primarily through the interactions of unequal-sized particles. Another way around this is to consider that for realistic velocity distributions, some particles of this size will collide at lower speeds than the RMS velocity, and so sticking can still occur.

10.6.2 Charging Barrier

This barrier is suggested to be important for relatively small dust grains, which have accumulated a charge (e.g, through frequent collisions with free electrons, see Draine 25.1). Charged grains would repel each other, which would act to prevent the close collisions that result in grain growth. However, it is currently thought that this barrier acts to slow but not ultimately prevent grain growth at these sizes.

10.7 Theories for Grain Growth

10.7.1 Fragmentation + Mass Transfer

If a disk can maintain a sufficient number of small particles, it may be possible to get around fragmentation if small grains collide with larger grains at moderate speeds of 25-50 m/s. Such collisions will result both in some transfer of mass to the larger grains, and will contribute to the maintenance of a smaller grain population, as the smaller-sized grain partially fragments as a result of the collision.

10.7.2 Dust Traps

One of the most active areas of both theoretical and observational research into dust grain growth and planetesimal formation is into ways to overcome the meter-sized barrier that results from radial drift. A popular idea is that disk inhomogeneities and small-scale structures might be able to effectively trap large dust grains in place, preventing their inspiral and allowing them to grow.

Examples of these structures include vortices or pressure traps, and there has been some observational confirmation of their existence through ALMA observations of asymmetric dust emission in disks:

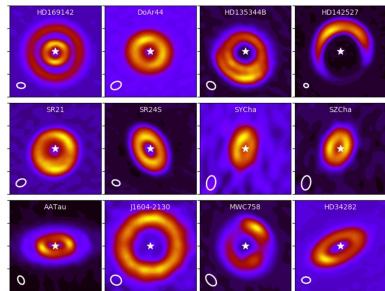


Figure 10.9: *MarelEtAl.jpg*

Chapter 11

The Neutral ISM

Table 1: The Milky Way Interstellar Medium

	Molecular Clouds	Cold Neutral Medium	Warm Neutral Medium	Warm Ionized Medium	Coronal Gas
Temperatures (K)	10-20	50-100	500-7000	8000-10,000	10^6-10^7
Densities (cm^{-3})	$10^2 - 10^6$	20-50	0.1-0.6	0.2-0.5	$10^{-3} - 10^{-2}$
Type of Hydrogen	H ₂	H I	H I	H II	H II
Scale					
Height (pc)	50-75	90-250	130-400	800-1800	>1000
Volume Filling Fraction	< 1%	1-5%	30-50%	~30%	30-70%
Observed Wavelengths	Radio/ Millimeter	Radio (21 cm)	Radio (21 cm)	Optical (H α)	X-ray
Primary Coolants	CO dust	C ⁺	H ₂ , C ⁺ Fine-structure	Ly α Metal ions	Fe Ions Free-free

Sources: Dalgarno & McCray (1972)
Brinks (1990)
Ferriere (2001)
Gaensler et al. (2008)

Figure 11.1: ISM.png

In the Milky Way, the relative amounts of these ISM components by mass are: * Molecular gas: $\sim 8 \times 10^8 M_{\odot}$ (3%) * Cold Neutral Medium: $\sim 1.5 \times 10^9 M_{\odot}$ (6%) * Warm Neutral Medium: $\sim 1.5 \times 10^9 M_{\odot}$ (6%) * Warm Ionized Medium: $\sim 10^9 M_{\odot}$ (4%) * Coronal Gas: $\sim 2 \times 10^{10} M_{\odot}$ (81%)

Recently we have focused a lot on HII regions and the WIM. This gas is important because it is a highly-visible tracer of the amount/location of recent star formation in a galaxy.

In contrast, the neutral gas tells us about reservoir of gas left in a galaxy for future star formation. It is most typical to see substantial amounts of both neutral and ionized gas in late-type (disky,

spiral) galaxies, as more evolved galaxies have exhausted this reservoir, and are primarily dominated by stars. Historically, this gas is an important kinematic tracer of galaxy mass and size because neutral gas extends to much larger radii than stars/star formation. Both the WNM and the WIM take up a substantial fraction of the volume of the Milky Way disk.

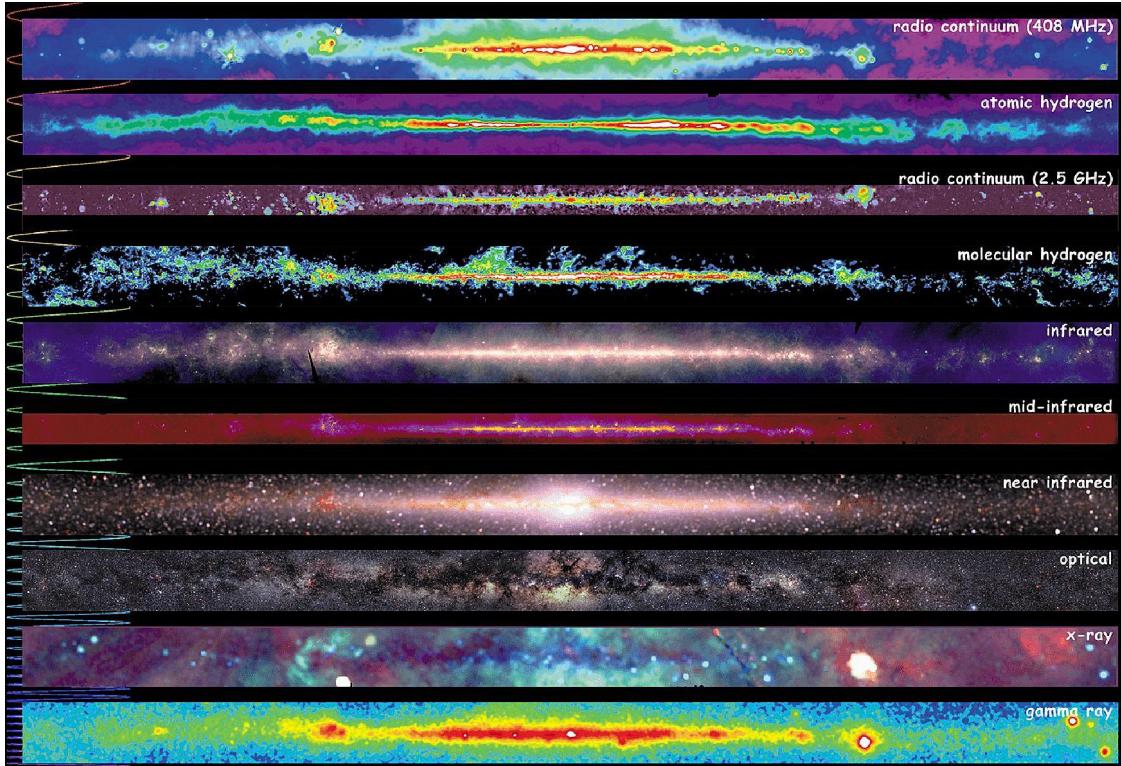


Figure 11.2: MW_plane.jpeg

11.1 21 cm emission and absorption

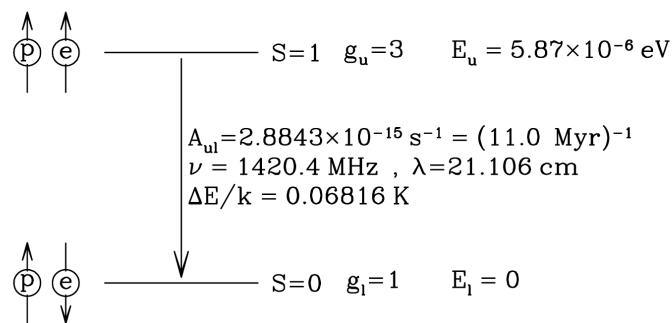


Figure 8.1 Hyperfine splitting of the $1s$ ground state of atomic H (Gould 1994).

Figure 11.3: spinflip.png

The direct detection of neutral Hydrogen is made possible by the existence of hyperfine structure in the ground state of neutral hydrogen. This is the result of the (incredibly small!) energy difference in the states in which the spin of the electron is aligned/anti-aligned with the spin of the proton, due to the interaction of their magnetic moments.

Emission of a photon when an electron changes its spin from aligned to anti-aligned occurs at a wavelength of 21 cm, and is sometimes referred to as the hydrogen “spinflip” transition.

There are a number of important properties of this emission. Firstly, the energy splitting between these states is extremely small:

$$E = \frac{hc}{\lambda} = \frac{2.998 \times 10^{10} (6.626 \times 10^{-34})}{21} = 9.45 \times 10^{-25} \text{ J, or in terms of temperature } E = kT, T = 0.068 \text{ K}$$

This means that the cosmic microwave background radiation with $T = 2.7 \text{ K}$ will be enough to excite hydrogen atoms into the upper spin state.

We can calculate the expected population of the upper (spin aligned) and lower (spin anti-aligned) states using Boltzmann statistics:

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} \exp\left(-\frac{0.068K}{T_{spin}}\right)$$

Since any reasonable temperature distribution for the gas ($T_{spin} > T_{CMB}$) will result in an exponential term that is ~ 1 , the relative population of the two states will simply be set by their degeneracies.

Recalling that $g = 2(S + 1)$, the lower state with $S = 0$ will have $g_l = 1$ and the upper state with $S = 1$ will have $g_u = 3$ and thus

$$\frac{n_u}{n_l} = 3.$$

Given the relatively high population of the upper energy level compared to the lower energy level (the spin ground state), stimulated emission will be an important consideration.

As the population of the upper and lower states is largely insensitive to the spin temperature, the emissivity of HI also does not depend on the spin temperature:

$$j_\nu = \frac{3}{16\pi} A_{ul} h \nu_{ul} n(H) \phi_\nu$$

where ϕ_ν is the observed line profile as a function of frequency, normalized so that $\int \phi_\nu d\nu = 1$

The observed intensity of HI emission is:

$$I_\nu = I_0 + \int j_\nu ds$$

which, recalling the definition of column density to be $\int nds = N$, means that the background-subtracted line intensity is directly proportional to the column density (essentially: the total number of hydrogen atoms along the line of sight):

$$I_\nu - I_0 = \frac{3}{16\pi} A_{ul} h \nu_{ul} N(H) \phi_\nu$$

11.2 HI Heating

Both cosmic ray heating and dust play extremely important roles in heating the neutral ISM. We will talk more about dust photoelectric heating in Section 10.2.

