

Physics 610: Interstellar Matter

John F. Wu[†]

[†]jfwu@physics.rutgers.edu
Rutgers University, Department of Physics
and Astronomy

These are my notes on the Rutgers graduate level course on interstellar matter (ISM; although this abbreviation is usually reserved for interstellar *medium*). This course was taught in Fall 2015 by Eric Gawiser, and based on lecture notes by Andrew J. Baker. The primary textbook is by Bruce T. Draine - *Physics of the Interstellar and Intergalactic Medium* (2011). Also relevant in the latter half of the course is *The Physics of Astrophysics Volume II: Gas Dynamics* by Frank H. Shu (1992).

I strive for, but do not guarantee, scientific (or grammatical accuracy) in these personal notes. If you find any errors, please let me know via email.

Friday, 4 September

Introduction

The first chapter of Draine, and a review of Rybicki & Lightman - *Radiative Processes in Astrophysics* (1979) suffices for an introduction.

A few definitions

A **specific intensity**, I_ν , can be defined via the differential energy passing through a differential area and viewed through a differential solid angle:

$$dE = I_\nu (\hat{n} \cdot \hat{r}) dv dA d\Omega dt. \quad (1)$$

Specific intensity is conserved along a ray (in the absence of everything).

Note that the usual units are $[I_\nu] = \text{erg cm}^{-2} \text{ster}^{-1} \text{Hz}^{-1} \text{s}^{-1} = 10^{23} \text{Jy}$.

Surface brightness, μ (or I , in Rybicki & Lightman), is simply the frequency-integrated specific intensity:

$$\mu = \int dv I_\nu, \quad (2)$$

in is likewise conserved, barring redshifts or matter.

If relative velocities are accounted for, then it turns out that I_ν / ν^3 is Lorentz invariant. In terms of cosmological redshifts, where observed and emitted frequencies (ν_o, ν_e) are related via $\nu_o = \nu_e / (1 + z)$. Thus,

$$\frac{I_{\nu_o}}{I_{\nu_e}} = \left(\frac{\nu_o}{\nu_e} \right)^3 = (1 + z)^{-3} \text{ and}$$
$$\frac{\mu_o}{\mu_e} = \left(\frac{\nu_o}{\nu_e} \right)^4 = (1 + z)^{-4}.$$

This latter factor of $(1 + z)^{-4}$ is the “cosmological surface brightness dimming.”

The **specific energy density**, u_ν , is the energy per volume per frequency:

$$u_\nu = \frac{1}{c} \int d\Omega I_\nu. \quad (3)$$

The **flux density**, F_ν , is the directional solid-angle integrated intensity:

$$F_\nu = \int d\Omega I_\nu \hat{n} \cdot \hat{r}. \quad (4)$$

Keep in mind that our telescope detectors actually measure specific intensity, but by assuming isotropic emission, intensity can be converted into a flux density.

The **flux**, F , is the flux density integrated over the entire spectrum, and can be expressed by :

$$\begin{aligned} F &= \int d\Omega \mu \hat{n} \cdot \hat{r} \\ &= \int d\nu F_\nu \\ &= \int d\lambda F_\lambda. \end{aligned} \quad (5)$$

The **luminosity**, L , is simply

$$L = \int dA F, \quad (6)$$

and has units of erg s^{-1} .

Another useful quantity is the **momentum flux**, p_ν , which includes yet another factor of $\hat{r} \cdot \hat{n}$, which I'll now define as $\cos \theta \equiv \hat{r} \cdot \hat{n}$. So then, the specific energy density, the flux density, and the momentum flux are the zeroth, first, and second **angular moments of specific intensity**:

$$\begin{aligned} u_\nu &= \frac{1}{c} \int d\Omega I_\nu \\ F_\nu &= \int d\Omega I_\nu \cos \theta \\ p_\nu &= \frac{1}{c} \int d\Omega I_\nu \cos^2 \theta. \end{aligned} \quad (7)$$

If we assume isotropy, then we can define a **mean intensity**, $J_\nu = (1/4\pi) \int_{4\pi} d\Omega I_\nu$. Under these assumptions, $I_\nu \Rightarrow J_\nu$, and then:

$$\begin{aligned} u_\nu &= \frac{4\pi}{c} J_\nu \\ F_\nu &= 0 \\ p_\nu &= \frac{4\pi}{3c} J_\nu. \end{aligned}$$

The last quantity should be obvious, since the factor of three stems out of our living in a three-dimensional world.

Radiative transfer

What does matter do to specific intensity along a ray?

Matter can spontaneously emit or emit photons with stimulation, as well as absorb photons.

The **monochromatic emission coefficient**, j_ν , is the energy per time per solid angle per frequency per volume: $dE = j_\nu dV d\Omega dt d\nu$. Therefore along some ray, ds , the spontaneous emission is $dI_\nu = j_\nu ds$.

The **absorption coefficient**, α_ν (or κ_ν in Draine), is the loss of specific intensity along some ray: $dI_\nu = -\alpha_\nu I_\nu ds$.

Note that stimulated emission is essentially a *negative* absorption coefficient—and so α_ν (κ_ν) includes that term. Thus, the **equation of radiative transfer** is

$$\frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu. \quad (8)$$

The solutions to this differential equation require an **optical depth** definition:

$$\tau_\nu = \int ds \alpha_\nu(s), \quad (9)$$

where s is a chosen path.

Also useful is to define a **source function** term, which is just the ratio of the emission to absorption coefficients:

$$S_\nu = \frac{j_\nu}{\alpha_\nu}. \quad (10)$$

This allows us to continue with equation (8) and rewrite it more generally:

$$\begin{aligned} \frac{dI_\nu}{d\tau_\nu} &= -I_\nu + S_\nu \\ e^{\tau_\nu} \frac{dI_\nu}{d\tau_\nu} + I_\nu e^{\tau_\nu} &= S_\nu e^{\tau_\nu} \\ \frac{d}{d\tau_\nu} (e^{\tau_\nu} I_\nu) &= S_\nu e^{\tau_\nu}. \end{aligned} \quad (11)$$

The solution is then

$$I_\nu = I_{\nu,0} e^{-\tau_\nu} + \int_0^{\tau_\nu} d\tau'_\nu S_\nu(\tau'_\nu) e^{\tau'_\nu - \tau_\nu}. \quad (12)$$

Finally, thermal emission requires the use the **blackbody** or **Planck function**:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \quad (13)$$

Wednesday, 9 September

Emission and absorption in a thermal plasma

In a thermal plasma, we can expect the following emission mechanisms:

- bound-bound (spontaneous & stimulated emission; note that there exist two-photon emission mechanisms)
- free-free (*bremssstrahlung*; inelastic “scattering” of an e^- off an ion)
- free-bound (recombination)

and likewise, all the similar absorption mechanisms. All of these are on Draine p. 92.

The emissivity of free-free radiation can be described by Draine equation (10.1):

$$j_\nu = \frac{C_4}{4\pi} n_e^2 T^{-1/2} \ln \left(C_3 T^{3/2} / \nu \right) \exp \left(-\frac{h\nu}{kT} \right). \quad (14)$$

The logarithmic term, $\ln C_3 T^{3/2} / \nu$, is called the **Gaunt factor**.

In a thermally emitting medium, $\alpha_\nu = j_\nu / B_\nu$. Empirically, $\alpha_\nu \propto n_e^2 T^{-1.35} \nu^{-2.1}$, and a proper form of it can be found in the review article by Condon (1992).

Regions of ionization hydrogen, or H II **regions**, are good places to find ionized gas and thermal plasma emission. We usually mean H II regions to mean the spheres nearby massive O, B stars ($M_\star \geq 5 M_\odot$; $T_{\text{eff}} \geq 15000$ K). The emission profile shows **nebular emission lines**, or some sort of “nebular continuum” due to all the free-free, recombination, and two-photon emissions. Draine figure 10.2 shows the continuous emission spectrum of a hydrogen plasma.

It is important to notice that stellar and galaxy spectral energy distributions (SEDs) show the *opposite*, in which case stellar photospheres *absorb* the continuum rather than emit it.

Synchrotron emission

Synchrotron radiation is the emission from relativistic (non-relativistic is called cyclotron radiation) particles accelerating in a magnetic field. Particles can be characterized by a **frequency of gyration**:

$$\omega_B = \frac{qB}{\gamma mc}. \quad (15)$$

The total power of synchrotron emission is

$$\frac{dE}{dt} = \left(\frac{2}{3} \right)^2 B^2 \gamma^2 \beta^2 c r_0^2, \quad (16)$$

where r_0 is the Bohr radius. This equation is the relativistic Larmor formula. Relativistic radiation will be beamed, thus narrowing the spatial extent of the power, and broadening the frequency profile of the spectrum. The cutoff frequency can be given by

$$\omega_c = \frac{3}{2} \gamma^3 \omega_B \sin \alpha_B, \quad (17)$$

where α_B is the angle between the magnetic field orientation and the direction of propagation.

If we (as usual) assume the electron energy distribution as a power law, e.g., $n(E)dE \propto E^{-p}dE$, then we can expect a power law synchrotron spectrum:

$$P_{\text{tot}} \propto \omega^{-(p-1)/2}. \quad (18)$$

We can use the synchrotron spectrum's shape to date the age of the emission source. Older synchrotron spectra will *steepen* the spectrum, whereas younger spectra (e.g., from recent supernova remnants, etc) will be flatter.

Einstein coefficients

If we assume that we have a two-level system separated by some energy $h\nu_{21}$, and where the ground state has energy E .

The spontaneous transition from energy $E + h\nu_{21}$ to E via a single photon can be described by the Einstein coefficient A_{21} , which has units of s^{-1} . A_{21} is the probability per unit time of spontaneous emission.

$B_{12}\bar{J}$ is the probability per unit time of excitation due to absorption of a photon with energy $h\nu_{21}$ from a radiation field with mean intensity $\bar{J} = \int_0^\infty J_\nu d\nu \phi(\nu)$, and $\phi(\nu)$ is the line profile (distribution) that integrates to unity. The units of B_{12} are therefore $\text{erg}^{-1} \text{cm}^2 \text{Hz ster}$.

Finally, $B_{21}\bar{J}$ is the probability per unit time of stimulated emission of a photon with energy $h\nu_{21}$ from a radiation field with mean intensity \bar{J} . Note that the emitted photon will force the emission of a photon with the same energy, direction, and polarization as the one in its field.¹

If we assume thermodynamic equilibrium, then we know that the rates of excitation and decay are equal. Let's define n_1, n_2 to be the number densities of atoms in these states. Then,

$$n_1 B_{12} \bar{J} = n_2 A_{21} + n_2 B_{21} \bar{J}. \quad (19)$$

Using basic thermodynamics (i.e., the Saha equation), we see that

$$\frac{n_1}{n_2} = \frac{g_1}{g_2} e^{h\nu_{21}/kT},$$

and

$$\bar{J} = \frac{A_{21}/B_{21}}{\frac{g_1 B_{12}}{g_2 B_{21}} e^{h\nu_{21}/kT} - 1}. \quad (20)$$

How do we match equation (20) to the Planck function? By matching, we can see that

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

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

Note that these are independent of temperature.

¹ One way to think of this is the fact that, just as fermions obey an exclusion principle, bosons find that mirroring another is energetically favorable; therefore the creation of a photon is enhanced in a radiation field (Shu, *Physics of Astrophysics Vol. I: Radiation*, 1991).

Using all of the above, we find that the source function is

$$S_\nu = \frac{2h\nu_{21}^3}{c^2} \left(\frac{g_2 n_1}{g_1 n_2} - 1 \right)^{-1}. \quad (21)$$

If $n_2 g_1 / n_1 g_2 = \exp(-h\nu_{21}/kT)$, we have thermal equilibrium. If $n_2 g_1 / n_1 g_2 \neq \exp(-h\nu_{21}/kT)$, then our emission is not in thermal equilibrium. In both cases, we require that $\exp(-h\nu_{21}/kT) < 1$. When $\exp(-h\nu_{21}/kT) > 1$, we enter the case of an inverted population where there is a *negative* excitation temperature (e.g., a maser).

Back to the Einstein coefficients: the B_{12} and B_{21} coefficients can be defined in terms of **oscillator strengths**, f_{12} and f_{21} , e.g.,

$$B_{12} = \frac{4\pi^2 e^2}{h\nu_{21} mc} f_{12}. \quad (22)$$

We can gain a little intuition of oscillator strengths by noting that $f_{12}, f_{21} \rightarrow 1$ in the classical regime. Thus, they represent quantum corrections.

Friday, 11 September

Line transitions

For a line emission or absorption, a line profile ought to be dependent on the intrinsic line shape (a Lorentz profile) and the Doppler broadening (a Maxwellian velocity distribution). The Lorentzian convolved with a gaussian is thus called a **Voigt** profile, or

$$\phi_\nu^{\text{Voigt}} = \frac{1}{\sqrt{2\pi}} \int dv \frac{1}{\sigma_v} e^{-v^2/2\sigma_v^2} \frac{4\gamma_{ul}}{16\pi^2 [\nu - (1 - v/c)\nu_{ul}]^2 + \gamma_{ul}^2} \quad (23)$$

which is marked by a gaussian core plus “damping wings.”

For electric dipole transitions, we have a series of **selection rules**, which require:

1. Parity change
2. $\Delta L = 0, \pm 1$
3. $\Delta J = 0, \pm 1$, although $J = 0 \rightarrow 0$ is forbidden
4. Only one single electron wavefunction $n\ell$ (e.g., $1s$ or $2p$) changes, with $\Delta\ell = \pm 1$
5. $\Delta S = 0$, spin cannot change

Forbidden transitions, e.g., [N II] break ≥ 1 rules of 1. through 4. **Semi-forbidden transitions**, e.g., N II] only break rule 5.

The 21 cm line of H I

The hyperfine (spin-flip) transition of neutral hydrogen emits the famous 21 cm (1420 MHz) line. See Fig 1 for a simple graphic. This transition translates to a kinetic temperature of $T_{kin} = 0.07$ K. However, the population ratio of excited to ground state n_2/n_1 is determined by the **spin temperature** or **excitation temperature**, $T_{exc} = T_{spin}$, which depend on:

1. collisions with other hydrogen atoms
2. absorption of thermal radiation (note that $T_{CMB} = 2.7$ K is a *floor*)
3. Scattering of Ly α photons—it can cause $T_{spin} \rightarrow T_{kin}$.

Typical values of $T_{spin} \sim 100$ K. Because this transition is a magnetic dipole transition, the Einstein coefficient $A_{21} = 3 \times 10^{-15} \text{ s}^{-1}$, giving the excited state a lifetime of $\sim 10^{11}$ yr.

Because $kT_{spin} \gg h\nu_{21}$, $n_2/n_1 \simeq g_2/g_1 = 3$; e.g., the upper and lower states are equilibrated and don't depend on the spin temperature.

To get a proper line profile in terms of velocity, we note that

$$\phi(v) = \phi(\nu) \frac{d\nu}{dv} = \phi(\nu) \frac{v_{21}}{c}.$$

Then, the optical depth of the 21 cm line is

$$\begin{aligned} \tau_\nu &= \frac{3c^2 h A_{21}}{32\pi k_B \nu_{21}^2} \frac{N_H}{T_{spin}} \phi(\nu) \\ &= 0.549 \left(\frac{N_{H I}}{10^{21} \text{ cm}^{-2}} \right) \left(\frac{\phi(\nu)}{[\text{km s}^{-1}]^{-1}} \right) \frac{100 \text{ K}}{T_{spin}}. \end{aligned} \quad (24)$$

where the velocity width matters, helping determine whether the line self-absorbs or is optically thin.

The Saha equation

The **Saha equation** tells us the population ratios of *ionization states*.² It depends on the temperature, T , and the number density of electrons, n_e .

Also, we will use the partition function

$$Z^j(T) = \sum_i g_i^j e^{-E_i/kT} = g_0^j n_{tot}^j / n_0^j. \quad (25)$$

So the Saha equation is:

$$\frac{n_{tot}^{j+1} n_e}{n_{tot}^j} = \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} \frac{2Z^{j+1}(T)}{Z^j(T)} e^{-h\nu_{j+1 \rightarrow j}/kT}. \quad (26)$$

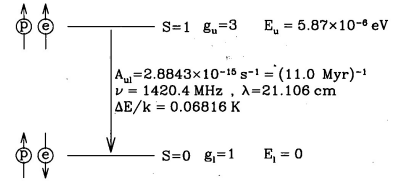


Figure 1: Draine Figure 8.1, depicting the neutral hydrogen 21 cm transition. Note that S represents the hyperfine total angular momentum, not the spin angular momentum. Originally from Gould (1994).

² Previously, we've used the **Boltzmann equation** to figure out *excitation* states. To get a full electron census, we'd want to use both in conjunction.

Pure hydrogen nebula

Surrounding a single, massive star (OB star), we find a radiation field with $h\nu \geq 13.6$ eV, corresponding to $\lambda < 912$ Å. Photons beyond the Lyman limit are called **Lyman continuum** or **Lyman leak** photons.

The rate of production³ of these ionizing photons is Q , where $[Q] = \text{s}^{-1}$:

$$Q = \int_{\nu_*}^{\infty} d\nu \frac{L_\nu}{h\nu}, \quad (27)$$

and where $\nu_* = c/(912 \text{ Å})$, the ionization energy.

Now, we can define the cross-section for ionization as a function of frequency,

$$a_\nu(\nu) = 6.3 \times 10^{-18} \text{ cm}^2 \left(\frac{\nu_*}{\nu}\right)^4 \frac{\exp[4 - (4/\epsilon) \arctan \epsilon]}{1 - \exp(-2\pi/\epsilon)}, \quad (28)$$

where $\epsilon \equiv \sqrt{\nu/\nu_* - 1}$ for convenience. Note that $a_\nu \propto \nu^{-3}$. Then, the rate of ionization per unit volume equals the local rate of recombination per unit volume when

$$n_{\text{H I}} \int_{\nu_*}^{\infty} d\nu \frac{4\pi J_\nu}{h\nu} a_\nu(\nu) = n_e n_p \alpha(T), \quad (29)$$

where $\alpha(T)$ is the recombination coefficient. This equation assumes ionization equilibrium.

H II regions can be **density-bounded**, where ionizing photons escape due to low H I gas density. They can also be **ionization-bounded**, where the edge of the H II region is defined by the extent beyond which no ionizing photons can reach. We can further define $\alpha_B(T)$, the rate of recombination into *all* states other than the ground state.

If we volume-integrate equation (29), the left-hand side becomes Q , as defined by equation (27), and the right-hand side becomes $n_{\text{tot}}^2 \propto \alpha_B(T) V_{\text{HII}} \sim \alpha_B(T) \cdot 4\pi r_S^3/3$, and where

$$r_S = \left(\frac{3Q}{4\pi n_{\text{tot}}^2 \alpha_B(T)} \right)^{1/3} \quad (30)$$

is the **Strömgren radius**.

Wednesday, 16 September

Fiducial H II regions

To get a sense of fiducial numbers, let's get some numbers for an H II region surrounding an O5 star:

Quantity	Value
T_{eff}	46120 K
Q	$3.39 \times 10^{49} \text{ s}^{-1}$
n_{H}	100 cm^{-3}
r_S	4.7 pc

³ We care about the *number* of ionizing photons (not the luminosity) because any photon can only ionize one atom. Excess energy (as counted in the luminosity) does not contribute to further ionizations, but rather becomes kinetic energy to the electron and/or proton.

Note that in this case, r_S was calculated by using $\alpha_B(T = 10^4 \text{ K})$. We've also assumed a high ionization fraction $y = n_e/n_H$ and of course $n_e = n_p$. The rate of recombination into $n \geq 2$, $n_e n_p \alpha_B = 2.6 \times 10^{-9} y^2 \text{ s}^{-1} \text{ cm}^{-3}$.

Thus, the volumetric rate of ionization is

$$\frac{n_{H\text{I}}}{4\pi r^2} \int d\nu \frac{L_\nu}{h\nu} e^{-\tau_\nu} \simeq \frac{n_{H\text{I}} Q a_\nu}{4\pi r^2} = 1.8 \times 10^{-4} (1-y) \left(\frac{r}{\text{pc}} \right)^{-2} \text{ s}^{-1} \text{ cm}^{-3}. \quad (31)$$

If we set $n_e n_p \alpha_B$ equal to the equation above, and input $y = 0.9999$ (e.g., where the H II region begins to become very neutral), we get $r = 3.8 \text{ pc}$. Recall that the entire H II region is 4.7 pc in radius.

The transition from 50% ionized to neutral happens over a mean free path λ_{MFP} of an ionizing photon,

$$\lambda_{\text{MFP}} = \frac{1}{0.5 n_H \nu_T} = 10^{-3} \text{ pc} \ll r_S. \quad (32)$$

This mean free path is comparable to the solar system size (whereas the entire ionized region is far larger).

Absorption lines

We primarily study the intergalactic medium via absorption lines illuminated by a background quasar. We begin by using the equation of radiative transfer,

$$I_\nu = I_{\nu,0} e^{-\tau_\nu} + B_\nu(T_{\text{ex}}) (1 - e^{-\tau_\nu}), \quad (33)$$

where the first term is absorption and the second term is emission at the excitation temperature. In absorption-line systems, we can assume emission to be unimportant (stimulated emission is negligible due to low T_{ex}). Then the absorption coefficient, using the oscillator strength f_{12} becomes

$$\alpha_\nu = \frac{h\nu_{21}}{4\pi} n_1 B_{12} \phi(\nu) = \frac{\pi e^2}{m_e c} f_{12} n_1 \phi(\nu), \quad (34)$$

so that the optical depth is

$$\tau_\nu = \int ds \alpha_\nu = \frac{h\nu_{21}}{4\pi} n_1 B_{12} \phi(\nu) = \frac{\pi e^2}{m_e c} f_{12} N_1 \phi(\nu), \quad (35)$$

where we have assumed that the number density is constant throughout the path in order to compute the column density $N_1 = \int ds n_1$.

We will assume a gaussian line profile,

$$\begin{aligned} \phi_\nu &= \frac{1}{\sqrt{2\pi}} \frac{c}{\sigma_\nu \nu_0} \exp \left(-\frac{1}{2} \left[\frac{\Delta\nu/\nu_0}{\sigma_\nu/c} \right]^2 \right) \\ &= \frac{1}{\sqrt{\pi}} \frac{\lambda}{b} \exp \left(-\left[\frac{\Delta\nu}{b} \right]^2 \right), \end{aligned} \quad (36)$$

and $b \equiv \sigma_v \sqrt{2}$ is the Doppler parameter. And thus the optical depth is

$$\tau_\nu = \frac{\sqrt{\pi} e^2}{m_e c} \frac{N_1 f_{12} \lambda}{b} \exp \left(- \left[\frac{\Delta \nu}{b} \right]^2 \right). \quad (37)$$

The optical depth at the *center* of the line profile, τ_0 , is

$$\begin{aligned} \tau_0 &= \tau_\nu \exp \left(\left[\frac{\Delta \nu}{b} \right]^2 \right) \\ &\approx 0.7580 \left(\frac{N_1}{10^{13} \text{ cm}^{-2}} \right) \left(\frac{f_{12}}{0.4164} \right) \left(\frac{\lambda_{12}}{1215.7} \right) \left(\frac{b}{10 \text{ km s}^{-1}} \right)^{-1}, \end{aligned} \quad (38)$$

and where the fiducial values of the Ly α transition are given.

So how do we measure tiny dips in some continuum at small optical depth? We can define the **equivalent width** of an absorption line (which is important, because it's *independent of spectrograph resolution*):

$$W_\lambda = \int d\lambda \left(1 - \frac{I_\nu}{I_{\nu,0}} \right) = \lambda_0 \int \frac{d\nu}{\nu_0} (1 - e^{-\tau_\nu}). \quad (39)$$

Draine prefers to use the **dimensionless equivalent width**, $W = W_\lambda / \lambda_0$.

The curve of growth

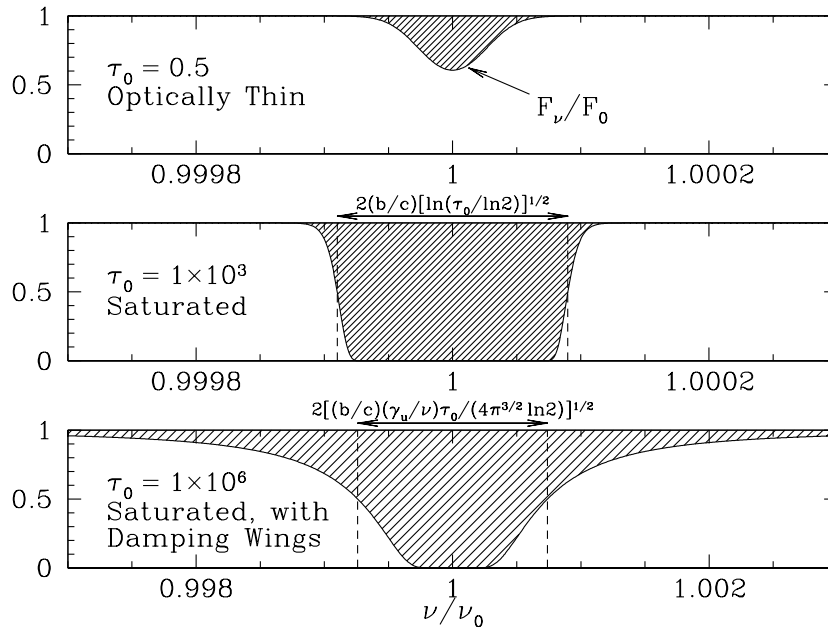


Figure 2: Draine figure 9.1 showing the three regimes of the curve of growth.

There are multiple regimes (three that are important) of optical depth, each of which specifically affect W_λ dependence on the column density N_1 :

1. *Linear regime* ($\tau_0 \leq 1$, so that $1 - e^{-\tau_\nu} \approx \tau_\nu$):

$$W = \int \frac{d\nu}{\nu_0} \tau_\nu = \frac{1}{\nu_0} \frac{\pi e^2}{m_e c} f_{12} N_1 \int d\nu \phi(\nu), \quad (40)$$

or

$$N_1 = \frac{m_e c^2}{\pi e^2} \frac{1}{f_{12} \lambda} \frac{W_\lambda}{\lambda}. \quad (41)$$

2. *Intermediate regime* ($10 \leq \tau_0 \leq 10^3$, so optically thick) We can approximate this by imagining a square profile between $\tau_\nu = 1$ points, so that $W \propto \sqrt{\ln N_1 / b}$.

This is called the “flat” part of the curve of growth, signifying a weak dependence of W_λ on N_1 .

3. *Saturated regime* ($\tau_0 \geq 10^5$). We can use the “damping wings” because the intrinsic broadening dominates over Doppler broadening. If we define the sum over all spontaneous transitions $\gamma_2 = \sum_i A_{2i}$, to get

$$W = \left(\lambda_0 \gamma_2 N_1 f_{12} \lambda \frac{e^2}{\pi m_e c^3} \right)^{1/2} \propto N_1^{1/2}. \quad (42)$$

Therefore, this is the “square root” portion of the curve of growth (or damped portion).

Spectroscopic notation

The wavefunction of a single electron is characterized by four quantum numbers

- n : the mean distance from nucleus
- ℓ : shape of orbital, eigenvalue of J^2 , ranges from 0 to $n - 1$. For example, for $\ell = 0, 1, 2, 3 \rightarrow s, p, d, f$
- m : the eigenvalue of J_z , ranges from $-\ell$ to ℓ
- s : the spin, either $\pm 1/2$

E.g., a notation of $n\ell^x$ means that there are x electrons in some orbital (n, ℓ).

For multiple electrons, we characterized unfilled shells/subshells:

- $L = \sum_i \ell_i$: orbital angular momentum. We use capital letters, e.g., $0, 1, 2, 3 \rightarrow S, P, D, F$
- $S = \sum_i s_i$: spin angular momentum, e.g., where $2S + 1 = 1, 2, 3, 4 \rightarrow$ singlet, doublet, triplet, quartet
- $\vec{J} = \vec{L} + \vec{S}$: total angular momentum, which has degeneracy $2J + 1$.

Electron energies depend on L, S, J : we split terms given by L and S —which then split further by spin-orbit coupling into fine structure levels given by J . For a given L and S , the number of fine structure levels is $\min(2S + 1, L + 1)$.

E.g., we label these as $^{2S+1}L_J$.