11.2.1 Cosmic Ray Heating

Like our derivation of the photoionization heating rate, the cosmic ray heating rate is a function of the cosmic ray ionization rate ζ_{CR} :

$$\Gamma_{CR} \sim [n_H + n_{He}] E_{heat}$$

Here E_{heat} is the amount of energy from the cosmic ray that goes into heating the gas. This is a function of the ionization fraction of the gas ($x_e = n_e/n_H$) and the typical cosmic ray energy:

$$E_{heat} \sim 6.5 \text{ eV} + 26.4 \text{ eV} \left(\frac{x_e}{x_e + 0.7} \right)^{0.5}$$

Cosmic rays also interact with free electrons, where

$$\Gamma_{CR,e} \propto n_e \zeta_{CR}$$

The constant of proportionality depends on the energy spectrum of the cosmic rays, but can be assumed to be on the order of 5×10^{-10} erg.

Typical values of ζ_{CR} are a few 10^{-16} s^{-1} .

11.2.2 Photoionization Heating

For fun, we will also consider a toy model in which the ionization of neutral carbon is the dominant heating source for the CNM. As we have discussed, this will ultimately heat the gas through the production of free electrons.

We can express the volumetric heating rate from the photoionization of carbon as:

$$\Gamma_{pi,C} = n_C \zeta_{pi,C} \langle E_{pe} \rangle$$

Once again, we will do our usual trick of assuming the region we are considering is in ionization equilibrium, so that we do not have to calculate $\zeta_{pi,C}$, which is a nasty function of things like the strength and spectrum of background radiation field. Assuming

$$n_C \zeta_{pi,C} = n_{C^+} n_e \alpha_C.$$

we can then just write

$$\Gamma_{pi,C} = n_{C^+} n_e \alpha_C \langle E_{pe} \rangle$$

To find the heating rate, we need to know the average energy $\langle E_{pe} \rangle$ that is imparted to the ejected photoelectron by a single photoionization. Here, we will note that there is a limited range of photon energies that contribute to the ionization of carbon, as photons with energies higher than 13.6 eV will simply ionize hydrogen instead. Assuming a ~ 50

$$\langle E_{pe} \rangle \sim 0.5(13.6 \text{ eV} - 11.3 \text{ eV}) \sim 1.2 \text{ eV}$$

11.3 HI Cooling

Here, species like C⁺ and O actually do play a dominant role in cooling the neutral ISM.

We can write an expression for the cooling rate of C⁺ (due to line emission following collisional excitation) as:

$$\Lambda_{ce}(C^+) = 8 \times 10^{-33} n_{C^+} n_e T_e^{-0.5} \exp\left(-\frac{92 \text{ K}}{T_e}\right) \text{ J m}^{-3} \text{ s}^{-1}$$

Note this is a very similar form as the expression for the cooling rate of O⁺ from the previous lecture:

$$\Lambda_{ce}(O^+) = 1.833 \times 10^{-30} y_{O^+} n_O n_e T_e^{-0.5} \exp\left(-\frac{3.89 \times 10^4}{T_e}\right) \text{ J m}^{-3} \text{ s}^{-1}$$

11.4 The Two-Phase ISM

The simultaneous existence of both a warm neutral medium (WNM) and a cold neutral medium (CNM) phase occurs due to the temperature dependence of the heating and cooling rates and because the temperatures and densities at which one encounters HI span a large range.

If the HI gas is in pressure equilibrium, then there exists a range of pressures where there are two equilibrium temperatures for low-density and high-density HI. Between these values, the temperature-dependence of the relevant heating and cooling rates causes gas at these temperatures to be driven toward the equilibrium values, as the temperature and density are linked, such that lowering the temperature increases the density, and vice versa.

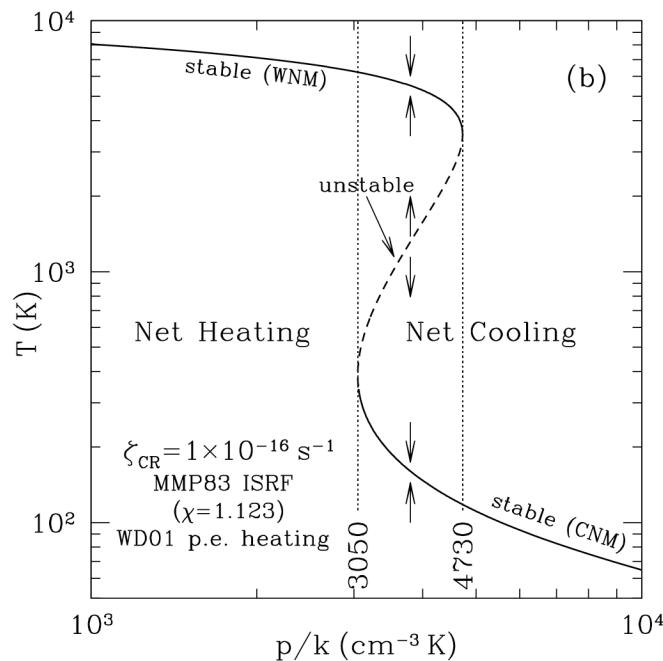


Figure 11.4: 2phase.png

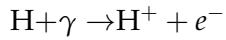
Chapter 12

Ionized Gas Processes

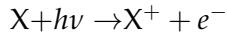
12.1 Ionization

12.1.1 Photoionization

We can express the "reaction" leading to photoionization as:



or alternatively / more generally (as in Draine)



This reaction can be characterized by a photoionization rate, which we can see from the reactants involved will depend on the density of ionizing photons - ionizing meaning $E > 13.6$ eV for hydrogen - and the photoelectric cross section of the atom.

We will start by defining the photoelectric cross section. Note that for hydrogen and hydrogen-like (one electron) ions, this can be derived analytically and written exactly as:

$$\sigma_{pe}(\nu) = \sigma_0 \left(\frac{Z^2 I_H}{h\nu} \right) \frac{e^{4-4\arctan(x)/x}}{1 - e^{-2\pi/x}}$$

where x is defined as

$$x = \sqrt{\frac{h\nu}{Z^2 I_H}} - 1$$

and the cross section exactly at the energy of photoionization is given by

$$\sigma_0 \equiv \frac{2^9 \pi}{3e^4} Z^{-2} \alpha \pi a_0^2 \text{ cm}^2$$

Here e is the number e not the electron charge, and α is the fine structure constant, not the recombination coefficient. The cross section σ_{pe} that we defined in this way also has units of area (e.g., cm^2). We can use this definition of the cross section to write a rate (photoionizations per time) of

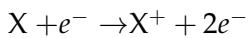
$$\zeta_{pi} = \int_{\nu_1}^{\infty} \sigma_{pe}(\nu) c \frac{u_{\nu}}{h\nu} d\nu$$

Note that as this is just a rate per time, to express this as a rate of photoionizations per time per unit volume (of neutral gas) you would multiply by n_H (or n_X , where X is the neutral species of interest). This would make this rate more directly comparable with some of the other rates (e.g., collision rates) that we have previously defined.

12.1.2 Collisional Ionization

Of course photons are not the only things that atoms encounter that can impart the energy necessary for ionization. While we will discuss collisional ionization due to shocks in neutral/molecular gas later on this semester, of particular interest for plasmas is ionization due to collision with energetic free electrons.

We can write such a reaction as follows:



Reaching back to our early definitions of collisional rates, we recall that generally we can express a collisional rate as $k = \langle \sigma v \rangle$, with units of volume/time. Here σ is a cross section for the interaction, and v is the typical interaction velocity.

We can similarly describe the rate at which ionizing collisions occur as:

$$k_{ci} = \int_I^\infty \sigma_{ci} v f_E dE$$

Assuming that the flux f_E of incident electrons with energy E comes from a thermal distribution we can rewrite this as:

$$k_{ci} = \left(\frac{8kT}{\pi m_e} \right)^{1/2} \int_I^\infty \sigma_{ci}(E) \frac{E}{kT} e^{-E/kT} \frac{dE}{kT}$$

Here, the term outside of the integral comes from the expected thermal velocity of electrons in the plasma.

Note that we must do an integral over the distribution of energies (expressed here in terms of E/kT) both because the cross-sections are energy dependent and because only sufficiently energetic electrons (that can impart energies larger than the ionization threshold I) are considered.

Recall that a collisional ionization rate per time (used for example to calculate how long it would take for a single atom of X to be ionized) would be given by $k_{ci} n_e$. A collisional rate per time per unit volume of neutral gas would be given by $k_{ci} n_e n_X$

12.1.3 Cosmic Ray Ionization

In addition to collisions with thermal electrons, an atom may undergo collisions with highly energetic electrons or ions, known as cosmic rays:



In many ways, this is a similar mechanism as just discussed above for collisional ionization. The difference is the energy spectrum of the colliding particles: while the velocities/kinetic energies of electrons in the thermal plasma can be assumed to be well-described with a Maxwell-Boltzmann

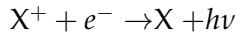
relation, cosmic rays have a nonthermal distribution of energies. This spectrum is much less well-characterized, which leads to some of the uncertain assumptions that lead to the relations given in Draine for the cross-section and ionization rate.

The details of these particular approximations are not so important. What is useful to note is just that again we follow the characteristic form of defining an interaction cross section (here, just the collisional cross section or σ_{ci}) as well as a primary ionization rate ζ_{CR} . We then see the first full expression given in this chapter for an ionization rate per unit volume:

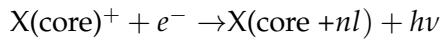
$$\left(\frac{dn_e}{dt} \right) \sim 1.1 n_H \zeta_{CR} \left(1 - \frac{x_e}{1.2} \right) [1 + \phi_s]$$

Note that this includes a term ϕ_s for secondary ionizations (if the initial CR leads to ejection of an electron that itself is sufficiently energetic to ionize another atom)

12.2 Recombination



More specifically, Draine write this process as



where the "core" is an initial bound configuration of electrons, and the captured electron occupies a bound energy state nl .

From the reactants on the left hand side of this expression, we can see that the recombination rate will depend on the collisional cross section of H with e^- and the number densities of both species.

Here, the cross section for this electron capture is $\sigma_{rr,nl}$, and the thermal rate coefficient, which describes the likelihood that a free electron will be captured into state nl , resulting in recombination and the emission of a photon, can be written as:

$$\alpha_{nl} = \left(\frac{8kT}{\pi m_e} \right)^{1/2} \int_0^\infty \sigma_{rr,nl}(E) \frac{E}{kT} e^{-E/kT} \frac{dE}{kT}$$

Note that more broadly, the recombination process is generally a cascade of emission resulting from the radiative decay of the initial bound state resulting from the initial electron capture that can proceed all the way down to the ground state of the atom. The spectrum of photons emitted through this process is generally described as recombination emission, or recombination lines.