The ordering of levels follows **Hund's rules** (for the ground state):

- higher S leads to lower energies
- given S , higher L leads to lower energy
- given L, S , if outer shell is less than half full, “normal” ordering where lower J leads to lower energy. Otherwise it's the other way around

Let's now consider the ground state of singly-ionized nitrogen, N II. The six electrons are then labeled: $1s^2 2s^2 2p^2$, where the last subshell is unfilled.

What if we have two electrons? See the table in the equation sheet—showing which states are allowed via the Pauli exclusion principle. We are allowed $S = 0$ or 1 , and $L = 0, 1, 2$, so we end up with $^1S_0, ^1D_2$, and $^3P_0, ^3P_1, ^3P_2$.

The full level diagram looks like Draine figure 4.1, which is shown in the margin.

Friday, 18 September

Recombination of hydrogen

A **recombination cascade** occurs when electrons have more than one recombination/de-excitation states to choose from, e.g., $n = 3 \rightarrow n = 2$ ($H\alpha$) or $n = 3 \rightarrow n = 1$ ($Ly\beta$).

There are two regimes for recombination:

- Case A recombination: all photons produced in the recombination cascade escape (an H II region), so that the region is optically thin, or equivalently, density-bounded.
- Case B recombination (usually more relevant): all non-Lyman series photons escape, but Lyman line and Lyman continuum photons (LyC) *besides* $Ly\alpha$ are instantly reabsorbed. The H II region is therefore ionization-bounded, or equivalently optically thick.

Remembering our selection rules, and considering hydrogen, we see that Lyman photons must go from $np \rightarrow 1s$. From $3p \rightarrow 1s$, the $Ly\beta$ photon is very easily absorbed when on-resonance by another electron in the $1s$ state (even more easily than LyC photons). Since Case B recombination is far more common in H II regions, the only way for photons to escape is via the cascade, where we get something like $3p \rightarrow 2s \rightarrow 1s$, producing $Ly\alpha$ and $H\beta$ photons that can escape.⁴

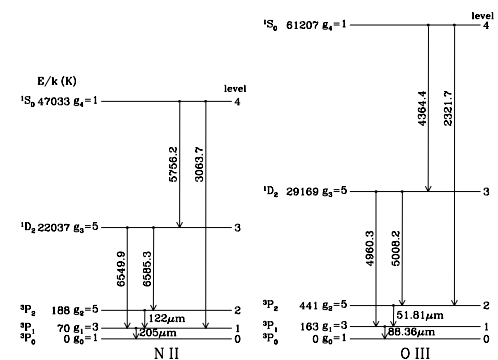


Figure 3: See Draine figure 4.1 showing the electron configurations for N II and O III.

⁴ Note that the $2s \rightarrow 1s$ transition is heavily forbidden, and in high density regions, $2s \rightarrow 2p \rightarrow 1s$ (via collisional excitation and then spontaneous emission, respectively) is more favored.

A generalization can be made: every Lyman photon with upper level $n \geq 3$ will be converted into either

- either a Ly α photon or two continuum photons (that sum to Ly α energy)
- one Balmer series photon
- sometimes photons from higher series

Ly α is resonantly scattered, because $2p \rightarrow 1s$ photons can easily excite another nearby neutral hydrogen. They also might get absorbed by dust. Ly α photons have very long path lengths in the ISM because they are always resonant scattering and random walking through a far greater column density (than, say, LyC photons, which do not resonantly scatter). This is called dust quenching.⁵

⁵ Actually, this depends on the dust morphology in or near an H II region. Recent literature have shown that Ly α photons do not get significantly more quenched than LyC photons.

Star formation rates

H α is nice because it will always leave the H II region. Therefore it traces the star formation rate quite well.

$$L_{H\alpha} \propto n_e n_p \alpha_B \Delta V \propto Q. \quad (43)$$

The rate of photons emitted from a *cluster* of stars is

$$Q_{\text{cluster}} = M_{\text{tot}} \int_{m_L}^{m_U} dm \int_{v_T}^{\infty} dv \frac{L_v(m)}{h\nu} \Phi(m), \quad (44)$$

where $\Phi(m)$ is the initial mass function (IMF), which determines the mass distribution of stars as they are formed. Kennicutt (1998) reviews SFR estimation techniques.

Collisional transitions and critical density

Let's assume a two (or more) level set of energy states, with $E_2 = E_1 + h\nu_{21}$. The relative populations can be affected by absorption and emission, as well as collisions with dominant “collider” species (electrons, H, H₂) that may impart $\Delta E > h\nu_{21}$ and excite the state (**collisional excitation**), or steal away exactly $h\nu_{21}$ (**collisional de-excitation**).

In order for collisional excitation, the collider must have a minimum kinetic energy $mv_{\text{min}}^2/2 \geq h\nu_{21}$. We can represent these as cross sections, where

$$\gamma_{12} = \langle \sigma_{12} v \rangle = \int_{v_{\text{min}}}^{\infty} dv \sigma_{12} f(v), \quad (45)$$

where $[\gamma_{12}] = [\text{cm}^3 \text{s}^{-1}]$. However, to get a volumetric rate of collisions, we'd want $n_{\text{collider}} n_1 \gamma_{12}$. For de-excitations, we'd imagine $v_{\text{min}} = 0$ and simply reverse the subscripts $1 \leftrightarrow 2$.

For a velocity distribution $f(v)$, we often will assume **Local Thermodynamic Equilibrium**, or LTE. In this case T_{kin} is well-defined and our velocity distribution is a Maxwellian distribution (properly normalized):

$$f(v) = \left(\frac{m_{\text{collider}}}{2\pi k T_{\text{kin}}} \right)^{3/2} v^2 \exp \left[-\frac{m_{\text{collider}} v^2}{2k T_{\text{kin}}} \right]. \quad (46)$$

Averaging σv over a Maxwellian leads to the simple relation $\langle \sigma v \rangle \propto T_{\text{kin}}^{-1/2}$.

In LTE, our excitation and de-excitations are equal ($n_{\text{collider}} n_1 \gamma_{12} = n_{\text{collider}} n_2 \gamma_{21}$), leading to our well-known

$$\frac{\gamma_{12}}{\gamma_{21}} = \frac{n_2}{n_1} = \frac{g_2}{g_1} \exp \left(-\frac{h\nu_{21}}{k T_{\text{kin}}} \right),$$

which makes sense, because T_{kin} acts in the same way as a thermal bath of photons (even though we're imagining a bunch of colliding particles). This is true in LTE without radiative transitions (that we never accounted for). However, these derived relationships for collisional transition coefficients are true *even when radiative transitions are included*—so long as we're in LTE.

What if we assume non-LTE with both collisional and radiative transitions important? Then, we combine their terms⁶

⁶ I'm abbreviating $n_{\text{collider}} = n_{\text{coll}}$

$$n_1 (n_{\text{coll}} \gamma_{12} + B_{12} \bar{J}) = n_2 (n_{\text{coll}} \gamma_{21} + B_{21} \bar{J} + A_{21}), \quad (47)$$

and still substitute the B coefficients with previously-derived definitions:

$$\frac{n_2}{n_1} = \frac{n_{\text{coll}} \gamma_{12} + (g_2/g_1)(c^2/2h\nu_{21}^3) A_{21} \bar{J}}{n_{\text{coll}} \gamma_{21} + (c^2/2h\nu_{21}^3) A_{21} \bar{J} + A_{21}}. \quad (48)$$

For now, if we assume $\bar{J} \rightarrow 0$, then, using a **critical density** $n_{\text{crit}} \equiv A_{21}/\gamma_{21}$:

$$\frac{n_2}{n_1} \rightarrow \frac{\gamma_{12}/\gamma_{21}}{1 + (n_{\text{crit}}/n_{\text{coll}})}. \quad (49)$$

Thus, we now find that

$$\frac{n_2}{n_1} \rightarrow \frac{g_2}{g_1} \frac{e^{-h\nu_{21}/k T_{\text{kin}}}}{1 + (n_{\text{crit}}/n_{\text{coll}})}. \quad (50)$$

In the *high-density* regime, where $n_{\text{coll}} \gg n_{\text{crit}}$, we get

$$\frac{n_2}{n_1} \rightarrow \frac{g_2}{g_1} \exp \left(-\frac{h\nu_{21}}{k T_{\text{kin}}} \right),$$

signifying that the excitation temperature is equal to the kinetic temperature ($T_{\text{exc}} = T_{\text{kin}}$). In the high-density regime,

$$n_2 n_{\text{coll}} \gamma_{21} \gg n_2 A_{21}.$$

In the *low-density* regime, where $n_{\text{coll}} \ll n_{\text{crit}}$, then

$$\frac{T_{\text{exc}} - T_{\text{kin}}}{T_{\text{exc}} T_{\text{kin}}} = \frac{k}{h\nu_{21}} \ln \left(\frac{n_{\text{coll}}}{n_{\text{crit}}} \right) < 0,$$

so that $T_{\text{exc}} < T_{\text{kin}}$, i.e., the excitations are *sub-thermal*. The collisions can't thermalize our population levels

Let's now drop our assumption, so $\bar{J} \neq 0$. We'll still have to make some other assumptions, namely that photons produced locally can only be absorbed locally (e.g., this can happen with some velocity gradient in the ISM). We will define the **escape probability**, $\beta(\tau_{\nu_{21}})$, that a photon formed at an optical depth $\tau_{\nu_{21}}$ from the "surface" manages to escape:⁷

$$\tau_{\nu_{21}} = \frac{c^3}{8\pi h \nu_{21}^3} A_{21} n_2 \left(\frac{n_1 g_2}{n_2 g_1} - 1 \right) \frac{\Delta s}{\Delta v}. \quad (51)$$

⁷ See Lecture 6 notes equations (21-24) for derivation. Note that we've approximated $\phi(v) = (c/v)\phi(v) \approx c\Delta v/v$.

This previous result is referred to as the "large velocity gradient" formalism for optically thick radiative transfer.

So, revisiting our previous equations, where we've set the number of photons absorbed equal to the number that haven't escaped, we will define a new $n'_{\text{crit}} = \beta(\tau_{\nu_{21}}) A_{21} / \gamma_{21}$, and find

$$\frac{n_2}{n_1} \rightarrow \frac{g_2}{g_1} \frac{e^{-h\nu_{21}/kT_{\text{kin}}}}{1 + (n'_{\text{crit}}/n_{\text{coll}})}. \quad (52)$$

Here, **radiative trapping** absorptions help colliders maintain thermal equilibrium (against spontaneous emission), and reduces the effective critical density. See Draine chapter 19 for values of escape probabilities, $\beta(\tau_{\nu})$.

Tuesday, 22 September

The [C II] fine structure line

Note that all fine structure lines, e.g., where the spin orbit coupling effect causes ΔL and $\Delta S = 0$. However, fine structure transitions don't change parity, and neither does a single electron wavefunction ($n\ell$) change.

The [C II] fine structure line at $\lambda = 158 \mu\text{m}$ is the dominant cooling line for diffuse gas in star-forming galaxies. Draine writes the spectroscopic notation as $^2P_{3/2}^o \rightarrow ^2P_{1/2}^o$ (the *o* is for odd parity; and a nothing for even parity) in his Figure 17.3.

Line luminosities per unit volume

Let's start from the equation of statistical balance:

$$n_1 (n_{\text{coll}} \gamma_{12} + B_{12} \bar{J}) = n_2 (n_{\text{coll}} \gamma_{21} + B_{21} \bar{J} + A_{21})$$

and multiple by $h\nu_{21}$ to get local rates of energy transfer per unit volume. These terms have taken into account all emission and absorption due to spontaneous, stimulated, and collisional effects.

Then the net line luminosity per unit volume (with units $\text{erg s}^{-1} \text{cm}^{-3}$) is:

$$\Lambda_{21} = h\nu_{21} (n_2 B_{21} \bar{J} - n_1 B_{12} \bar{J} + n_2 A_{21}). \quad (53)$$

Note: in general, Λ is a volumetric cooling rate. There is also an analogous volumetric heating rate, Γ . However, some authors use $\Lambda \rightarrow n^2 \Lambda$ and $\Gamma \rightarrow n \Gamma$.

Why did we drop the terms involving collisions? They aren't associated with *cooling*—i.e., there's no radiative energy transfer.

However, not all these photons will leave our H II region; that would depend on the optical depth, $\tau_{\nu_{21}}$ (we define the optical depth equal to zero at the cloud's surface, and increasing with penetration). We will use the idea of the *escape probability* defined last time, $\beta(\tau_{\nu_{21}})$ to summarize the complex radiative transfer, now giving us

$$(n_1 B_{21} - n_2 B_{21}) \bar{J} = n_2 (1 - \beta(\tau_{\nu_{21}})) A_{21}. \quad (54)$$

Substituting this back into our definition of the net line luminosity per unit volume, we get nice cancelations:

$$\Lambda_{21} = h\nu_{21} \beta(\tau_{\nu_{21}}) n_2 A_{21}. \quad (55)$$

This makes sense,⁸ because the cooling rate is equal to the rate of emitted photons times the escape fraction.

If we wanted to calculate the emergent luminosity from the entire cloud, we simply integrate: $L_{21} = \int dV \Lambda_{21}$, where obviously Λ_{21} is dependent on position.

⁸ We're assuming that absorption dominates over stimulated emission, which is true *except* in the case of an inverted population. In the limit of small optical depth, absorption is negligible, and likewise stimulated emission, so that only A_{21} is important.

Heavier elements in H II regions

Helium can ionize and form its own region called a He II region. Because the ionization energy for He ($h\nu \geq 24.6$ eV) is larger than for H, then the He II region will be smaller (only for the most massive stars—O1 to O3 stars will the H I region be of comparable size). The two regions are actually coupled, though, because He recombination lines can ionize H I, but we won't go through those calculations.

What about metals? Elements such as O, Ne, N, S, etc. are in low abundance, but have energy levels near the ground state due to fine structure ($\vec{L} + \vec{S}$) and term (L, S) splitting of multiple electrons. They can easily be populated by collisional excitations. But collisional de-excitation doesn't happen often, so *forbidden lines do the job of cooling—and they can rapidly cool the cloud this way*.⁹

From equation (55), we see that the strength of cooling depends on n_2 and $\beta(\tau_{\nu_{21}})$. n_2 is governed by electrons (over protons, due to their greater velocities). In a neutral or molecular region, H I or H₂ will be the respective dominant colliders. Metal lines are optically thin so $\beta \rightarrow 1$. We can set $\bar{J} = 0$ because the frequency-averaged local radiation field is negligible for a line.