12.2.1 Case A Recombination

In Case A recombination, one assumes that photons emitted as a result of recombination are able to escape from the neutral gas cloud. This is the optically-thin limit.

For Case A recombination, the total rate coefficient for radiative capture of an electron can be obtained by summing over the rates for recombination into all possible levels nl .

In many cases, an excited state of an atom resulting from recombination has several options as to how it can decay to a lower-excitation state. One must then employ branching ratios to determine the likelihood that a recombining atom will follow a decay cascade that takes it through a specific level $n'l'$.

12.2.2 Case B Recombination and Lyman α Resonant Scattering

Case B recombination treats the scenario when all of the photons cannot be assumed to freely escape, and those with energy instead have a significant probability of being reabsorbed by a nearby neutral atom. The levels affected are those corresponding to recombinations directly to $n = 1$ or the Lyman series, which occurs in the UV. Photons emitted in these $n \rightarrow 1$ transitions as a result of recombination can essentially be assumed to be immediately reabsorbed by a nearby neutral atom (which then becomes ionized). As a result, such “recombinations” do not lead to any net change in the overall number of ionized atoms in the gas. To determine the total rate coefficient for radiative capture of electrons in Case B, one then instead sums over all levels nl *except* those with $n = 1$.

This process of emission and reabsorption of Lyman photons is known as resonant scattering, and we will work through an example of this scattering below.

In general, Case B recombination is a good approximation to most situations where the gas is photoionized (e.g., HII regions around stars). Case A recombination is most applicable to collisionally-ionized (shock-heated) regions.

Example: Draine 14.2

The Einstein A coefficients for all of the allowed transitions of hydrogen from levels $n \leq 3$ are given in the following table:

u	ℓ	$A_{u\ell}(\text{s}^{-1})$	$\lambda_{u\ell}(\text{\AA})$
3d	2p	6.465×10^7	6564.6
3p	2s	2.245×10^7	6564.6
3s	2p	6.313×10^6	6564.6
3p	1s	1.672×10^8	1025.7
2p	1s	6.265×10^8	1215.7

Figure 12.1: Ly_B.png

(a) Consider a hydrogen atom in the $3p$ state as the result of radiative recombination: $p + e^- \rightarrow H(3p)$. What is the probability p_β that this atom will emit a Lyman β photon?

To solve this, we need to calculate the branching ratio:

$$\frac{A_{3p \rightarrow 2s}}{A_{3p \rightarrow 1s}} = \frac{2.245 \times 10^7 \text{ s}^{-1}}{1.672 \times 10^8 \text{ s}^{-1}} = 0.134$$

This is the probability of emitting $H\alpha$; the probability of emitting $Ly\beta$ is ~ 87

(b) In an HI region where hydrogen is the only important opacity source, and averaged over many atoms “prepared” in the $3p$ state, what is the mean number $\langle n_i \rangle$ of times a Lyman- β photon is “scattered” (that is, absorbed and then re-emitted) before an $H\alpha$ photon is emitted?

Here, since we are dealing with scattering, note that we are implicitly assuming Case B recombination.

We are looking for the expectation value $E(n)$ of the number of scatterings n that it takes before an $H\alpha$ photon is emitted (and escapes).

We will call the probability that an $H\alpha$ photon is emitted p , where $p = 0.134$. The probability of instead emitting a $Ly\beta$ photon is $q = 1 - p = 0.866$.

Note that the probability of emitting an $H\alpha$ photon initially is p , after one scattering it is pq , and after two scatterings it is pq^2 so that we can generally express the probability of emitting a photon after n scatterings as pq^n .

We can then write the expectation value of n as

$$E(n) = \sum npq^n$$

Using the result helpfully provided in Draine that

$$\sum nq^n = \frac{q}{(1-q)^2}$$

We can say that

$$E(n) = \frac{pq}{(1-q)^2} = \frac{(0.134)(0.866)}{(1-0.866)^2} = 6.5$$

So, typically there are 6-7 scatterings before $H\alpha$ is emitted.

Note that one result of this resonant scattering of $Ly\beta$ photons is to enhance the brightness of the $H\alpha$ line above what would be expected for Case A recombination.

12.3 Heating and Cooling Processes

Why do we care about cooling the gas in the ISM?

- We have to cool gas if we want it to contract and form stars. Otherwise the thermal motion of the gas will support it against gravitational collapse.
- We also need to cool gas if we want to form molecules.

What are some ways that we can cool a gas?

Ultimately, we need to get energy out of it. How can we do this?

- Conduction and Convection: Though effective in terrestrial conditions (or in stars), these are not efficient at the densities we are considering here.
- Radiation: Giving off photons via spontaneous emission, recombination, or free-free emission.
- Collision-Excitation-Radiation: Gas particles collide with other gas particles, and so the kinetic energy of the gas is transferred to internal energy states, which then usually decay quickly by emitting a photon. We will also talk later about collisions with dust, which it turns out can generally cool more efficiently than the gas via the emission of continuum radiation.

To be efficient, Collision-Excitation-Radiation cooling requires the following:

1. Collisions must be frequent, not requiring a low-abundance partner (say, F)

2. The excitation energy must be comparable to or less than the thermal kinetic energy of the gas particles.
3. There must be a high probability of excitation from collision (generally this can be assumed to be true)
4. There must be a high probability of emitting a photon after being excited
5. The gas must be optically thin so that the photons are not immediately re-absorbed

And so, we also see a practical reason to care about gas cooling as observers: the same processes that are largely responsible for cooling gas (radiating away photons) are also the processes that let us directly observe the interstellar gas.

Why do we care about how gas in the ISM gets heated?

- Heating gas can make it less likely that it will form stars (feedback)
- Heat it enough and the gas particles can actually be moving so fast that they will escape the galaxy
- Heating the gas can greatly change its composition (e.g., more energy to remove molecules and ices from dust grain surfaces will greatly change the gas chemistry)

Macroscopically, what can be responsible for heating the gas?

- Stars (photons and direct injection of momentum from winds)
- Supernova remnants (cosmic rays, high-energy photons, and shocks)
- Supermassive black holes (release of gravitational potential energy into X-rays)

Microscopically, how does heating happen?

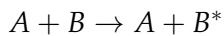
Here, I want to call attention to a fundamental difference between heating and cooling, as these processes can look somewhat similar. Situations in which a gas particle gives off a photon almost always results in cooling, because gas is typically optically thin (at least at the long wavelengths at which molecules tend to give off cooling radiation). Any photons emitted are unlikely to interact, and have a large probability of escape. To be effective, a heating process then generally needs to result in an increase of the bulk kinetic energy (not just energy of internal states, which tends to be quickly radiated away).

This is generally accomplished by the incidence of very high-energy photons knocking SOMETHING off of a particle. This can be:

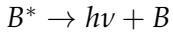
- Photoionization, creating free electrons having excess energy above the energy used to free them from the atom (the ionization potential)
- Photodissociation, breaking a chemical bond and increasing the energy of the remaining particles
- Photoelectric heating in which photons incident on dust grains also release free electrons.

Important reactions for the balance of heating and cooling in HII regions:

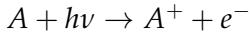
(1) Collision: energy goes into internal energy of particle



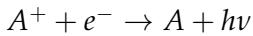
(2) De-excitation: collision energy is quickly radiated away cooling the gas



(3) Ionization: results in a free electron, which heats the gas



(4) Recombination: will cool the gas



12.3.1 HII Region Heating

Here, we will focus on photoionization, which ultimately heats the gas through the production of free electrons (and subsequent elastic collisions of these electrons with other species).

Not surprisingly, the heating rate due to photoionization relies on another rate, which we have seen before: the photoionization rate!

The rate of photoionizations per time is given as:

$$\zeta_{pi} = \int_{\nu_1}^{\infty} \sigma_{pe}(\nu) c \frac{u_\nu}{h\nu} d\nu$$

We previously wrote the ionization rate (photoionizations per volume per time) for species X as:

$$\frac{dn_e}{dt} = n_X \zeta_{pi}$$

Using this, we can define a volumetric heating rate (energy per unit volume and time) based on the ionization of atom X. We will call this heating rate Γ , with units of energy per volume per time. To determine the net energy gain from ionization, we have to remember that some of the photon energy went into the ionization potential, so we define the energy gain from photoionization as:

$$E_{pi} = h\nu - I$$

Where I is the ionization potential (13.6 eV for hydrogen).

We can then express the rate of energy input (heating) per time per unit volume (the volumetric heating rate) as:

$$\Gamma_{pi} = n_X \int_{\nu_1}^{\infty} \sigma_{pe}(\nu) c \frac{u_\nu}{h\nu} E_{pi} d\nu$$

To find the heating rate, we need to know the average energy of a single photoionization. Draine gives a more general/exact treatment of this by defining a variable ϕ where:

$$\phi = \frac{E_{pi}}{k_B T_{eff}}$$

where T_{eff} is the blackbody temperature of the star.

However, since ϕ is of order unity, it is fine for our purposes to just approximate

$$E_{pi} \sim k_B T_{eff}$$

12.3.2 HII Region Cooling

The cooling rate due to recombination can be similarly determined.

Recall that we wrote the recombination rate (recombinations per volume per time) for species X as:

$$\frac{dn_e}{dt} = -n_{X^+} n_e \alpha_B$$

where α_B is the (case B) recombination rate, which for a species with charge Z can be approximated as

$$\alpha_B = 2.54 \times 10^{-13} Z^2 \left(\frac{T_4}{Z^2} \right)^{-0.8163 - 0.0208 \ln(T_4/Z^2)}$$

The volumetric cooling rate due to recombination can be written as

$$\Lambda_{rr} = -n_{X^+} n_e \alpha_B \langle E_{rr} \rangle$$

where $\langle E_{rr} \rangle$ is the average energy removed from the gas via recombination.

Similar to the determination of the typical photoionization energy, Draine gives a more exact treatment of this value for various assumptions. In general though,

$$\langle E_{rr} \rangle \sim k_B T_e$$

This is just an approximate electron kinetic energy (modulo some additional likelihood that a given interaction with an electron will result in a recombination). Taking the cross-section for recombination properly into account yields, for case B recombination of hydrogen:

$$\langle E_{rr} \rangle = \left[0.684 - 0.0416 \ln \left(\frac{1}{Z^2} \frac{T_e}{10^4 \text{ K}} \right) \right] k_B T_e$$

12.3.3 Thermal Equilibrium

What happens if we only consider cooling due to recombination, and set $\Gamma_{pi} = \Lambda_{rr}$ to define an equilibrium electron temperature for the nebula?