Using the level population ratios at *low density*,

$$\frac{n_2}{n_1} = \frac{g_2}{g_1} \exp\left(-\frac{h\nu_{21}}{kT_{\text{kin}}}\right) \frac{n_{\text{coll}}}{n_{\text{crit}}}, \quad (56)$$

and we find that (referencing equations 8-10 on the Lecture 7 equation sheet):

$$\Lambda_{21} = h\nu_{21} n_{\text{coll}} n_1 \gamma_{12}. \quad (57)$$

⁹ In high density environments such as our labs, collisional excitations easily depopulate via collisional de-excitation. But not so in the near-vacuum of space.

In a **low density region**, every excitation will lead to a spontaneous emission. That is why the result does not depend on A_{21} !

In the **high density region**, we can follow Lecture 7 equations (11-14) to find

$$\Lambda_{21} = h\nu_{21}n_{\text{crit}}n_1\gamma_{12}\frac{1}{1 + n_{\text{crit}}/n_{\text{coll}}} \quad (58)$$

$$= h\nu_{21}n_{\text{coll}}n_1\gamma_{12}\frac{1}{1 + n_{\text{coll}}/n_{\text{crit}}}. \quad (59)$$

Here, A_{21} appears implicitly because it's factored into n_{crit} . The point is that this cooling rate will be lower than in the low density regime, because collisional de-excitations are taking up the energy which would otherwise be radiatively cooled (and collisional de-excitation does *not* cause energy to leave the cloud).

For electron collisions, we will define **collision strengths** Ω_{21} and Ω_{12} implicitly via

$$\gamma_{21} = \frac{h^2}{(2\pi m_e)^{3/2}(kT_{\text{kin}})^{1/2}} \frac{\Omega_{21}}{g_2} \quad (60)$$

$$\gamma_{12} = \frac{h^2}{(2\pi m_e)^{3/2}(kT_{\text{kin}})^{1/2}} \frac{\Omega_{21}}{g_1} \exp\left(-\frac{h\nu_{21}}{kT_{\text{kin}}}\right). \quad (61)$$

Regarding **dust extinction**, we use the term to refer to *both* attenuation (absorption of photons) and scattering of photons out of our line-of-sight. To measure the dust content, we assume a foreground dust screen (which is a good assumption for the Milky Way). Dust extinction is usually worse at shorter wavelengths.

Generally extinction, A_λ , is parametrized in *magnitudes* via

$$m_\lambda = -2.5 \log\left(\frac{F_\lambda}{F_{\lambda,0}}\right) + A_\lambda. \quad (62)$$

For AB magnitudes, $m_\lambda = 23.9 - 2.5 \log(F_\lambda / \text{mJy}) + A_\lambda$, where $1 \mu\text{Jy} = 10^{-29} \text{erg cm}^{-2} \text{s}^{-1} \text{Hz}^{-1}$.

The **dust law** is given by

$$k(\lambda) = \frac{A_\lambda}{A_B - A_V} \equiv \frac{A_\lambda}{E(B - V)}, \quad (63)$$

which is in terms of the **reddening**, $E(B - V) = A_B - A_V$ (because of the preferential extinguishing of bluer light).

To study via emission lines, we could use $H\alpha$ and $H\beta$, which is called the **Balmer decrement**:

$$\left[\frac{F_{H\alpha}}{F_{H\beta}}\right]_{\text{obs}} = 2.86 \times 10^{0.4(A_{H\beta} - A_{H\alpha})}, \quad (64)$$

which, for case B recombination, $T_{\text{kin}} = 10^4 \text{ K}$, has a value of 2.86. To derive $A_{H\alpha}$ or $E(B - V)$:

$$A_{H\beta} - A_{H\alpha} = A_{H\alpha} \left(\frac{k(H\beta)}{k(H\alpha)} - 1 \right).$$

Dust in H II regions systematically causes us to systematically underestimate the intrinsic Q value, and hence, SFR.

Friday, 25 September

Measuring kinetic temperatures in H II regions

Remember that if we're looking to characterize the dust extinction, we must measure two different lines of the same species (same element and ionization state). Those lines should be far apart, so as to be able to generalize a dust slope. If we want to measure the *temperature* (also useful for characterizing the dust), then we must measure the ratio of nearby lines to avoid dust sensitivity.

One set of options is¹⁰

[N II] 5755 ($^1D_2 - ^1S_0$)

[N II] 6548 ($^3P_1 - ^1D_2$)

[N II] 6583 ($^3P_2 - ^1D_2$)

¹⁰ see Draine Fig 6.1.

These lines are forbidden because parity $\Pi_i(-1)^{\ell_i}$ are unchanged, *and* because no single electron wavefunctions change.

The latter two are considered a **doublet** because they are close in wavelength, although they aren't necessarily related quantum mechanically. They also straddle the H α line, which is convenient. These lines are collisionally excited by can only be de-excited above the critical density, $\gtrsim 100\text{cm}^{-3}$. So we still see the lines.

The cooling rate of the [N II] 5755 line is given by

$$\Lambda_{SD} = h\nu_{SD}n_en_p\gamma_{PS}\frac{A_{SD}}{A_{SD} + A_{SP}},$$

where n_p is the number density of ionized nitrogen in the 3P state.

If we treat the latter as a single line, we can find the cooling rate

$$\Lambda_{DP} = h\nu_{DP}n_en_p\left(\gamma_{PD} + \gamma_{PD}\frac{A_{SD}}{A_{SD} + A_{SP}}\right).$$

Intermediate steps (equations 3-5 on the Lecture 8 sheet) lead to the ratio of line intensities:

$$\frac{\Lambda_{5755}}{\Lambda_{6548} + \Lambda_{6583}} = \frac{\Lambda_{SD}}{\Lambda_{DP}} \simeq \frac{\nu_{SD}}{\nu_{DP}} \frac{\gamma_{SP}}{\gamma_{DP}} \frac{g_S}{g_D} \frac{A_{SD}}{A_{SD} + A_{SP}} \exp\left(-\frac{h\nu_{SD}}{kT_{\text{kin}}}\right). \quad (65)$$

The final bit of this equation allows us to figure out the kinetic temperature ($\propto e^{-25000\text{K}/T_{\text{kin}}}$, so it's sensitive to temperatures *below* 25000 K). Note that the first line is hardly observed outside of our galaxies.

Another option, which is more commonly used, is O III (which has the same number of electrons and similar transitions to N II)

$$[\text{O III}] 4363 (^1D_2 - ^1S_0)$$

$$[\text{O III}] 4959 (^3P_1 - ^1D_2)$$

$$[\text{O III}] 5007 (^3P_2 - ^1D_2)$$

Even the latter two lines here form a (crude) doublet. Energy splittings are similar to [N II].

We can derive general solutions (or rather, Osterbrock & Ferland) that are valid above low-density, i.e., $n_e \gtrsim 10^5 \text{ cm}^{-3}$. These are given as equations (7-8) on the Lecture 8 equation sheet for [N II] and [O III] respectively. Qualitatively, we see that a *high temperature* gives us a smaller line ratio, whereas *lower temperatures* give us a larger line ratio (because fewer electrons will be in the higher energy 1S_0 state). From these equations, we see that H II region temperatures are $T \approx 7000 - 14000 \text{ K}$.

Measuring density in H II regions

In this case, we want transitions with the same *lower* level. A common choice is to use

$$[\text{O II}] 3726 (^4S_{3/2} - ^2D_{3/2})$$

$$[\text{O II}] 3729 (^4S_{3/2} - ^2D_{5/2})$$

The degeneracy factors are given in the spectroscopic notation ($g_i = 2J_i + 1$), so in this case, our three states $^4S_{3/2}$, $^2D_{5/2}$, and $^2D_{3/2}$ have respective degeneracy factors of $g_1 = 4$, $g_2 = 6$, and $g_3 = 4$.

In the *low density limit*, $n_e \ll n_{\text{crit}}$, we get a single photon for every collisional excitation. Note that $A_{32} \ll A_{31}$ because the gap is so extremely small between the second and third fine-structure states of O II.

Similarly to before, we can call equations (9), (10) from the Lecture 8 equation sheet. A **line ratio** of the two gives

$$\frac{\Lambda_{3729}}{\Lambda_{3726}} = \frac{\gamma_{12}}{\gamma_{13}} = 1.5,$$

which shouldn't be surprising—the energies are so similar, so the line ratio is simply the ratio of degeneracies ($g_2/g_3 = 1.5$).

What about in the *high density limit* ($n_e \gg n_{\text{crit}}$)? Now things are a little more complicated. Equations (12-16) show all this, culminate in

$$\frac{\Lambda_{3729}}{\Lambda_{3726}} \simeq \frac{g_2 A_{21}}{g_3 A_{31}} \approx 0.35.$$

The Einstein coefficients are the difference-maker here. This is because our lines are thermalized, and now transitions between the second to third levels may occur far more commonly due to collisions.

See Draine Fig 18.4 in the margin. This diagnostic is sensitive to $n_e (T/10^4 \text{ K})^{-1/2}$ over a range of $10^2 - 10^4 \text{ cm}^{-3}$. So we can first measure the temperature (as shown above), and then use it to constrain the density of the gas.

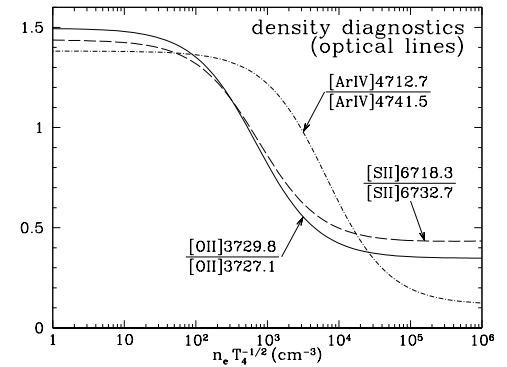


Figure 4: Draine Figure 18.4, showing optical line intensity ratios useful for determining the density.

Measuring (distant) ionization sources

Notice the **BPT Diagram**, shown in Draine Fig 18.7—which allows distinction of AGN from star forming sources.

So why are we using [O III] 5007 / H β 4861 and [N II] 6583 / H α 6563? The ionization potentials of N I, N II, and O II are 14.5, 29.6, 35.2 eV, respectively. The reason that the BPT star forming component curves to the upper left is because [N II] decreases and gives way to [N III] as [O III] becomes stronger.

Measuring metallicity

Metallicity is usually defined via the relative oxygen to hydrogen abundance ratio. However, ionized oxygen is very common (as we've seen), and we need to sum up *all* ionization states of oxygen to get a proper metallicity. One common method (Pagel et al. 1979) is an empirical indicator:

$$R_{23} = \frac{I_{[\text{O II}]3727} + I_{[\text{O III}]4959,5007}}{I_{\text{H}\beta}}. \quad (66)$$

Unfortunately, note that [O/H] is actually non-monotonic with R_{23} . Metal lines are efficient for cooling (fine structure lines in the far-infrared), which is why at high metallicity, R_{23} decreases. Then how do you break this degeneracy? We can use the [N II] 6583 / H α 6563 line ratio, although this is imperfect too (no [N III], etc).

The **ionization parameter** U is

$$U = \frac{Q}{4\pi r^2 c n_H}, \quad (67)$$

which is basically $U = n_\gamma / n_H$.¹¹ Empirically, we can use

¹¹ Think of $n_\gamma = Q dt / 4\pi r^2 dr$.

$$O_{32} = \frac{I_{[\text{O III}]4959,5007}}{I_{[\text{O II}]3727}}$$

to calibrate R_{23} .

Tuesday, 29 September

Thermal equilibrium in H II regions

We kept assuming fiducial values of $T \sim 10^4$ K. How did we derive that?

In the case of thermal equilibrium, heating balances cooling:

$$\Gamma_{\text{tot}} = \Lambda_{\text{tot}}.$$

The imparted excess kinetic energy from photoionization will be the main heating mechanism. The average extra energy is $\langle \Delta E \rangle = h \langle \nu \rangle - h\nu_0$, where $h\nu_0$ is the ionization energy.

Accounting for *all* photons that are around, we find that the heating rate is

$$\begin{aligned}\Gamma_{\text{PI}} &= n_e n_p \alpha_A \langle \Delta E \rangle \\ &= n_e n_p \alpha_A (T_{\text{kin}}) \frac{\int_{\nu_0}^{\infty} d\nu \frac{4\pi J_\nu}{h\nu} h(\nu - \nu_0) a_\nu}{\int_{\nu_0}^{\infty} d\nu \frac{4\pi J_\nu}{h\nu} \alpha_\nu},\end{aligned}\quad (68)$$

where we've included the un-normalized distribution of ionization events by averaging $h(\nu - \nu_0)$. Other species behave similarly, but H dominates. Of course, this quantity will depend on the shape of the local ionization spectrum, J_ν for $\nu > \nu_0$.

The characteristic initial temperature T_{init} of photoionized electrons is

$$\langle \Delta E \rangle = \frac{3}{2} k T_{\text{init}}.$$

Note that $T_{\text{kin}} \neq T_{\text{init}}$. However, Γ_{PI} depends on T_{kin} via α_A (recombination coefficient), with $\alpha_A \propto T_{\text{kin}}^{-1/2}$.

The timescale for an ejected electron to collide with another electron is ~ 100 s for $n_e \simeq n_H \sim 100 \text{ cm}^{-3}$, and this timescale is inversely proportional to number density. Compare with the timescale for ejected electrons to recombine, $t_{\text{rec}} = (\alpha_B n_p)^{-1} \sim 1200$ yr. Because newly ejected electrons collide many times before recombining, the entire H II regions are basically in a *thermal bath* of electrons.

There are three main processes that cool ionized gas:

1. Recombination

We assume that the initial kinetic energy ($\frac{1}{2} m_e v^2$) of an electron is lost when it recombines with a nucleus. If $\sigma_n(T_{\text{kin}})$ is a cross-section for recombination into level n , and $f(v)$ is a normalized Maxwellian velocity distribution,

$$\Lambda_{\text{rec}} = n_e n_p \sum_{n=1}^{\infty} \int_0^{\infty} dv v \sigma_n(T_{\text{kin}}) \frac{1}{2} m_e v^2 f(v). \quad (69)$$

Note also that $\sigma_n \propto v^{-2}$, such that the *slowest* electrons are recombined first. This tends to make $T_{\text{kin}} > T_{\text{init}}$ (which might seem weird for “cooling”). The temperature dependence of recombination is $\Lambda_{\text{rec}} \propto T_{\text{kin}}^{1/4}$.