Assuming that the nebula is in ionization equilibrium,

$$n_X \zeta_{pi} = n_{X^+} n_e \alpha_B.$$

We can then express the heating rate as:

$$n_X \zeta_{pi} k_B T_{\text{eff}} = n_{X^+} n_e \alpha_B k_B T_{\text{eff}}$$

And the cooling rate as:

$$\Lambda_{rr} = n_{X^+} n_e \alpha_B k_B T_e$$

we would find $T_e \sim T_{\text{eff}}$.

This then predicts that the temperature of a pure-hydrogen HII region should match that of a typical O star: 30,000 - 60,000 K. However, we observe them to be much cooler! What's up with that?

Answer: our heating rate isn't wrong (we understand the energy spectra of stars pretty well) but our cooling rate ***IS*** wrong: HII regions are not just hydrogen, and line cooling from other ions (He, O, N, etc.) is super-important.

This means if we want to find an equilibrium temperature, we need a more complete expression for the cooling rate. Actually, one must incorporate two effects: the loss of energy due to free-free continuum emission as well as the emission from excited ions of other species.

We can express the per-electron rate of energy loss from free-free emission for each electron created as a result of photoionization as:

$$\Lambda_{ff} = \langle E_{ff} \rangle n_X \zeta_{pi}$$

And if the nebula is in ionization equilibrium, this is the same as

$$\Lambda_{ff} = \langle E_{ff} \rangle n_{X^+} n_e \alpha_B$$

For free-free emission, we can write the average energy emitted by an electron as

$$\langle E_{ff} \rangle = 0.54 \left(\frac{T}{10^4 \text{ K}} \right)^{0.37} k_B T_e$$

Generally, the cooling rate from collisionally-excited line radiation from all species X with excitation states *i* is

$$\Lambda_{ce} = \sum_X \sum_i n(X, i) \sum_{j < i} A_{ij} (E_i - E_j)$$

For a specific case, we can take O⁺.

$$\Lambda_{ce}(\text{O}^+) = 1.833 \times 10^{-30} y_{\text{O}^+} n_O n_e T_e^{-0.5} \exp\left(-\frac{3.89 \times 10^4}{T_e}\right) \text{ J m}^{-3} \text{ s}^{-1}$$

As before, this is a volumetric cooling rate (energy per time per volume). y_{O^+} is the ionization fraction of oxygen, so that

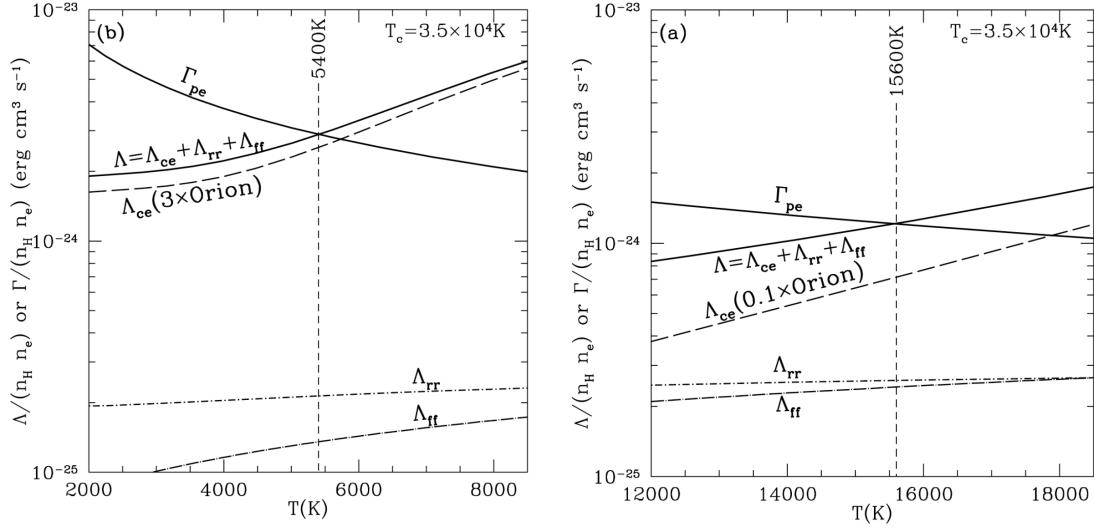
$$n_{\text{O}^+} = y_{\text{O}^+} n_O$$

The rest of the numbers here are just an expression of the temperature-dependent collisional rate coefficient.

Note that this cooling rate will depend strongly on the abundance of oxygen (n_O/n_H), which is set by the overall metallicity of the gas.

To find a better approximation of the equilibrium temperature we would then set $\Gamma_{pi} = \Lambda_{rr} + \Lambda_{ff} + \Lambda_{ce}(\text{O}^+)$.

There are a number of assumptions that can be made to solve this (assume $n_H \sim n_e$ and $y_{\text{O}^+} = 1$) but it still turns out that this equivalency will need to be solved numerically for T_e . **Draine Figure 27.2** shows an example of the equilibrium temperature for the Orion Nebula, taking into account a larger number of cooling lines, and adopting two different metallicities.

Figure 12.2: *thermal_eq.png*

12.4 Nebular Diagnostics

This is one of my favorite parts as an observer, because it is the payoff for all of the detailed treatment we have done of the quantum mechanics that determines the energy level structure of atoms and ions (their energies, angular momenta, and degeneracies), and the rates of excitation and de-excitation of these energy levels due to collisions, photon absorption, and photon emission (and the corresponding radiative transfer involved in these last two).

Thanks to this groundwork, we can start to use these atoms not just as individual laboratories of atomic microphysics, but as probes of their environment: the temperatures, densities, and chemical abundances and ionization fractions in the surrounding gas, as well as the background radiation field.

As we are now looking at things primarily from an observer's perspective, there are a number of practical considerations when choosing the lines that will be useful probes.

First, the species must be reasonably abundant: without a large number of the atom of interest present in the gas, the number of photons coming from its various energy transitions will not be easily detectable.

Second, the transitions must involve energy levels that are populated under the current gas conditions (For example, in an HII region, energy levels that are $\gg 10^4 \text{ K}$ are unlikely to be populated, given a thermal distribution of atoms, ions, and electrons).

Third, these energy levels must have radiative transitions that allow for their observation. As we discussed previously, most of these "allowed" and so strong transitions will be electric dipole transitions with relatively large Einstein A values, though some weaker electric quadrupole or magnetic dipole transitions can also be observed as "forbidden" or "semi-forbidden" lines.

Taking all of these into account, we will focus on three types of common diagnostics: temperature diagnostics, density diagnostics, and ionization/excitation diagnostics that probe the hardness of the background radiation field.

12.4.1 Temperature-Sensitive Line Ratios

For a pair of lines to give a good measurement of the gas kinetic temperature, the most important requirement is the both lines be primarily collisionally excited ($C_{lu} > B_{lu}$). If, instead, radiative excitation is important, then the level populations will reflect the radiation temperature more than the kinetic temperature of the overall gas.

The second important requirement is that the densities not be too high. After a collisional excitation to the upper level, we want the level to decay radiatively and not collisionally, so that the number of photons emitted will be proportional to the number of collisions ($A_{ul} > C_{ul}$). We have seen this condition before: this is our definition of the critical density from Section 6.3. (Note that relatively low densities should also ensure that this level can decay radiatively before it would undergo a collision that would further excite it to a higher energy level).

Third, the energy difference kT should be comparable in scale to the kinetic gas temperatures to be probed. (The energies of the individual levels should also not be so high that they are unlikely to be populated by collisions in $T \sim 10^4$ K gas, functionally this typically means $E < 70,000$ K.) The species of interest must also be present in the nebula: if the temperature is too high or too low, one might not have a sufficient abundance of the atom/ion of interest.

Commonly-used lines that fulfil these criterion are atoms or ions with np^2 and np^4 . configurations. The simplest examples of these are 6-electron ($2p^2$) and 8-electron ($2p^4$) ions. As a reminder, the lowest electron orbitals are:

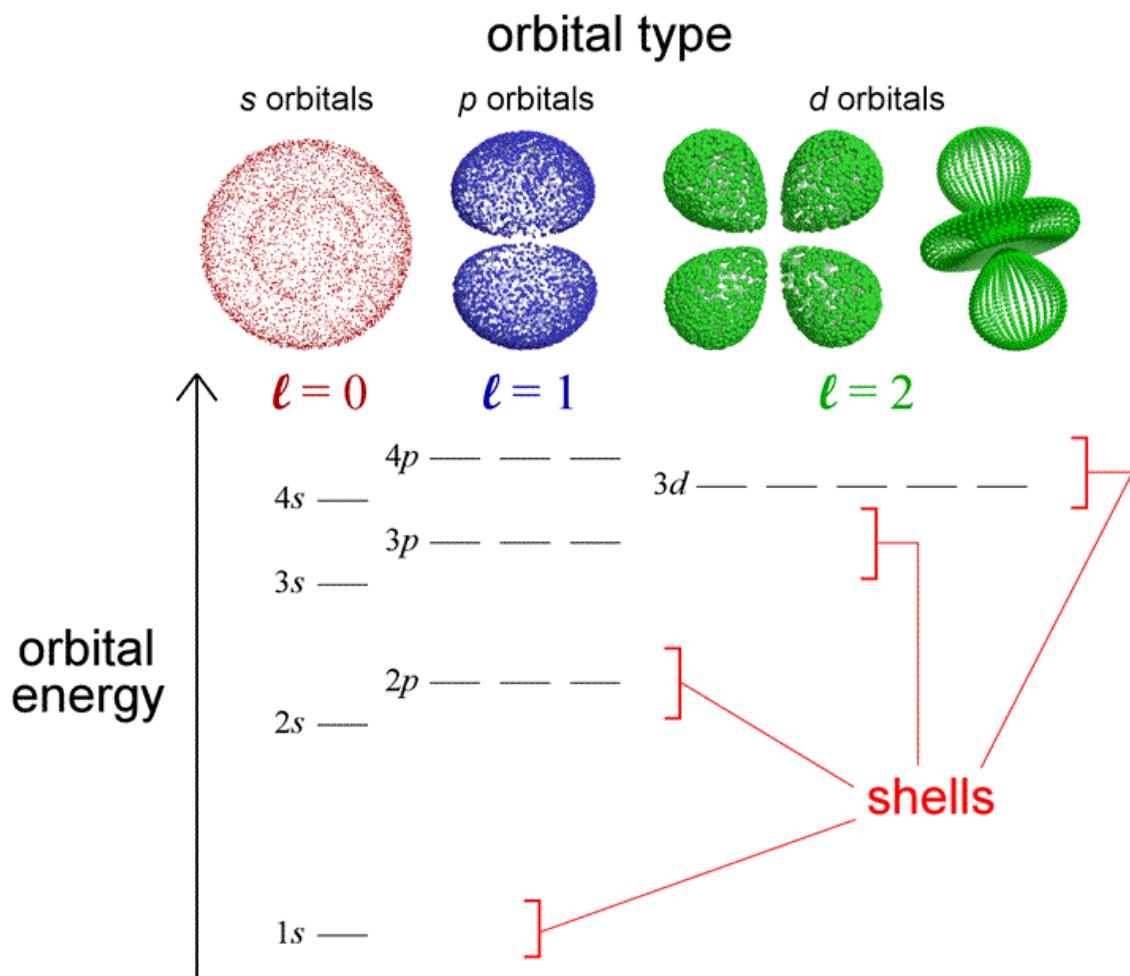


Figure 12.3: orbital_energy_and_type_diagram.png

The 1s and 2s orbitals can hold 2 electrons each, and the 2p orbital can hold 6 electrons.