2. Free-free emission

We've already discussed the power density of free-free radiation,

$$\frac{dW}{dv dV dt} = C_4 n_e^2 T_{\text{kin}}^{-1/2} \ln \left(C_3 T_{\text{kin}}^{3/2} / \nu \right) \exp \left(- \frac{h\nu}{k T_{\text{kin}}} \right).$$

To get a cooling rate, we must integrate over all ν (and we'll treat the Gaunt factor as approximately constant): $\Lambda_{\text{ff}} \propto n_e^2 T_{\text{kin}}^{1/2}$.

Aside: the mean energy decreases as we go further away from the star. This is because the photon-electron cross section *decreases* as we increase in energy above the ionization threshold.

3. Metal lines

These turn out to be the dominant cooling mechanisms in H II regions and planetary nebulae. The result from collisional excitation followed by forbidden line radiation.¹² The net effect converts kinetic energy to radiation, which is then *optically thin* in the region.

As we've calculated before,

$$\begin{aligned}\Lambda_{21} &= h\nu_{21}n_2A_{21} \\ &= h\nu_{21}n_en_1\gamma_{12}\frac{1}{1 + (n_e/n_{\text{crit}})}.\end{aligned}\quad (70)$$

So if we assume low $\tau_{\nu_{21}}$ with $\beta(\tau_{\nu_{21}}) \rightarrow 1$, then the total forbidden line cooling rate is the summed cooling rates over all species. It depends on

the local abundance of a given element

the distribution of elements' ionization states

the local density relative to a transition's critical density

For an H II region with solar metallicity, Z_{\odot} , $T_{\text{kin}} = 10^4$ K, $n_H = 100 \text{ cm}^{-3}$, dominant cooling lines are

- [O II] 3726+3729 $n_{\text{crit}} = 3300, 630 \text{ cm}^{-3}$
- [O III] 4959+5007 $n_{\text{crit}} = 6.5 \times 10^5 \text{ cm}^{-3}$
- [N II] 6458+6583 $n_{\text{crit}} = 7.8 \times 10^4 \text{ cm}^{-3}$

These are *strongly* sensitive to T_{kin} via $n_2 \propto \exp(-h\nu_{21}/kT_{\text{kin}})$.

So far, all cooling and heating rates are proportional to n^2 , even in the case of species like oxygen, e.g., $n_en_O = n_e^2(n_O/n_H) \propto n_H^2$. So T_{equil} is *nearly* independent of density. However, the caveat is the n_{coll} vs n_{crit} terms given in the metal line cooling rates.

So how do we get to an H II region temperature? We plot the heating and cooling lines and see at what temperature they balance, i.e., see Draine Fig. 27.1 (top panel). If we change the metal abundances (Draine Fig. 27.2), then of course the metal line cooling (which dominates) will be change, so that the heating-cooling balance will affect the equilibrium temperature. And if we change the densities, n_H , we change the collisional de-excitation rates as shown in Draine Fig. 27.3. All of these are of order $T \sim 10^4$ K.

Thermal equilibrium in neutral atomic gas

In a primarily neutral region, there are multiple *heating* mechanisms and only one main cooling mechanism.

The main coolant is [C II] ($^3\text{P}_{3/2}^0 \leftarrow ^2\text{P}_{1/2}^0$) at $158 \mu\text{m}$. Carbon is the second most abundant element in the ISM, typically one-third is in the gas

¹² Hydrogen and helium don't operate this way, because their ground states are in the 1s orbitals, whereby a 2s state requires a relatively large increase in energy. And even when that does happen, the photons can resonantly scatter.

Note that there is some non-zero ionized fraction of hydrogen.

phase, and the rest is in molecular gas and dust. Most of the atomic hydrogen is C^+ . The [C II] splitting of energy levels is $h\nu_{21}/k = 92$ K.

The dominant collider population in neutral gas is *still* electrons—if and only if we're at low densities, $n_H < 3000 \text{ cm}^{-3}$ for typical ionization fraction $y \sim 10^{-3}$. But neutral atoms are the dominant colliders if n_H is higher than that.

The rate of [C II] colling in low-density limit ($n_{\text{coll}} < n_{\text{crit}}$) is given by

$$\frac{\Lambda_{[\text{C II}]}}{\text{erg cm}^{-3} \text{ s}^{-1}} \simeq 3 \times 10^{27} \left(\frac{n_H}{\text{cm}^{-3}} \right)^2 \left(\frac{n_C/n_H}{1.4 \times 10^{-4}} \right) \left(1 + 0.42(y/10^{-3}) \right) \exp \left(-\frac{92 \text{ K}}{T_{\text{kin}}} \right). \quad (71)$$

Draine Fig. 30.1 shows the cooling rates for different cooling mechanisms, where we see that [C II] dominates at low temperatures and lower densities.

What about heating? Photons with energy less than 13.6 eV can ionize elements with low ionization potentials, e.g., C, S, Si (with potentials 11.3, 10.4, 8.2 eV). Most of these will be ionized in the ISM. But this is a *small* source of heating.

Dominant sources of heating will be by

1. Cosmic rays

We typically talk about highly energetic protons with $1 - 10$ MeV, and subtracting $\text{Ly}\alpha$ radiations, we get a primary energy of $E_{\text{prim}} = 7$ eV per ionization. The heating rate is simply

$$\Gamma_{\text{CR}} = n_{\text{H I}} E_{\text{prim}} \zeta_{\text{CR}} \sim 4.6 \times 10^{-27} \left(\frac{n_{\text{H I}}}{\text{cm}^{-3}} \right) \text{ erg cm}^{-3} \text{ s}^{-1}, \quad (72)$$

where canonical value of the rate of cosmic ray primary ionizations is $\zeta_{\text{CR}} \simeq 3 \times 10^{-17} \text{ s}^{-1}$.

2. X-rays

Hard X-rays have $E > 250$ eV, and $\zeta_{\text{XR}} \simeq 10^{-18} \text{ s}^{-1}$. They're not a huge deal but still relevant.

3. Dust

Photoelectric heating of small dust grains (including polycyclic aromatic hydrocarbons, **PAHs**) dominates the heating mechanism of the neutral atomic gas. The net effect of dust absorptions is the creation of free electrons, which add kinetic energy to the gas.

The **yield** is the probability that the dust escapes multiplied by the fraction of the photon's energy that is carried off as kinetic energy. For large grains, the yield is about 5%, but close to 1% for small grains. For PAHs, the yield is $\sim 15\%$. All this leads to a local heating rate (by the photoelectric effect on dust)

$$\Gamma_{\text{PE}} \simeq 5 \times 10^{-26} \left(\frac{n_H}{\text{cm}^{-3}} \right) \text{ erg cm}^{-3} \text{ s}^{-1}. \quad (73)$$

The interstellar radiation field, or **ISRF** is mostly produced by local stars with intensity G_0 . In the FUV, we measure the intensity using units of **Habing**: $1 \text{ Habing} = 1.2 \times 10^{-4} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ ster}^{-1}$. The local ISRF has $G_0 = 1.7$ Habing.

Friday, 2 October

Neutral atomic gas

Thermal equilibrium means that heating and cooling are balanced, such that

$$\frac{\Gamma}{n_H^2} = \frac{\Lambda}{n_H^2}.$$

If we assume the ideal gas law applies, and we know that these rates are functions of temperature, e.g., $\Lambda/n_H^2 = f(T)$, so we can write

$$\frac{\Gamma}{n_H} \frac{kT}{P} = f(T).$$

Because most heating is due to dust, and dust heating depends on the density of atoms and density of photons, so that Γ/n_H is nearly independent of density (and neither is it dependent on temperature).

Now, *assuming a low density limit* (for convenience), we know that we can use (57) and find

$$\Lambda_{21} \propto T_{\text{kin}}^{-1/2} \exp\left(-\frac{92 \text{ K}}{T_{\text{kin}}}\right)$$

We examine Draine figures 30.2 (a) and (b), taking note of panel (b) especially. At the intersection of the cold and warm media, there is *pressure equilibrium* but *not thermal equilibrium* (between 3050 and $4730 \text{ cm}^{-3} \text{ K}$): this **two-phase ISM** has compressed cold pockets of cold neutral medium (CNM) surrounded by warm neutral medium (WNM). There is also a warm ionized medium (WIM).

This thermal instability is also called an **isobaric criterion**, and is formally defined as when

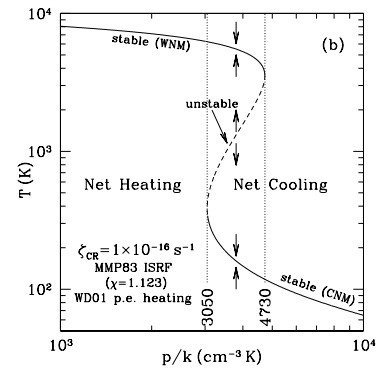
$$\left. \frac{\partial(\Lambda - \Gamma)}{\partial T} \right|_P < 0. \quad (74)$$

From above, we can find

$$n_H = \left(\frac{\Lambda}{f(T)} \right)^{1/2}.$$

It turns out that we get a two-phase neutral medium IFF $1000 \text{ cm}^{-3} \text{ K} \leq P/k \leq 3600 \text{ cm}^{-3} \text{ K}$. Or, assuming $P/k = 3000 \text{ cm}^{-3} \text{ K}$:

	n_H [cm ⁻³]	T_{kin} [K]	y	t_{rec} [yr]	t_{therm} [yr]	t_{sc} [Myr]	Filling factor	MW scale height [pc]
CNM	60	50	10^{-3}	6×10^4	3600	0.5	1%	94
WNM	0.4	7500	0.1	$\sim 6 \times 10^5$	1.6×10^6	16	30%	~ 300
WIM	0.1	8000					25%	~ 900



Some comments:

- Why is the WNM ionization fraction (y) so much higher? It's because the CNM recombination rates are far higher (due to higher n_H).
- The **recombination timescale** (of a proton) is given by

$$t_{\text{rec}} = (\alpha_B n_e)^{-1} \sim 200 \left(\frac{T^{0.6} / \text{K}}{n_e / \text{cm}^{-3}} \right) \text{ yr.} \quad (75)$$

- The **thermal timescale** is the time period for any phase of the ISM to change its energy per unit volume:

$$t_{\text{therm}} = \frac{nkT_{\text{kin}}}{\Lambda}. \quad (76)$$

- The **sound crossing time** is

$$t_{\text{sc}} \simeq \frac{N_H}{n_H c_s}, \quad (77)$$

where the sound speed is $c_s \simeq 1.4 \times 10^4 (T_{\text{kin}} / \text{K})^{1/2}$.

- The WIM cannot have a filling factor of 25% and still have a scale height of ~ 900 pc via stellar winds and SNe alone. They must remain ionized due to OB stars. But OB stars have scale height of 70 pc. This means that there *must* be an appreciable fraction ($\sim 15\%$) of photons escaping H II regions—from *density bounded* H II regions.

Dust

What are some observed phenomena of dust?

- extinction at short wavelength, e.g., dark clouds blocking the disk of the MW—caused by dust grains absorbing light
- reflection nebulae—dust grains scattering starlight into our line of sight (preserves stellar spectra)
- reddening of starlight due to differential extinction at different wavelengths
- polarization of starlight—elongated dust grains align with B fields
- infrared emission from dense clouds—dust re=emitting absorbed energy
- **depletion**—the absence of particular elements (especially Fe, Si, C, Ca, etc.) from the gas phase because they're absorbed onto dust grains

To be clear, some of these are degenerate. *Extinction* is actually the sum of *absorption* (a.k.a attenuation) and *scattering*. The old version of radiative transfer was given by

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

where α_ν is the lost specific intensity due to *absorption*. Now, we write

$$dI_\nu = -(\alpha_\nu + \sigma_\nu) I_\nu ds, \quad (78)$$

where σ_ν is the scattering coefficient. We explicitly keep track of scatterings because we need to conserve photon number!

Previously, we parametrized emission via

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

but now we note

$$dI_\nu = (j_\nu + \sigma_\nu J_\nu) ds. \quad (79)$$

This assumes that scattering is coherent (in frequency, e.g., preserves ν) and isotropic. Of course, the latter is tricky, because the photons change direction compared to before, but it simplifies our problem. Explicitly, this assumption is reflected in the use of the averaged radiation field is $J_\nu = (4\pi)^{-1} \int d\Omega I_\nu$.

Now, we *assume emission is thermal* so that $\alpha_\nu = \alpha_\nu B_\nu(T)$, and combine terms:

$$\begin{aligned} \frac{dI_\nu}{ds} &= -(\alpha_\nu + \sigma_\nu) I_\nu + \alpha_\nu B_\nu(T) + \sigma_\nu J_\nu \\ &= -\alpha_\nu (I_\nu - B_\nu) - \sigma_\nu (I_\nu - J_\nu) \end{aligned} \quad (80)$$

So we'll need to define a new version of the source function, S'_ν (versus the original S_ν), as

$$S'_\nu = \frac{\alpha_\nu B_\nu(T) + \sigma_\nu J_\nu}{\alpha_\nu + \sigma_\nu}, \quad (81)$$

and

$$\frac{dI_\nu}{ds} = -(\alpha_\nu + \sigma_\nu) (I_\nu - S'_\nu). \quad (82)$$

So now we can define **extinction optical depth**, $d\tau'_\nu$,

$$d\tau'_\nu = (\alpha_\nu + \sigma_\nu) ds \quad (83)$$

so that we get our new equation of radiative transfer

$$\frac{dI_\nu}{d\tau'_\nu} = -I_\nu + S'_\nu. \quad (84)$$

Observational perspective of dust

Let's now assume a foreground dust screen. We can model the extinction

For a great review of empirical relations of dust extinction, check out Calzetti (2001).

$$F_\lambda = F_\lambda^{\text{no dust}} e^{-\tau'_\lambda}. \quad (85)$$

Be careful with the F_λ and F_ν switches, and don't forget that the primed quantities are *not* derivatives. Our fractional transmission in *magnitudes* are

$$F_\lambda = F_\lambda^{\text{no dust}} 10^{-0.4A_\lambda},$$

$$A_\lambda = -2.5 \log \left(F_\lambda / F_\lambda^{\text{no dust}} \right).$$

And it turns out that (because $1 / \ln(2.512) \approx 1.086$),

$$\frac{A_\lambda}{\text{mag}} = 1.086 \tau'_\lambda.$$

In a more general case, we might have a distribution of sources mixed with a distribution of dusty clouds. But we'll still assume that the population of sources is uniformly mixed with dust within some slab of thickness L . The optical depth through the entire slab is

$$\tau'_\lambda = \int_0^L ds (\alpha_\nu + \sigma_\nu). \quad (86)$$

Then the observed flux density is

$$F_\lambda = F_\lambda^{\text{no dust}} \left(\frac{1 - \exp(-\tau'_\lambda)}{\tau'_\lambda} \right). \quad (87)$$

For example, if $\tau'_\lambda = 5$ in a *uniformly distributed dust population*, then the fraction of dust-free emission received is 20% (what's happening here). If $\tau'_\lambda = 5$ for a *complete foreground screen*, we'd received 0.7% of the emission. However, with the Calzetti law, we usually assume an *effective* foreground screen for simplicity and give us good fits to correct for the dust attenuation.

Our color excess is parametrized by $E(\lambda_1 - \lambda_2) = A_{\lambda_1} - A_{\lambda_2}$: the extra redness is due to the (effective) foreground dust screen of $\lambda_1 - \lambda_2$ color.

We often use $E(B - V)$, where the B, V filters are centered at $\sim 4400 \text{ \AA}$ and 5500 \AA respectively. The extinction curve (as we've seen before) is

$$k(\lambda) = \frac{A_\lambda}{E(B - V)}, \quad (88)$$

and the normalization given by $R_V \equiv k(V) = A_V / E(B - V)$. This is also called the **total-to-selective extinction ratio**, and it's just a standard. The dust extinction curves are shown in Draine Figures 21.1 and 21.2. Note the 2175 \AA "bump"—which of course is actually an observed trough in our spectra.

Finally, we can apply all this by using the observed Balmer decrement, $F_{H\alpha} / F_{H\beta}$, to calculate the dust law:

$$E(B - V) = \frac{1}{k(H\beta) - k(H\alpha)} \cdot 2.5 \log \left(\frac{F_{H\alpha} / F_{H\beta}}{2.86} \right)$$

Tuesday, 6 October

Dust extinction (continued)

The dust column [density] is measured to be higher from hydrogen recombination lines than from UV/optical continuum. This is because the OB stars that

ionize hydrogen don't live long enough to drift away from their birth clouds, as opposed to the general UV/optical continuum from potentially older stars. We model this with a simplified geometry of an "extra" foreground screen in front of H II regions via

$$E(B - V)_{\text{gas}} = \frac{E(B - V)}{0.44}.$$

This factor of 0.44 is accepted for the local universe but may not be applicable at higher redshift.

The integrated light from a galaxy can also be described by an "effective" screen geometry (when dust attenuation is accounted for), e.g., the Calzetti attenuation curve. Star forming regions in the LMC (such as 30 Doradus) do *not* mirror the MW extinction curve (e.g., the 2175 Å bump).

Also in the MW extinction curve, we see **diffuse interstellar bands** (DIBs; possibly caused by C-C stretching modes). We also see a feature at 3.4 μm due to C-H stretching, and wide 9.7 μm Si-O stretching and 18 μm O-Si-O bending mode lines. These features can be found in Draine Figures 23.2 and 23.3 (the MW extinction curve is shown in Draine Figure 21.1).

In general, there is a good correlation between color excess and total gas column, implying that the dust-to-gas ratio is nearly constant averaged over large scales. This allows us to write down

$$\begin{aligned} \frac{E(B - V)}{\text{mag}} &= \frac{N_H + 2N_{H_2}}{5.8 \times 10^{21} \text{ cm}^{-2}} \\ \frac{A_V}{\text{mag}} &= \frac{N_H + 2N_{H_2}}{1.9 \times 10^{21} \text{ cm}^{-2}}, \end{aligned}$$

where obviously there is the usual factor of 3.1 difference between A_V and $E(B - V)$.

If we look into a dusty region towards a central source, we see further at longer wavelengths. However, this will paradoxically cause us to see a *higher* $E(B - V)$, or dust column, because of the longer optical depth allowed at longer λ . So we need to measure different wavelength pairs to carefully account for this (e.g., we observe stars near the central super-massive black hole in the MW at NIR wavelengths).

Physical properties of dust grains (theoretical)

We define **efficiency factors**, Q , to be the ratio of the extinction cross section and the effective physical cross section of a dust grain. For dust scattering,

$$Q_{\text{sca}}(\lambda) = \frac{\sigma_{\text{sca}}(\lambda)}{\pi a^2},$$

where $a \equiv a_{\text{eff}} = (3V/4\pi)^{1/3}$. For dust absorption,

$$Q_{\text{abs}}(\lambda) = \frac{\sigma_{\text{abs}}(\lambda)}{\pi a^2}.$$

This absorption can give rise to thermal emission, so that the **thermal emission coefficient** is given by

$$j_\nu = n_{\text{dust}} Q_{\text{abs}}(\lambda) \pi a^2 B_\nu(T_{\text{dust}}). \quad (89)$$

The total extinction is $Q_{\text{ext}}(\lambda) = Q_{\text{sca}}(\lambda) + Q_{\text{abs}}(\lambda)$, and we can find our total extinction optical depth, including scattering (denoted with primes):

$$\tau'_\lambda = \int_0^L ds \int_{a_{\text{min}}}^{a_{\text{max}}} da n_{\text{dust}}(a) \sigma_{\text{ext}}(\lambda, a) \quad (90)$$

$$\simeq Q_{\text{ext}}(\lambda) \pi a^2 N_{\text{dust}}. \quad (91)$$

This result is made only by assuming that the extinction cross section, σ_{ext} is independent of position.

The **albedo** is calculated

$$\tilde{w}(\lambda) = \sigma_{\text{sca}} / \sigma_{\text{ext}} = Q_{\text{sca}} / Q_{\text{ext}} \leq 1.$$

In order to calculate cross sections, we need to account for

- the ratio of λ to grain size characterized by $x \equiv 2\pi a / \lambda$
- the grain shapes, because they're generally not spherical
- the grain composition, which determines electric and magnetic polarizabilities:

$$\alpha_E = \frac{p}{E_{\text{app}}} \quad (92)$$

$$\alpha_H = \frac{m}{H_{\text{app}}} \quad (93)$$

based on applied electric and magnetic fields. Recall that $H = B - 4\pi M$, where M is the magnetization within a grain. Furthermore, these depend on their permeabilities ϵ and μ , which are functions themselves of λ .

To properly solve this crazy last bit, we will need to define a **complex refractive index** $m = \sqrt{\epsilon}$ (and ϵ is now the complex dielectric function; see Draine §22.2). This leads to **Mie theory**, as derived by Mie (1908) and Debye (1909). We refer to Draine figures 22.1-22.3, which plot the efficiency factors against $2\pi|m - 1|a / \lambda$, which is the phase shift of a wave traveling a distance within grain.

Mie theory is hard stuff, but phenomenologically,

- for large grains, $Q_{\text{abs}} \rightarrow 1$, $Q_{\text{sca}} \rightarrow 1$, so $Q_{\text{ext}} \rightarrow 2$. Small angle diffractions actually cause this total extinction efficiency to approach *double* the expected absorption (Draine figure 22.3).¹³
- For intermediate grain sizes, scattering (and hence extinction) Q factors show series of ripples. These “beats” are caused by interference between the incident and forward-scattered waves (Draine figure 22.2).

¹³ Even small angles in trajectory along astrophysical path lengths will cause photons to miss their original targets, so this is important in space, but not in a lab.

- The peak efficiency is for grain sizes comparable to λ . This helps us estimate (crudely) the characteristic dust grain size. It turns out that much of the dust mass is due to dust with $a \approx 0.15 \mu\text{m}$.

The **MRN model** (Mathis, Rumpl, and Nordsieck 1977) is a good, albeit not cutting-edge, model for dust extinction. They proposed power-law distributions of Si and graphite (carbonaceous) grains from 50-2500 Å in size. Those distributions are given by

$$\begin{aligned}\frac{dn_{\text{sil}}}{da} &= 7.8 \times 10^{-26} \left(\frac{n_H}{\text{cm}^{-3}} \right) \left(\frac{a}{\text{cm}} \right)^{-3.5} \\ \frac{dn_{\text{gra}}}{da} &= 6.9 \times 10^{-26} \left(\frac{n_H}{\text{cm}^{-3}} \right) \left(\frac{a}{\text{cm}} \right)^{-3.5}.\end{aligned}$$

One result of this model is that it assumes two-thirds of all C to be depleted, $\sim 20\%$ of all O, and nearly all Mg and F. We do see strong magnesium lines in the NUV, so if this model is correct, then Mg is a terrible predictor of metallicity (and we must be careful using the others). Another implication is that $M_{\text{dust}}/M_H \simeq 0.01$, suggesting that most of the metals are locked up in dust grains.

The total dust mass (or volume) is then

$$M_{\text{dust}} \propto \int da a^3 \frac{dn}{da} \propto a_{\text{max}}^{0.5} - a_{\text{min}}^{0.5},$$

so that the mass (volume) is dominated by the largest grains. The surface area is

$$A_{\text{dust}} \propto \int da a^2 \frac{dn}{da} \propto a_{\text{max}}^{-0.5} - a_{\text{min}}^{-0.5},$$

so that the surface area is dominated by the smallest grains. Logically, we also conclude that the number density is dominated by the smallest dust grains.

Dust emission

We'll want to calculate the dust temperature to determine its emission. The volumetric heating rate from the ambient radiation field, $J_\lambda(\lambda)$ is

$$\Gamma_{\text{gr}} = 4\pi\sigma_d \int_0^\infty d\lambda Q_{\text{abs}}(\lambda) J_\lambda(\lambda).$$

The emission per volume is then

$$\Lambda_{\text{em}} = 4\pi\sigma_d \int_0^\infty d\lambda Q_{\text{abs}}(\lambda) B_\lambda(T_d).$$

We rewrite the optical depth

$$\tau_\lambda = \int dl n_d \pi a^2 Q_{\text{abs}}(\lambda), \quad (94)$$

where the integral resolves $\tau_\lambda \propto \lambda^\beta$. This β is the **emissivity index**, and $\beta \simeq 2$ for metals, $\beta \simeq 1$ for amorphous carbon grains. The observed intensity

will be $B_\lambda(T_d)$ modulated by $1 - e^{-\tau_\lambda}$, or

$$1 - e^{-\tau_\lambda} = 1 - \exp \left[- \left(\frac{\lambda_0}{\lambda} \right)^\beta \right] \simeq \left(\frac{\lambda_0}{\lambda} \right)^\beta,$$

which also defines the **emissivity**. Note that λ_0 is simply the wavelength at which the optical depth is 1.

So then, we can find the expected spectrum in B_λ or B_ν .¹⁴

$$B_\lambda(T_d)(\lambda_0/\lambda)^\beta \propto \lambda^{-5-\beta} \frac{1}{\exp(hc/\lambda k T_d) - 1}$$

$$B_\nu(T_d)(\nu/\nu_0)^\beta \propto \nu^{3+\beta} \frac{1}{\exp(h\nu/kT_d) - 1},$$

where now we see the steep *rise* in the spectrum ($3 + \beta$), which causes the well-known *positive k-correction*.

Nearby bright young stars, we can get T_d via

$$T_d \simeq 33.5 \left(\frac{a}{0.2 \mu\text{m}} \right)^{-0.2} \left(\frac{G_0}{10^4 \text{ Habing}} \right)^{0.2} \text{ K}$$

in terms of grain size and the ISRF, G_0 .

Measuring and modelling dust SEDs reveals

- diagnose whether an obscured source has dust luminosity powered by star formation (SF) or AGN (via T_d —warm or hot dust)
- for AGN, the shape of SEDs can be used to model dust geometry near the black hole (BH)
- limit to T_d : above $T_d \sim 1500$ K, “sublimation temperature” would destroy dust grains
- shape of SED of extragalactic FIR background shows contributions of AGN vs. SF
- this MRN model fails to explain 2-25 μm emission excess caused by transiently heated, small dust grains.

Friday, 9 October

Dust grains

Dust grains are charged and elongated. Why should dust be charged?

- Emission of photoelectrons – volumetric rate $R_{pe} \propto n_e$
- Accretion of (collisions with) electrons – volumetric rate $R_{ae} \propto n_e^2$
- Accretion of (collisions) with positive ions – volumetric rate $R_{ai} \propto n_e^2$

¹⁴ Recall that, despite all of our $\lambda \leftrightarrow \nu$ shenanigans,

$$\int d\lambda B_\lambda(T_d) = \int d\nu B_\nu(T_d) \frac{c}{\lambda^2}.$$

Let's assume for now 2 charge states, Z_d and $Z_d + 1$. Then the fractions $f(Z_d)$ and $f(Z_d + 1)$ are given by

$$f(Z_d) [R_{pe}(Z_d) + R_{ai}(Z_d)] = f(Z_d + 1) R_{ae}(Z_d + 1). \quad (95)$$

For a multi-charge state, we'd have to balance many terms, etc, similar to our multi-state atom.

In *dense clouds*, R_{pe} is negligible because of the dust shielding. Then the electrons have far higher cross-sections due to their higher velocities, so grains will often have negative charges. This *peak* charge in the charge distribution is like

$$Z_{\text{high density}}^* \simeq -1.5 \left(\frac{a}{1000} \right) \left(\frac{T_{\text{kin}}}{100 \text{ K}} \right).$$

In *diffuse clouds*, the photoelectric effect will dominate and our grains should have *positive charge*, with charge distribution peaked at

$$Z_{\text{low density}}^* \simeq +36 \left(\frac{a}{1000} \right) \left(\frac{G_0}{1 \text{ Habing}} \right) \left(\frac{T_{\text{kin}}}{10^4 \text{ K}} \right)^{1/2} \left(\frac{n_e}{0.1 \text{ cm}^{-3}} \right)^{-1}.$$

Some general subtleties: Coulomb repulsions/attractions are included, and the Coulomb barrier depends on the charge and size of dust grain. This explains why both of the above scale as $\propto a$ and not $\propto a^2$.