Two examples of $2p^2$ ions that can be observed in HII regions are NII and OIII.

These ions have a ground state term 3P (as a reminder, the 3 is related to the total spin, $2S + 1$, and P indicates a total orbital angular momentum $L = 1$). The first excited states have terms 1D and 1S , as shown in the energy-level diagram for NII below:

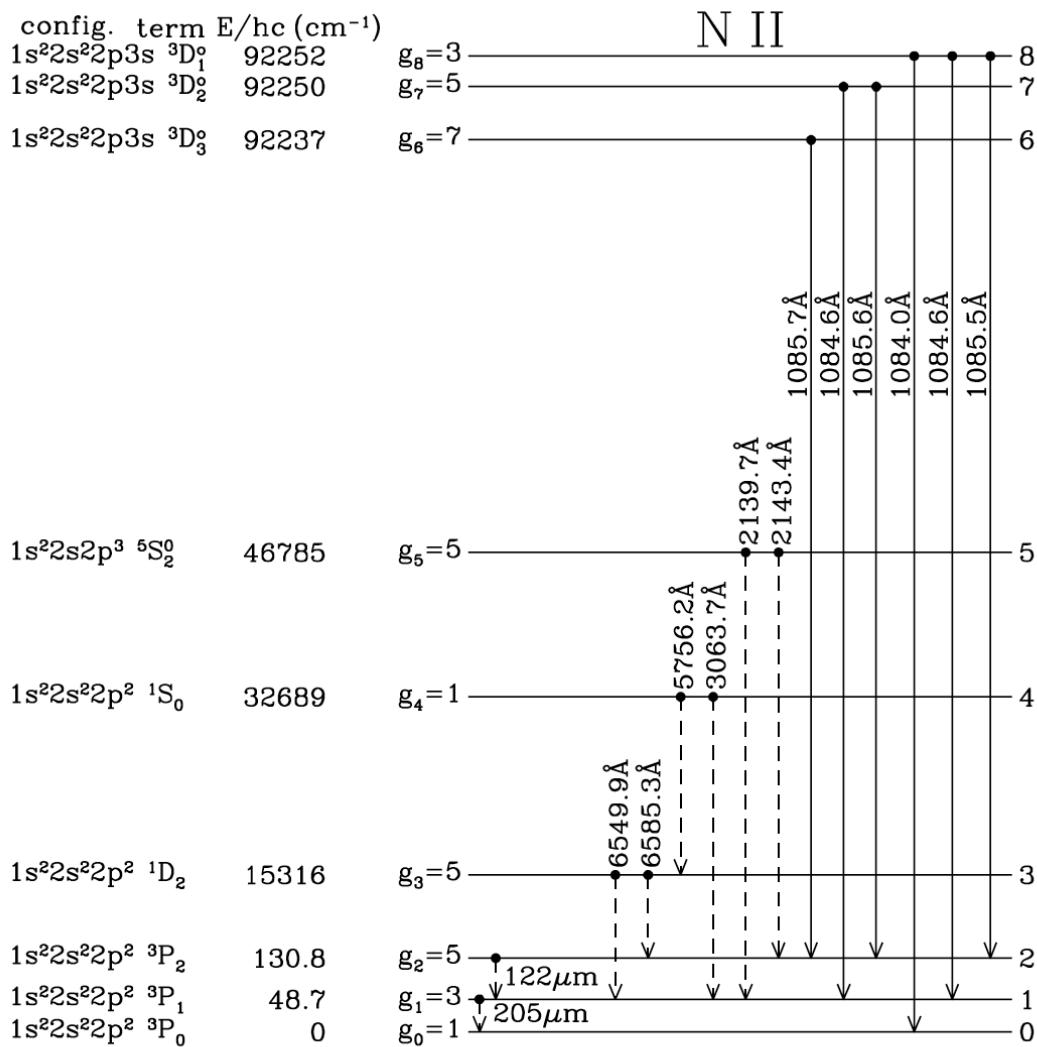


Figure 6.1 First nine energy levels of N II. Forbidden transitions are indicated by broken lines, and allowed transitions by solid lines; forbidden decays are not shown from levels that have permitted decay channels. Fine-structure splitting is not to scale. Hyperfine splitting is not shown.

Figure 12.4: NII_energy.png

Note that the number of electrons in each orbital does not change between these states, but the spin orientation and angular momentum values have changed.

Example

Now, consider using the 4364 angstrom and 5008 angstrom [OIII] lines as a temperature diagnostic. In **Draine 18.2**, we are given an example in which

$$\frac{I([OIII]4364)}{I([OIII]5008)} = 0.003$$

Using Figure 18.2 from Draine:

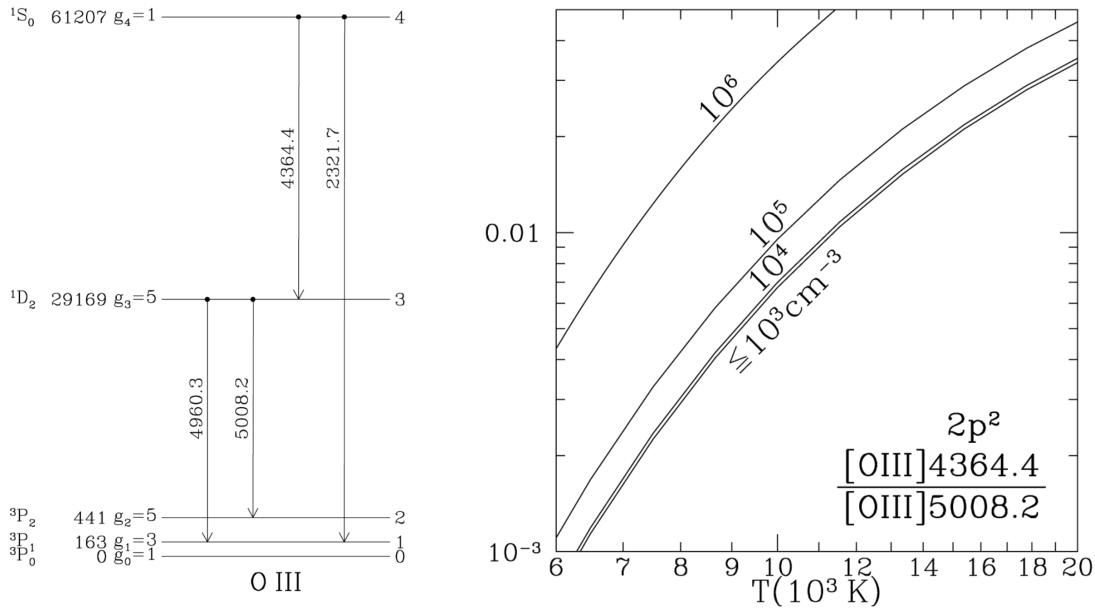


Figure 12.5: OIII_temperature.png

We can see that for densities $n_e \lesssim 10^4 \text{ cm}^{-3}$ this indicates that the temperature of the nebula is $\sim 8000 \text{ K}$.

12.4.2 Density-Sensitive Line Ratios

Other types of atomic lines are instead sensitive to the gas density.

Here, the key is to find lines for which the ground state is a singlet and the first excited state is a doublet. The ratios of the intensities of the transitions from each of the doublet lines down to the ground state is sensitive to the density (proportional to the collision rate) for $n < n_{crit}$. This happens because the level separation in the doublet is typically $\ll kT$, and so the thermal energy of the colliding particles will be nearly the same for both lines. Thus, instead of being sensitive to the gas temperature, the relative energy level populations will be largely set by the **number** of the collisions as determined by the density of collision partners.

Example

Now, consider using the 3729.8 angstrom and 3727.1 angstrom [OII] lines as a density diagnostic. In **Draine 18.2**, we are given an example in which

$$\frac{I([OII]3729.8)}{I([OII]3727.1)} = 1.2$$

Using Figure 18.4 from Draine:

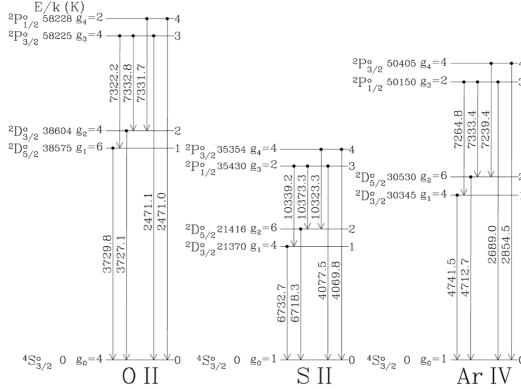


Figure 18.3 First five energy levels of the $2p^3$ ion O II, and the $3p^3$ ions S II and Ar IV. Transitions are labeled by wavelength in vacuo.

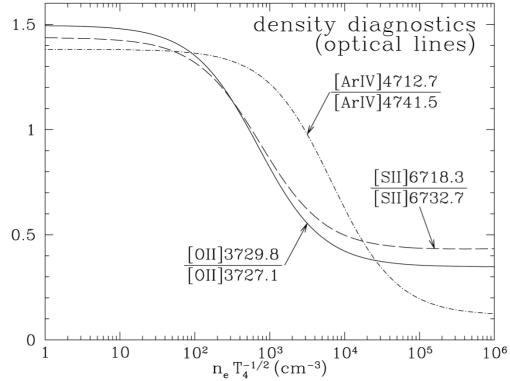


Figure 18.4 [O II], [S II], and [Ar IV] optical line intensity ratios useful for density determination. Wavelengths are in vacuo.

Figure 12.6: OII_density.png

We can see that for a temperature of $T_e \sim 10^4$ K, the electron density is $\sim 10^2 - 10^3$ cm⁻³

12.4.3 BPT Diagram

One of the most commonly used application for observations of ionized gas line ratios (at least in an extragalactic context) is to distinguish between sources with hard (X-ray dominated) and soft (UV-dominated) spectra. The typical photon energies of such hard spectra are large enough ($kT >> 10^4$ K) that they can be assumed not to have a thermal origin. Thus, one can devise a criterion in which one can separate sources with thermal spectra (galaxies powered primarily by star formation) and sources with nonthermal spectra (galaxies with substantial X-ray fluxes from a central active black hole).

The classic form of this diagnostic is known as the Baldwin-Phillips-Terlevich or BPT diagram.

Here, one plots the ratio of the [NII] 6585 angstrom line with H α against the ratio of the [OIII] 5008 angstrom line with H β . Both [NII] and [OIII] are sensitive to the degree of ionization in the gas (and so the hardness of the radiation field). The hydrogen recombination lines provide an important normalization for the total amount of thermal (low-ionization) flux in the galaxy, so that one can clearly distinguish an excess nonthermal (high-ionization) contribution, if it is present.

These specific line ratios are primarily chosen because these are all strong and so easily-observed lines. Additionally, the very similar wavelengths of the two pairs of lines ensure that they should have nearly the same extinction due to dust.

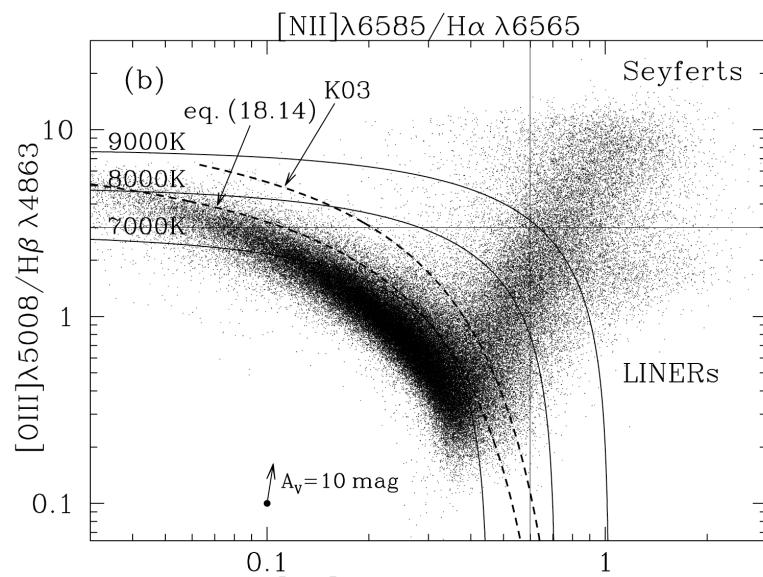


Figure 12.7: BPT.png

Chapter 13

Fluid Dynamics

13.1 The Basic Assumptions

Here, we revisit a few underlying concepts from the glossary in Chapter 1.

The Ideal Gas Equation of State applied to the ISM:

- The gas consists of a large number of molecules that are in random motion and obey Newton's laws of motion. So far, so good!
- The particles are far enough apart (so that the volume of the individual particles is much smaller than the volume of the gas). A quick check of this assumption: For a density of 10^8 hydrogen atoms per cubic meter, assuming an effective radius for a hydrogen atom to be 10^{-10} m, the total volume of the atoms would be $10^8(4/3)\pi(10^{-10})^3 = 4 \times 10^{-22}$ m³
- The particles are not TOO far apart (so that they undergo many collisions before crossing a region, and the velocity distribution can be described by kinetic temperature). Checking this: we saw MFP was large (105 km) but this isn't actually so large compared to typical scales (1 pc $\sim 10^{13}$ km).
- All collisions are perfectly elastic. Is this always true? No— as atoms and molecules have internal energy states, some of the collision energy can be absorbed (and chemical reactions can occur). However, it turns out that, energetically, these processes are negligible. This is not because the typical changes in internal energy are small compared the bulk kinetic energy of the particles— in fact, they can be quite comparable (Take a hydrogen atom at a temperature of 10000 K. $U = 3/2kT = 2 \times 10^{-19}$ J. The energy of the H alpha transition at 656 nm is $E = hc/\lambda = 3 \times 10^{-19}$ J). Instead, it is because a typical collision is unlikely to have exactly the correct energy to result in a quantum transition: most collisions do not excite internal energy states (or result in a chemical reaction) and so can be considered perfectly elastic.
- Apart from brief elastic collisions, there are no intramolecular forces. This is also not true! As we have discussed, many of the particles are either charged or even if neutral, have a dipole moment. However, the extreme low density combined moderate temperatures keeps particles far enough apart and interactions sufficiently brief that the intermolecular forces aren't too important.

We can then describe ISM gas with the ideal gas equation of state:

$$P = nk_B T \text{ or } P = \frac{\rho}{\mu} k_B T$$

The Mean Molecular Weight μ is a dimensionless quantity that can be expressed as

$$\mu = \frac{\bar{m}}{m_H}$$

\bar{m} can be either the mass of a single particle, or the average mass of an ensemble of particles. If we know the concentration of different species by number, we can write:

$$\bar{m} = \%A(m_A) + \%B(m_B) + \%C(m_C) + \dots + \%N(m_N) / 100\%$$

Maxwell-Boltzmann Statistics and T_{kin}

A typical particle in the ISM will undergo many collisions before traversing a ‘significant’ distance. Thus, we can say that the gas will essentially always be well-described by Maxwell-Boltzmann statistics. This also assumes:

- (1) Individual gas particles are distinguishable (Unlike for Bose-Einstein statistics, in which individual particles are not distinguishable)
- (2) There is no limit on the number of particles in a system that can share the same internal energy state (Unlike for Fermi statistics, in which each energy state of the system can be occupied by at most one particle).

From the kinetic theory of gases, assuming that the motion of the particles is homogeneous in 3 dimensions, we can then write a probability function that describes the likelihood of finding a particle in our system having a velocity v :

$$f(v) dv = \left(\frac{\bar{m}}{2\pi k T} \right)^{3/2} 4\pi v^2 \exp\left(-\frac{\bar{m}v^2}{2kT}\right) \text{ where } k \text{ is the Boltzmann constant.}$$

By definition, the temperature in this equation is the gas kinetic temperature T_{kin} .

The mean value of this distribution describes the typical random speed of a particle. This is often called the thermal velocity, and it can be expressed as

$$v_{th} = \sqrt{\frac{8k_B T}{\pi \bar{m}}} \text{ or } \sqrt{\frac{8k_B T}{\pi \mu m_H}}$$

Beyond being able to apply Maxwell-Boltzmann Statistics (and the ideal gas law), another important consequence of the fact that mean free paths are small in the ISM is that when streams of gas collide, diffusion is negligible, and a distinct boundary between the streams remains.

13.2 Conservation Laws

To derive the equations of astrophysical fluid mechanics, we will focus on the conservation of three key quantities: mass, momentum, and energy. For a full magnetohydrodynamic (MHD) approach, we would also want to consider the conservation of magnetic flux, but we will neglect this in a basic treatment. In general, applying conservation of energy is much more complicated (hello, thermodynamics) and we will just consider two specific and simplifying cases.

We will adopt several other important simplifications, beginning with that of a simple, plane-parallel flow, where we only consider motion in an (arbitrary) x -direction.

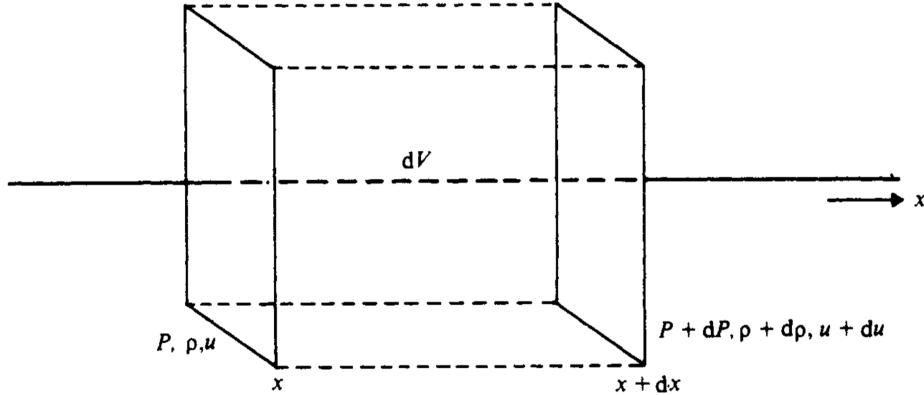


Figure 13.1: flow.png

We define the variables of interest as the initial and final pressure P , density ρ , and bulk (flow) velocity (u in this figure but generally defined as v) at the entry and exit boundary along the x direction for our volume element dV .

13.2.1 Conservation of Mass

The conserved quantity, the mass, is defined as

$$M \equiv \rho dV.$$

As we are considering only motion in the x direction, we can also consider ρdx to be a conserved quantity.

Since this must remain constant, we can write the equation of conservation as:

$$\frac{D}{Dt} (\rho dx) = \frac{\delta \rho}{\delta t} + v \frac{\delta \rho}{\delta x} + \rho \frac{\delta v}{\delta x} = 0$$

This is typically referred to as the continuity equation. These derivatives are all partial derivatives, so time derivatives are evaluated at a fixed point in space, and space derivatives are evaluated at a fixed instant in time.

Here we assume there are no mass sinks or sources. Note that this is not always a valid assumption in the ISM, as mass can be lost by accretion onto sources that do not behave as fluids (individual protostars, stars, or black holes) or added to the fluid via stellar winds or explosions. Actual hydrodynamic simulations that (for example) follow the evolution of a gas cloud to form stars, will explicitly include sink particles to deal with this.