Dust grains are elongated, not spherical. In general, \vec{J} to align with the principal axis of the largest moment of inertia, I_1 *firstly*, and secondly, \vec{J} will align with any \vec{B} fields. This leads to polarization of starlight due to differential extinction, given by¹⁵

$$p(\lambda) = p_{\text{max}} \exp \left(-K [\ln(\lambda/\lambda_{\text{max}})]^2 \right) \quad (96)$$

Typically, $\lambda_{\text{max}} \simeq 5500 \text{ \AA}$, $K \simeq 1$, and we find

$$p_{\text{max}} \leq 0.03 \frac{A_{\lambda_{\text{max}}}}{\text{mag}}.$$

The polarization is less if \vec{B} is disordered or if it's aligned partially along the line of sight. Also, λ_{max} increases with gas density.¹⁶

Molecular energy levels: order of magnitude

There are three types of energy levels, and we can calculate order-of-magnitude estimates of their energy scales:

1. Electronic Given a molecular size a , use de Broglie's equation ($\lambda = h/p$) to get $p \sim \hbar/a$:

$$E_{\text{ele}} \sim \frac{p^2}{m} \sim \frac{\hbar^2}{m_e a^2}, \quad (97)$$

where typical $a \sim 10^{-8}$, so we find energy scales of a few eV. Transitions result in *optical/UV wavelength* photons.

¹⁵ This is called the Serkowski law, from Serkowski (1973).

¹⁶ Because our fiducial polarization peak wavelength is in the optical rather than in the UV, it suggests that large grains with higher absorption tend to either be more spherical, or in regions of disordered \vec{B} fields.

2. Vibration We will assume a diatomic molecule and treat it like a simple harmonic oscillator, where the potential is $\frac{1}{2}M\omega^2\zeta^2$, where ζ is the displacement from some usual separation, R_0 , and M is the nuclear mass and ω the vibrational frequency. Then the electron energy differences are given by

$$\frac{1}{2}M\omega^2a^2 \sim \frac{\hbar^2}{2m_e a^2} \quad (98)$$

, so if $E_{\text{vib}} \sim \hbar\omega$, and

$$E_{\text{vib}} \sim \left(\frac{m_e}{M}\right)^{1/2} \frac{\hbar^2}{m_e a^2}, \quad (99)$$

putting our energy scales $\sim 0.1 - 0.01$ eV and wavelengths in the infrared (MIR).

3. Rotation

Rotational transitions depend on J , with angular momentum $J\hbar$ and

$$E_{\text{rot}} \sim \frac{\hbar^2 J(J+1)}{2I}. \quad (100)$$

For a diatomic molecule, $I \simeq \frac{1}{2}Ma^2$ (so long as the two atom masses are comparable), and for small J ,

$$E_{\text{rot}} \sim \left(\frac{m_e}{M}\right) \frac{\hbar^2}{m_e a^2}. \quad (101)$$

So rotational energy transitions have $E_{\text{vib}} \sim 101^{-3}$ and λ is at radio wavelengths.

This allows us to use the Born-Oppenheimer approximation because $E_{\text{ele}} \gg E_{\text{vib}} \gg E_{\text{rot}}$.

Electronic bonds

We will consider H_2^+ , in which the energetically favorable configuration will look something like $p^+ \cdot e^- \cdot p^+$. We will use the **Linear Combination of Atomic Orbitals** (LCAO), so that the protons (labeled A and B) can be written in their ground states

$$\psi_A(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} \exp\left(-\frac{|\vec{r} - \vec{R}_A|}{a_0}\right), \quad (102)$$

etc. The Bohr radius is $a_0 = 0.529 \times 10^{-8}$ cm, and we know that normalized, $\int d^3\vec{r} |\psi_{A,B}(\vec{r})|^2 = 1$. We can construct molecular states, where $g \rightarrow \text{gerade}$, or even, and $u \rightarrow \text{ungerade}$, which is odd:

$$\chi_g(\vec{r}) = C_+ (\psi_A(\vec{r}) + \psi_B(\vec{r})) \quad \chi_u(\vec{r}) = C_- (\psi_A(\vec{r}) - \psi_B(\vec{r})). \quad (103)$$

Note that

$$C_{\pm}^{-2} = \int d^3\vec{r} |\psi_A(\vec{r}) \pm \psi_B(\vec{r})|^2 = 2 \pm 2 \left(1 + \frac{R}{a_0} + \frac{1}{3} \left(\frac{R}{a_0} \right)^2 \right) e^{-R/a_0}. \quad (104)$$

The Hamiltonian is

$$H = -\frac{\hbar^2}{2e} \nabla^2 - \frac{e^2}{\vec{r} - \vec{R}_A} - \frac{e^2}{\vec{r} - \vec{R}_B} - \frac{e^2}{\vec{R}_A - \vec{R}_B}, \quad (105)$$

with expectation value

$$\langle H_{\pm} \rangle = \int d^3\vec{r} \chi_{\pm}^*(\vec{r}) H \chi_{\pm}(\vec{r}) \quad (106)$$

and net binding energy $E_{\pm} - E_{1s}$ given by equation 20 given by the Lecture 20 equation sheet. A negative binding energy is energetically favorable.

Only the *even* state dips negative, resulting in a **bonding orbital** (dip is to ~ 2.8 eV at $R \simeq 1.0$ Å). The odd state remains at positive energy, so that it's an **anti-bonding orbital**.

Now, for neutral H_2 , we can build orbitals by combining χ_g orbitals (adding in Coulomb repulsion of two electrons and watching out for the Pauli exclusion principle). Therefore, we must keep the electron pair at opposite spins, which is the **singlet anti-symmetric state**:

$$\chi_{\text{singlet}} \propto (\psi_A(\vec{r}_1) + \psi_B(\vec{r}_1)) (\psi_A(\vec{r}_2) - \psi_B(\vec{r}_2)) \phi_{\text{singlet}}, \quad (107)$$

where we can ground cross terms given by “equations” (22) and (23) on the Lecture 20 equation sheet. Equation (22) describes an H^- combining with another proton, but that's unlikely; instead we want equation (23), which is the H_2^+ scenario. Then the *singlet* and *triplet* states are bound and unbound, respectively, and equations (24-25) on the equation sheet.

Tuesday, 13 October

Electronic states for homonuclear diatomic molecules

Quantum numbers of relevance include

- S , the total spin
- Λ , the orbital angular momentum of electrons *along internuclear axis*
- Ω , (where $\vec{\Omega} = \vec{S} + \vec{\Lambda}$) the total angular momentum of electrons *along internuclear axis*
- u/g , odd/even parity, defined only for homonuclear molecules
- $+/-$, reflection symmetry along a plane containing internuclear axis, defined only if $\Lambda = 0$

The eigenvalue $\lambda = |m_\ell|$, labeled as $\sigma, \pi, \delta, \phi$ for $\lambda = 0, 1, 2, 3$. Then the total $M_L = \sum_i m_{\ell,i}$, and $\Lambda \equiv |M_L|$.¹⁷ These Λ terms are labeled as $\Sigma, \Pi, \Delta, \Phi$ for $\Lambda = 0, 1, 2, 3$.

If we have spin-orbit coupling, then $\Omega = \Lambda + M_S$, where M_S is the projected spin along the internuclear axis. So the full *term* symbol looks like $^{2S+1}\Lambda_\Omega$.

For H_2 , the four lowest-energy electronic states are

- X $^1\Sigma_0^+$ $g+$: both electrons are in lowest σ_g $1s$ orbital and have opposite spins
- b $^3\Sigma_1^-$ $u+$: “repulsive” electrons in σ_g $1s$ and σ_u $1s$ orbitals with same spins
- B $^1\Sigma_0^+$ $u+$: electrons in σ_g $1s$ and σ_u $1s$ orbitals with opposite spin
- C $^1\Pi_1^-$ u : electrons are in σ_g $1s$ and π_u $2p$ orbitals with opposite spins

See Draine Figure 5.1 (margin) to see some of the electron potentials of the H_2 electronic configurations. Observationally, transitions linking X to B are **Lyman bands**, and transitions linking X to C are Werner bands—both sets of line transitions are in the rest-UV.

Rotational spectra for linear molecules

In cold, dense environments, we can assume that the molecules are in their lowest electronic and vibrational states, such that only the rotational excitations are allowed. We diagonalize the moment of inertia tensor to get the principle moments, I_x, I_y, I_z . Typically we see the following molecular shapes:¹⁸

- Linear $I_x = I_y, I_z \simeq 0$, e.g., CO
- Oblate (symmetric top) $I_x = I_y < I_z$, e.g., ammonia (NH_3).

Prolate also works, $I_x < I_y = I_z$, e.g., methyl acetylene (CH_3CCH)

- Spherical top $I_x = I_y = I_z$, e.g., methane (CH_4)
- Asymmetric top $I_x \neq I_y \neq I_z$, e.g., formaldehyde (H_2CO)

Classically, we can write the total energy and angular momentum:¹⁹

$$E = \frac{\vec{J}_x^2}{2I_x} + \frac{\vec{J}_y^2}{2I_y} + \frac{\vec{J}_z^2}{2I_z} \quad (108)$$

$$\vec{J}^2 = J_x^2 + J_y^2 + J_z^2 \quad (109)$$

The hamiltonian is simply $H_{\text{rot}} = J^2/2I$, and we find energy levels

$$E_{\text{rot}} = \frac{h^2 J(J+1)}{2\mu R^2}, \quad (110)$$

where sometimes we define a **rotation constant** $B = h^2/8\pi^2 I$, so that $E_{\text{rot}} = hBJ(J+1)$.

Selection rules for electronic dipole transitions:

¹⁷ But note that $\Lambda \neq \sigma_i |m_{\ell,i}|$.

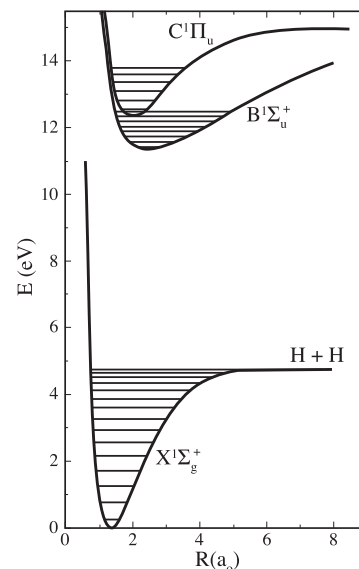


Figure 5: The energy spectrum of different H_2 molecular states (Draine Figure 5.1).

¹⁸ We call them *tops* because they tend to spin

¹⁹ Why do we use J as the single quantum number? Only because this tends to behave like a classical system, but it has nothing to do with the total spin-orbit angular momentum of an atom.

1. $d \neq 0$. It must have a permanent electric dipole moment, i.e., d cannot equal zero (this is impossible for homonuclear molecules)

2. $\Delta J = \pm 1$

Energy levels can easily be written in terms of frequency $\nu_{J+1 \rightarrow J} = 2B(J+1)$. The quadratically spaced energy levels have linearly spaced energy separations. This is exactly how the CO rotational ladder works.

Remember way back when we derived the A_{21} Einstein coefficient,

$$A_{21} = \frac{64\pi^4}{3hc^3} \nu_{21}^3 \left| \langle \psi_1 | \vec{d} | \psi_2 \rangle \right|^2.$$

For rotational transition $J+1 \rightarrow J$,

$$\left| \langle \psi_{J+1} | \vec{d} | \psi_J \rangle \right|^2 = d^2 \frac{J+1}{2J+3}.$$

All these have assumed a **rigid rotator**, which ignores centrifugal stretching (a $\sim 10\%$ effect).

The moment of inertia and dipole are affected by **isotopic composition** because the rotation coefficient $B \propto I^{-1} \propto \mu^{-1}$. So changing the mass leads to

$$\begin{aligned} \nu &\propto \frac{M_1 + M_2}{M_1 M_2} \\ \frac{d\nu}{dM_1} &\propto -\frac{1}{M_1^2} \\ \frac{\Delta\nu}{\nu} &= -\frac{\Delta M_1}{M_1} \frac{M_2}{M_1 + M_2} \end{aligned}$$

So, for the case of ^{13}CO versus ^{12}CO , we find a shift in frequency

$$\frac{\nu(^{13}\text{CO}) - \nu(^{12}\text{CO})}{\nu(^{12}\text{CO})} \simeq -\frac{1}{12} \frac{16}{28} \approx -0.048,$$

again assuming a rigid rotator.

Rotational spectra for more complex molecules

First, we'll examine the energy of a prolate molecule ($I_B = I_C > I_A$):

$$H_{\text{rot}} = \frac{J^2}{2I_B} + J_z^2 \left(\frac{1}{2I_A} - \frac{1}{2I_B} \right). \quad (111)$$

This form is nice because the eigenvalues of J^2 and J_z are known:

$$J^2 \psi = \hbar^2 J(J+1) \psi \quad (112)$$

$$J_z \psi = \hbar K \psi, \quad (113)$$

where $K \in [-J, -J+1, \dots, J]$. So, plugging back into the hamiltonian,

$$E = \frac{\hbar^2}{2I_B} J(J+1) + \left(\frac{\hbar^2}{2I_A} - \frac{\hbar^2}{2I_B} \right) K^2 \quad (114)$$

$$= hBJ(J+1) + h(A-B)K^2, \quad (115)$$

where B was the old rotational constant and A is a new one. The *oblate* top is exactly the same, except that the constant A is swapped out for another one.

So now we come upon these new selection rules for electric dipole transitions:

1. $\Delta J = 0, \pm 1$
2. $\Delta K = 0$

The K quantum number *cannot* change via dipole transitions, so, e.g., a molecule starting out with $J = 0$ will *always* have $K = 0$.

Draine Figure 5.4 shows the spectrum of NH_3 , which is an oblate spheroid. The nitrogen can tunnel through the three hydrogen molecules, causing a parity difference and a ~ 23 GHz transition.

Finally, nuclear wavefunctions can also break degeneracies in energy splittings. They are called *ortho*- when they are symmetric, and *para*- when anti-symmetric. For water (Draine Figure 5.5) or formaldehyde, we can observe the difference in line populations. The conversion between ortho and para takes \sim a million years.