For a time-independent flow,

$$v \frac{d\rho}{dx} + \rho \frac{dv}{dx} = 0$$

and we can integrate to find $\rho v = \text{constant}$

13.2.2 Conservation of Momentum

The conserved quantity, the momentum, is defined as

$$p \equiv \rho v dV.$$

Starting from $F = ma$, taking $dV = dx dA$, and considering just the force due to pressure (for now we ignore any forces due to gravity, electromagnetism, or viscosity) we can then write

$$(\rho dx) \frac{D}{Dt}(v) = -\frac{dF}{dA} = -P$$

We can rewrite this as:

$$\frac{\delta v}{\delta t} + v \frac{\delta v}{\delta x} = -\frac{1}{\rho} \frac{\delta P}{\delta x}$$

This is sometimes referred to as Euler's equation.

For a time-independent flow,

$$v \frac{dv}{dx} = -\frac{1}{\rho} \frac{dP}{dx}$$

and we can integrate to find $P + \rho v^2 = \text{constant}$

13.2.3 Conservation of Energy

An equation for the conservation of energy can be quite complicated, so here we will consider two limiting cases for which the energy equation is a relatively simple expression of several thermodynamic variables.

13.2.4 Adiabatic Flow

One case that has many important astrophysical applications is that in which a parcel of gas can be assumed not to exchange heat with its environment—this is an adiabatic flow. Any energy change is due solely to work (either the element expanding and doing work on its surroundings, or vice versa). In this scenario, the entropy of the gas remains constant along the flow and the energy equation takes the form:

$$P \propto \rho^\gamma$$

Here γ is the adiabatic index, which for a monoatomic ideal gas is just 5/3.

13.2.5 Isothermal Flow

Alternatively, if the gas efficiently exchanges heat with its surroundings (for example, due to absorbing and emitting radiation) you will have an isothermal flow. Here, efficiently generally means that the heating and cooling timescales are much less than the dynamical time scale.

The gas temperature is set by the balance between the heating and cooling rate ($\Gamma = \Lambda$) and the energy equation can be reduced to the extremely simple form

$$T = \text{constant}$$

13.3 Sound Waves and Propagation

Generally, for both adiabatic and isothermal conditions we can relate the pressure and density by:

$$P \propto \rho^n$$

For an adiabatic gas, $n = \gamma = 5/3$, and for an isothermal gas $n = 1$ (which, if we substitute into the ideal gas equation, is equivalent to saying $T = \text{constant}$).

We are now interested in the time scale on which pressure changes occur in the gas (recall that we already know how to find the time scale on which the temperature changes: this is governed by the heating or cooling rate). Essentially, how fast does a pressure wave move through the gas? This ought to sound familiar: what we are looking for here is the speed of sound!

We can do a lazy professor's derivation of the speed of sound as follows: Assume that we have a gas at rest, with a constant pressure P_0 and a constant density ρ_0 . We then make a small change (a perturbation) in the density and pressure so that:

$$P = P_0 + dP$$

$$\rho = \rho_0 + d\rho$$

We are going to do a linear analysis, which means that we assume dP and $d\rho$ are so small that we can ignore any terms of dP^2 , $d\rho^2$, or higher order as being essentially zero.

We then define these perturbations using the above relationship between P and ρ and taking the derivative of the pressure with respect to density, which gives us the following relationship between the differential quantities dP and $d\rho$:

$$dP = \gamma \rho^{\gamma-1} d\rho$$

Substituting $\rho_0 + d\rho$ for ρ :

$$dP = \gamma(\rho_0 + d\rho)^{\gamma-1} d\rho$$

Now, we need to expand the term $(\rho_0 + d\rho)^{\gamma-1}$ using the binomial theorem:

$$(\rho_0 + d\rho)^{\gamma-1} = \rho_0^{\gamma-1} + O(d\rho) + O(d\rho^2)$$

Where $O(d\rho^n)$ indicates a term proportional to $d\rho$ raised to the power n .

Luckily, since we are multiplying by $d\rho$ and we get to ignore terms of $O(d\rho^2)$ we only keep the first term of this expansion, yielding:

$$dP = \gamma \rho_0^{\gamma-1} d\rho$$

We then do a little arithmetic trick and say that $x^{n-1} = x^n / x$:

$$dP = \gamma \left(\frac{\rho_0^\gamma}{\rho_0} \right) d\rho$$

But ρ_0^γ is just P_0 so:

$$dP = \gamma \left(\frac{P_0}{\rho_0} \right) d\rho \quad \text{or} \quad \frac{dP}{d\rho} = \gamma \left(\frac{P_0}{\rho_0} \right) = c_s^2$$

The quantity c_s is the speed of sound in this medium, or the speed at which changes in pressure propagate through a gas:

$$c_s = \left(\frac{dP}{d\rho} \right)^{1/2} = \left(\frac{\gamma P}{\rho} \right)^{1/2}$$

(Again, this is really half a derivation of the speed at which pressure disturbances propagate through this medium— for a full derivation you would plug these perturbations into a set of equations for fluid flow, which would yield a differential equation of a form consistent with wave propagation, from which you would see this term defined as the wave speed.)

Recall that for an isothermal gas, $\gamma = 1$ so we can define a special case, the isothermal sound speed, as

$$c_i = \sqrt{\frac{p}{\rho}}$$

Then the adiabatic sound speed in an ideal non-relativistic gas is just 1.3 times the isothermal speed. We can also rewrite this expression for c_i using the ideal gas equation and our definition of number density:

$$c_i = \sqrt{\frac{k_B T}{\mu m_H}}$$

For an isothermal gas the sound speed is then the same as the root-mean-square of the 1D thermal velocity (basically the thermal velocity that we previously defined above!).

Using the sound speed, we can define two useful quantities. One is the Mach number:

$$M = \frac{v}{c_s}$$

This is a unitless quantity that describes how fast the gas in a flow is moving as a ratio with the speed of sound in that gas.

The other is the sound crossing time for a source of length L :

$$t_{SC} = \frac{L}{c_s}$$

This is the time it takes a sound wave to traverse a region. This is the typical or characteristic timescale on which changes in pressure can occur in a system (or on which a system can respond to some outside stimulus by adjusting the pressure).

To get an idea of whether a system is adiabatic or isothermal, you can compare the cooling time to the sound crossing time.

For an adiabatic process, $t_{SC} < t_{cool}$. This means that either changes to the system are happening fast, or that it just takes a long time to cool.

For an isothermal process $t_{SC} > t_{cool}$. Now the time scale for pressure adjustment is much longer than the time it takes to cool.

As we previously saw, isothermal conditions are a good assumption for gas in an HII region, and indeed for much of the ISM where opacity is low and heat exchange via radiation is efficient. In contrast, adiabatic conditions tend to become more important when cooling is inefficient, particularly when opacity increases to the point that radiation cannot easily escape a source (the insides of stars is a great example of this!).

13.4 Shocks

13.4.1 Shock Formation

What is a shock? Like a piston in a car motor, this is a rapid (and so adiabatic) change in the pressure: specifically, a compression wave.

Imagine that our gas is in a tube, and we push the piston to the right.

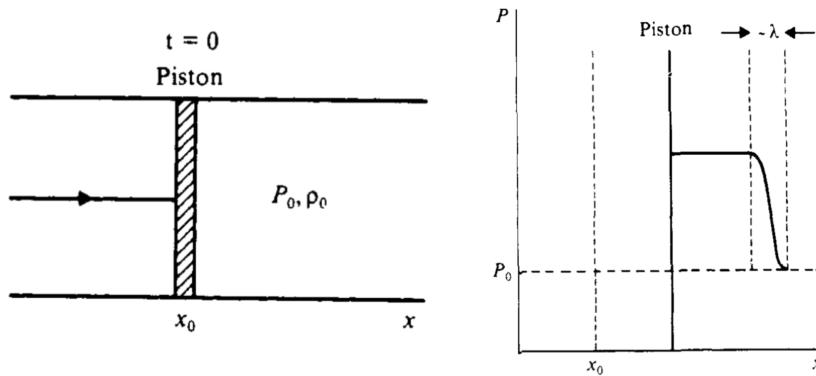


Figure 13.2: *shock_structure.png*

Let's imagine that we do this incrementally. Our first push results in an increase of the initial pressure P_0 . This pressure wave propagates to the right at the (adiabatic) speed of sound c_s , and results in the compression (increased density) of the gas it moves through. We then push the piston again, and it again leads to a wave of higher pressure, but the gas it is pushing into has already been compressed, and is at higher pressure and density. How does this change the speed of sound, since it is related to the ratio of the pressure and density? Well, we can use the adiabatic equation of state ($P \propto \rho^\gamma$) and the relation we used to 'derive' the speed of sound ($dP = \gamma \rho^{\gamma-1} d\rho$) to show that while both the pressure and the density increase, the pressure increases faster:

$$c_s^2 \propto \frac{dP}{d\rho} \propto \gamma \rho^{\gamma-1}$$

For an adiabatic index of $=5/3$, we can say

$$c_s^2 \propto \rho^{2/3} \quad \text{or} \quad c_s \propto \rho^{1/3}$$

This means that the speed of sound in this gas is higher than it was, and so this pressure wave moves faster than the one in front of it and starts catching up with it, increasing the magnitude of the pressure change. This keeps happening, and results in a pile-up effect as each of these subsequent pressure waves tries to overtake the first. All of these incremental pressure changes then essentially combine into one drastic pressure change: a shock wave.

13.4.2 Examples of shocks in the ISM:

- Collisions of parcels of turbulent gas - Heats gas to temperatures of up to a few hundred K. Increases molecular complexity (adds grain-chemistry species like SiO, CH₃OH, and complicated molecules to the gas phase).

- Protostellar jets - Drive shocks in to the surrounding molecular clouds. Causes localized enhancements of SiO, and excites maser transitions where the jets collide with the surrounding gas.
- HII region expansion - Drives shocks into the surrounding molecular clouds. Helps dissipate clouds.
- Stellar winds (and colliding-wind binaries) - Relatively localized shocks, as the natal cloud has dissipated by this time. Emit X-rays.
- Supernova blast waves - drives shocks into the diffuse (ionized and neutral) ISM. Heats the this gas to a super-hot plasma.
- Spiral arm shocks - large scale shocks in galaxies, responsible for converting gas from atomic to molecular
- Bar shocks - collisions of gas on self-intersecting orbit removes angular momentum, increases molecular fraction, and drives gas toward the inner regions of a galaxy

13.4.3 Shock Structure

The discontinuity between the initial (pre-shock) and final (post-shock) pressure has a typical size-scale over which the pressure jump occurs that is characteristic of the mean free path in the gas. In this area, typical relations used to describe the gas flow break down, and the microscopic processes we have talked about in this class (collisions which increase the viscosity, energy loss through ionization and even heat transfer through conduction) act to keep the shock from becoming infinitely steep (note that while these processes transform the type of energy in the shock, energy overall is still conserved across the shock).

We then want to describe the effects of a shock in terms of the conditions on either side of the shock front: the initial (upstream) and final (downstream) conditions:

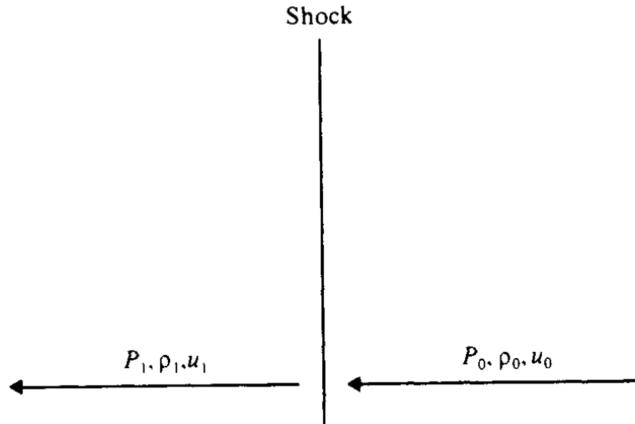


Figure 13.3: *shock_variables.png*

As before, we encounter the variable u , which is commonly used to refer to the x-component of the velocity, v_x .

Noting that we still conserve mass, energy, and momentum across the shock, we can find expressions (the so-called “jump conditions”) relating the initial and final pressure P , density ρ , and the velocity in the x-direction which for simplicity we will just call v .

13.4.4 Jump Conditions

We again adopt the same (1D flow) geometry as in Section 13.2, and further consider only a steady-state flow.

As we saw, the first two equations (The continuity equation and the Euler equation) gave us the following conserved properties for steady state conditions:

$$\phi = \rho v = \text{constant} \quad (\text{this is the } \mathbf{\text{mass flux}})$$

$$\zeta = P + \rho v^2 = \text{constant} \quad (\text{this is the } \mathbf{\text{momentum flux}})$$

We need to add one more requirement based on the energy:

$$E = \rho v \left(\frac{1}{2} v^2 + \frac{3}{2} \frac{P}{\rho} \right)$$

where the first term in the parentheses represents the kinetic energy per unit mass, and the second term represents the internal energy per unit mass, and we have assumed $\gamma = 5/3$.

Across the shock, the energy will change due to the work done on the gas by the pressure difference:

$$\Delta E = -\Delta Pv.$$

Combining these two expressions, we can write an expression for the quantity which will be conserved across the shock:

$$\xi = \frac{1}{2} v^2 + \frac{5}{2} \frac{P}{\rho} = \text{constant} \quad (\text{this is the } \mathbf{\text{specific total energy}})$$

These three quantities (mass flux, momentum flux, and specific total energy) will be conserved across the shock:

$$\rho_0 v_0 = \rho_1 v_1$$

$$P_0 + \rho_0 v_0^2 = P_1 + \rho_1 v_1^2$$

$$\frac{1}{2} v_0^2 + \frac{5}{2} \frac{P_0}{\rho_0} = \frac{1}{2} v_1^2 + \frac{5}{2} \frac{P_1}{\rho_1}$$

Together, these three equations are called the Rankine-Hugoniot relations. We will use them to determine pre- and post-shock values of the quantities which are discontinuous (or "jump" in value) across the shock: P , ρ , v , and T .

13.4.5 Strong Shocks

Deriving the relationships between the pre- and post-shock values of P , ρ , and v involves a lot of algebraic gymnastics.

To begin with, we will note that the ratio between the momentum flux and the mass flux has units of velocity, so we will use this to define a reference velocity

$$u \equiv \frac{\zeta}{\phi}$$

Additionally, for $\gamma = 5/3$,

$$c_s^2 = \frac{5}{3} \frac{P}{\rho}$$

We can use the quantity u to rewrite the Mach number:

$$M^2 = \frac{v^2}{(5/3)(P/\rho)} = \frac{3}{5} \frac{v}{u - v}$$

From the Rankine-Hugoniot conditions, we can then write the specific energy as:

$$\xi = \frac{1}{2}v^2 + \frac{3}{2}c_s^2 = v \left(\frac{5}{2}u - 2v \right)$$

And rewrite the momentum equation as:

$$v^2 + uv + \frac{3}{5}c_s^2 = 0$$

The two roots of this quadratic equation will then give us the pre- and post-shock flow velocities.

Together, this yields the relation:

$$v_0 + v_1 = \frac{5}{4}u$$

We can eliminate u by using our definition of the Mach number to finally write:

$$\frac{v_1}{v_0} = \frac{M_0^2 + 3}{4M_0^2}$$

Everything becomes wonderfully more simple if we assume we are dealing with a strong shock, in which $M_0 \gg 1$ (Here, the upstream or pre-shock gas is strongly supersonic, while the downstream or post-shock gas will be subsonic). In this case,

$$\frac{v_1}{v_0} = \frac{1}{4}$$

Then, from the continuity equation, we also know that

$$\frac{\rho_1}{\rho_0} = 4$$

and we can further derive:

$$P_1 = \frac{3}{4}\rho_0 v_0^2$$

and from the ideal gas equation,

$$T_1 = \frac{3}{16} \frac{\mu m_H}{k_B} v_0^2$$

13.4.6 Fixed Reference Frame

The above results were derived in a reference frame in which the shock is stationary ($v_s = 0$). It is often more convenient to rewrite these results for a fixed reference frame, in which the shock propagates at a speed v_s .

Then,

$$v_0 = v_l - v_s$$

and

$$v_1 = v_f - v_s$$

(where v_i and v_f are the upstream and downstream velocities measured in the fixed frame)

The speed of the gas after the shock has passed through is then:

$$v_f = \frac{3}{4}v_s$$

The post-shock pressure in terms of the shock speed is:

$$P_f = \frac{3}{4}\rho_i v_s^2$$

The post-shock temperature in terms of the shock speed is:

$$T_f = \frac{3}{16} \frac{\mu m_H}{k_B} v_s^2$$

Note that shocks with a speed of at least 100 km/s will generally fully ionize the gas!

Finally, we can also write the specific internal energy of the shocked gas as

$$\xi_f = \frac{9}{32} \rho_i v_s^2 = \frac{3k_B T_f}{\mu m_H}$$

13.4.7 Cooling Time and Cooling Length

After the shock has passed through, the post-shock gas can cool radiatively. This introduces another important scale: the cooling time, or the length of time it takes for the post-shock gas to roughly cool to its pre-shock value, radiating the bulk of the thermal energy it gained from the shock's passage.

We can define the cooling rate per unit volume in the post-shock gas as:

$$L_V = n_f^2 \Lambda(T_f) \quad \text{with units of energy per time per unit volume}$$

The cooling rate per unit mass is:

$$L_M = \frac{n_f \Lambda(T_f)}{\mu_f m_H} \quad \text{with units of energy per time per unit mass}$$

(Note that μ_f may not be the same as μ_i if the post-shock gas is ionized!)

The cooling time is then the time to radiate away the specific internal energy (energy per unit mass):

$$t_c = \frac{\mu_f m_H \xi_f}{n_f \Lambda(T_f)} = \frac{3k_B T_f}{n_f \Lambda(T_f)}$$

Another useful quantity is the cooling length, or how far away from the shock front the gas will achieve this temperature:

$$l_c = \frac{v_s t_c}{4} = \frac{3k_B T_f v_s}{4n_f \Lambda(T_f)}$$

13.4.8 Shocks and the 3-Phase ISM

A typical supernova might have initial shock speeds up to (and above) 10,000 km/s. Assuming the post-shock gas is ionized hydrogen, this gives:

$$T_f = \frac{3}{16} \frac{0.5(1.67 \times 10^{-27} \text{ kg})}{1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}} (10^7 \text{ m s}^{-1})^2 = 10^9 \text{ K}$$

Such high temperatures are enough to permit nucleosynthesis of heavy elements in the supernova remnant.

Even with the lower speeds of these shocks far away from the supernova, as they pass through diffuse atomic gas, they are sufficient to raise the temperature of that gas above 10^4 K – the rough limit to which gas can be radiatively heated.

The ‘coronal gas’ phase of the ISM has temperature of 10^5 K, and as such is likely heated primarily by supernova-driven shocks:

Region	Particle density (m^{-3})	Temperature (K)	Pressure (N m^{-2})
Diffuse nebula	$\gtrsim 10^8$	8000	$\gtrsim 1.1 \times 10^{-11}$
Diffuse atomic cloud	3×10^7	70	2.9×10^{-14}
Intercloud	3×10^5	6000	2.5×10^{-14}
Cool molecular cloud	10^9 – 10^{10}	20	2.8×10^{-13}
			2.8×10^{-12}
Coronal gas	$< 10^4$	5×10^5	$< 6.9 \times 10^{-14}$

Figure 13.4: *ISM_table.png*

We can compare the sound-crossing time of the coronal gas (which typically occupies a galaxy’s halo) to the intercloud gas (the warm ionized medium) and the diffuse HI clouds (the warm neutral medium), which have typical sizes of 5 pc.

For a typical HI cloud,

$$c_s = \sqrt{\frac{5 k_B T}{3 \mu m_H}} = \sqrt{\frac{5 (1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1})(70 \text{ K})}{0.5(1.67 \times 10^{-27})}} = 1.4 \text{ km/s}$$

The sound-crossing time is then

$$t_s = \frac{5(3.086 \times 10^{16})}{1.4 \times 10^3} = 3 \times 10^6 \text{ years}$$

This is plenty of time for a sound wave to move a proportionately larger distance through the intercloud gas:

$$c_s = \sqrt{\frac{5 (1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1})(6000 \text{ K})}{0.5(1.67 \times 10^{-27})}} = 12.9 \text{ km/s}$$

$$d = (1.29 \times 10^4 \text{ m s}^{-1})(9.4 \times 10^{13} \text{ s}) = 40 \text{ pc}$$

and the coronal gas:

$$c_s = \sqrt{\frac{5}{3} \frac{(1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1})(5 \times 10^5 \text{ K})}{0.5(1.67 \times 10^{-27})}} = 117 \text{ km/s}$$

$$d = (1.17 \times 10^5 \text{ m s}^{-1})(9.4 \times 10^{13} \text{ s}) = 350 \text{ pc}$$

which supports the observation that these three phases are in equilibrium.