Friday, 16 October

Vibrational spectra

For a diatomic molecule, we can use a *Morse potential*:

$$U(r) = U_0 + B_n \left[1 - e^{-\beta_n(r-r_0)} \right]^2. \quad (116)$$

These constants, B_n, β_n, r_0 are fit to each particular molecule (and n refers to the electronic state—impacting the separation of the atomic nuclei). Energy eigenvalues are given by

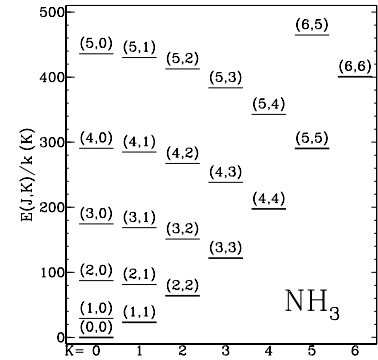
$$E_{nv} = \hbar\omega_{n0} \left(v + \frac{1}{2} \right) - \frac{\hbar\omega_{n0}^2}{4B_n} \left(v + \frac{1}{2} \right)^2, \quad (117)$$

where $\omega_{n0} = \beta_n \sqrt{2B_n/\mu}$ and μ the reduced mass. Note that we get a vibrational quantum number v , which is an integer within:²⁰

$$0 \leq v \leq \frac{\sqrt{2\mu B_n}}{\beta_n \hbar} - \frac{1}{2}.$$

A slight digression: Suppose we have a molecule with N atoms—then the number of degrees of freedom, \mathfrak{D} , is $3N$ (in configuration space). We “lose”

Figure 6: The energy spectrum of ammonia molecules (Draine Figure 5.5).



Observing formaldehyde, we can probe its formation history (by default, we'd see a ratio of 3:1 for ortho:para).

²⁰ The upper energy limit is $E = B_n$, which will basically lead to the unbinding of the molecule.

3 to center of mass translations, 2 or 3 or more to rotations (if linear, etc), leaving $\mathfrak{D} = 3N - 5$ vibrational degrees of freedom if linear (or $\mathfrak{D} = 3N - 6$ or fewer for more general molecules).

Selection rules for electric dipole vibration-rotation transitions of diatomic molecule:

1. $d \neq 0$ (from before—we require a permanent dipole)
2. $\Delta d|_{r=r_0} \neq 0$, i.e., the dipole moment must change during a change in vibrational state
3. $\Delta v = \pm 1$
4. $\Delta J = \pm 1$ if $\Lambda = 0$, or $\Delta J = \pm 1, 0$ if $\Lambda \neq 0$ (for rotations). For pairs of vibrational levels with $v_2 > v_1$, $\Delta J = J_{v_1} - J_{v_2}$

The notation for the $\Delta J = \{+1, 0, -1\}$ transitions are known as the {P, Q, R} branches. It is expressed as $v_2 - v_1 \{P, Q, R\} (J_{v_1})$. For example, for CO:

1-0 R(0) at 4.657 μm links $(v_2, J_{v_2}) = (1, 1)$ to $(v_1, J_{v_1}) = (0, 0)$

2-1 P(10) at 4.814 μm links $(v_2, J_{v_2}) = (2, 9)$ to $(v_1, J_{v_1}) = (1, 10)$

H_2 with $d = 0$ can undergo *electric quadrupole* transitions with selection rule

1. $\Delta J = \{+2, 0, -2\}$ called {O, Q, S} branches.
2. No rules on Δv

We observe strong transitions in the NIR, e.g., 1-0 S(1) at 2.12 μm and 2-1 S(1) at 2.25 μm . Can be seen in shocks or near young stars as UV photons excite molecules electronically, but as they hit the electronic ground state in some arbitrary ro-vibrational²¹ state, they cascade down the ro-vibrational ladder. In molecular clouds, the lack of collisional excitations prevent these H_2 transitions from occurring often.

²¹ or vibro-rotational

The statistical weights for (v, J) ro-vibrational levels of H_2 are

$$g_{(v,J)} = g_{o/p}(2J + 1),$$

where $g_{o/p} = 3$ for odd J (ortho- H_2) and $g_{o/p} = 1$ for even J (para- H_2).

Electronic-vibrational-rotational transitions

The full energy level is given by

$$E_{nvJ} = V_{n0} + \gamma_n \hbar^2 \Lambda^2 + \alpha_n \hbar J(J + 1) + \left(v + \frac{1}{2}\right) \hbar \omega_n. \quad (118)$$

We have the following selection rules:

1. $\Delta \Lambda = \{0, \pm 1\}$

2. $\Delta J = \{0, \pm 1\}$ but $J = 0 \rightarrow 0$ excluded, and $\Delta J = 0$ if $\Lambda = 0 \rightarrow 0$

3. $\Delta v = \{\text{any non-zero integer}\}$

See Rybicki & Lightman Figures 11.7 and 11.8 for PQR branches of these electro-vibro-rotational transitions.

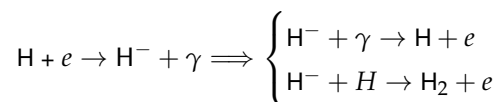
A decay of H_2 down to $v > 14$ (or $J > 5$ for $v = 14$) in the ground electronic state, the molecule will unbind. It's not fully intuitive, but the potential for a molecule will change depending on its electronic state.

Formation of H_2

There are gas-phase interactions and grain surface interactions to forming molecular hydrogen. First we'll examine the gas-phase interaction (more accurately than when we used the LCAO method).

Can H_2 form via radiative association (i.e., $\text{H} + \text{H} \rightarrow \text{H}_2 + \gamma$)? Nope, because there's only one photon emitted via a permanent dipole, but $d = 0$ for H_2 .

The best gas-phase formation channel involves



where the rates of reaction are given by equations (12-14) in the Lecture 14 equation sheet for the left side, and the upper/lower right side, respectively. Note that the upper part is dissociation of H^- and the lower part is the creation of H_2 .

If we set equal the rates of H^- formation and destruction equal, we can find the rate of H_2 formation

$$R_{\text{H}^-}(\text{H}_2) = 7.5 \times 10^{-23} \frac{T_{\text{gas}}}{100 \text{ K}} \frac{x}{1.4 \times 10^{-4}} n_{\text{H}}^3 \text{ cm}^{-3} \text{ s}^{-1}.$$

But we observe much more H_2 than this channel can explain. H can also be *adsorbed* onto a dust grain at two types of sites:

physisorption, where the atom is weakly bound by Van der Waals forces

chemisorption, where the atom is strongly bound by wavefunction overlap (like a chemical bond)

Schematically, the formation rate of $R_d(\text{H}_2)$ is

$$R_d(\text{H}_2) = \frac{1}{2} S(T_{\text{gas}}, T_d) \eta(T_d) n_d \sigma_d n_{\text{H}} v_{\text{H}}, \quad (119)$$

where S is the probability that a H atom with velocity v_{H} will stick to grain, $\eta(T_d)$ probability it will combine with another H atom and "evaporate" back to the gas phase.

We can simplify this by substituting $v_H \propto T_{\text{gas}}^{1/2}$, noting equation (19) from the equation sheet, finding that our dust-catalyzed formation rate of H_2 is:

$$R_d(\text{H}_2) \simeq 5 \times 10^{-17} \left(\frac{T_{\text{gas}}}{100 \text{ K}} \right)^{1/2} S(T_{\text{gas}}, T_d) \eta(T_d) n_H^2 \text{ cm}^{-3} \text{ s}^{-1}.$$

This formation mechanism dominates over the gas-phase creation of H_2 .

If we define

- $r_{\text{ads}} = n_H v_H \pi a^2 / N_s$ to be the rate of H adsorption per site, $N_s \simeq 2 \times 10^{15} \text{ cm}^{-2}$ along the grain surface
- θ_p, θ_c as the occupation fractions of physisorption, chemisorption sites
- $k_{\text{pc}}, k_{\text{ev}}$ as the rates at which H atoms on the surface jump from physisorption to chemisorption sites, and the rate at which atoms/molecules evaporate (or desorb)
- μ as the fraction of formed molecules that do not evaporate

Then the equilibrium equations (or balancing the two *at physisorption sites*) is

$$r_{\text{ads}} [1 - \theta_p(\text{H}_2)] = k_{\text{pc}} \theta_p(\text{H}) + k_{\text{ev}} \theta_p(\text{H}). \quad (120)$$

Again, we've balanced absorption onto grains versus migration and evaporation out at physisorption sites.

At chemisorption sites equilibrium requires

$$k_{\text{pc}} \theta_p(\text{H}) [1 - 2\theta_c(\text{H})] = 0, \quad (121)$$

so that $\theta_c(\text{H}) = \frac{1}{2}$ at equilibrium. And finally, for the formation of H_2 , the equilibrium is

$$\mu k_{\text{pc}} \theta_p(\text{H}) \theta_c(\text{H}) = k_{\text{ev}}(\text{H}) \theta_p(\text{H}_2). \quad (122)$$

There are three regimes:

1. low temperature ($T_d < 5 \text{ K}$): H atoms adsorb but don't migrate so physisorption sites fill without making H_2
2. high temperature ($T_d > 30 \text{ K}$): H atoms evaporate from physisorption sites before jumping to form H_2
3. just right ($5 \text{ K} \leq T_d \leq 20 \text{ K}$): nearly 100% efficiency of H atoms adsorbing, migrating, and partnering to form H_2 before evaporation

Tuesday, 20 October

Formation of non- H_2 molecules

Gas-phase chemistry is more feasible for the formation of molecules when the electric dipole $d \neq 0$. There are several kinds of important reactions.²²

²² Grain-phase chemistries are slow because large molecules can be too massive to jump across sites on the surface of a dust grain. Small molecules can form via these mechanisms, but large molecules must be built up slowly.

1. *Photodissociation*: e.g., $\text{CO} + \gamma \rightarrow \text{C} + \text{O}$. Note that, in H II regions, high energy UV photons can come from H_2 excitations due to passing cosmic rays
2. *Neutral-neutral reactions*: e.g., $\text{C} + \text{OH} \rightarrow \text{CO} + \text{H}$
3. *Ion-molecule reactions*: e.g., $\text{H}_3^+ + \text{CO} \rightarrow \text{HCO}^+ + \text{H}_2$
4. *Charge transfer*: e.g., $\text{O} + \text{H}^+ \rightarrow \text{O}^+ + \text{H}$ (key step in gas-phase oxygen chemistry)
5. *Radiative association*: e.g., $\text{C}^+ + \text{H}_2 \rightarrow \text{CH}_2 + \gamma$
6. *Dissociative electron recombinations*: e.g., $\text{OH}^+ + \text{e}^- \rightarrow \text{O} + \text{H}$

Molecular clouds

The CO (1-0) line traces molecular hydrogen despite being optically thick. The evidence is

1. Correlation between $I_{\text{CO}(1-0)}$, in units of K km s^{-1} , e.g., brightness temperature integrated over velocity, and estimates of N_{H_2} , from (optically thin) molecular absorption towards background stars proportional to $A_V - A_V(N_{\text{H I}})$
2. Correlation between $I_{\text{CO}(1-0)}$ and model for N_{H_2} based on maps of H I, dust emission proportional to $I_{\text{dust}} - I_{\text{dust}}(N_{\text{H I}})$
3. Correlation between extinction to background stars and optically thin emission like ^{13}CO (1-0)
4. Global correlation among CO (1-0), H I, and diffuse gamma ray emission proportional to $F_\gamma - F_\gamma(N_{\text{H I}})$

Taken together, the mean MW value of $X_{\text{CO}} = N_{\text{H}_2} / I_{\text{CO}}$ is given by

$$X_{\text{CO}} \simeq 2 \times 10^{20} \text{ cm} \left(\text{K km s}^{-1} \right)^{-1}.$$

If we examine diffuse clouds ($n_{\text{H}_2} \sim 50 \text{ cm}^{-3}$, CNM) versus dense clouds ($n_{\text{H}_2} \geq 200 \text{ cm}^{-3}$), we compare *abundance ratios*:

$$\text{CO}/\text{H}_2 \simeq 8 \times 10^{-5}$$

$$^{13}\text{CO}/\text{H}_2 \simeq 1.2 \times 10^{-6}, \text{ i.e., } ^{13}\text{CO}/\text{CO} \sim 1/70$$

$$\text{C}^{18}\text{O}/\text{H}_2 \simeq 1.7 \times 10^{-7}, \text{ e.g., } \text{C}^{18}\text{O}/\text{CO} \sim 1/500$$

We know that *emission line ratios* of these CO species give us closer to $^{13}\text{CO}/\text{CO} \sim 1/20$, signifying that the CO (1-0) line is optically thick. Remember that

$$I_\nu(\tau_\nu) \rightarrow B_\nu(T) (1 - e^{-\tau_\nu}), \quad (123)$$

with $\tau_\nu \propto N_\nu$. Then if both CO, ^{13}CO $J = 1$ states have low N , then

$$\frac{I_{\text{CO}(1-0)}}{I_{^{13}\text{CO}(1-0)}} \simeq \frac{\tau_{\nu,\text{CO}(1-0)}}{\tau_{\nu,^{13}\text{CO}(1-0)}} \simeq \frac{N_{\text{CO}(J=1)}}{N_{^{13}\text{CO}(J=1)}} \simeq 70.$$

So how does this all work out? Different isotopes of CO trace different depths into a cloud, which may also change the temperature of the probed region! The reason our conversion factor works is because we assume *virialized clouds*.

Molecular gas (in the MW, at least) are in gravitationally bound, discrete clouds. We observe a cloud with brightness temperature, T_{CO} , with velocity width, Δv_{CO} , and radius, r_{CO} . So our integrated intensity is $I_{\text{CO}} \simeq T_{\text{CO}} \Delta v_{\text{CO}}$. Integrate this over the projected surface area of the cloud to get “luminosity” L'_{CO} as in

$$L'_{\text{CO}} = \pi r_{\text{CO}}^2 T_{\text{CO}} \Delta v_{\text{CO}}, \quad (124)$$

which has units of $\text{K km s}^{-1} \text{pc}^2$. So now, assuming that our cloud is spherical and virialized, then

$$\begin{aligned} \Delta v_{\text{CO}} &= \sqrt{GM_{\text{gas}}/r_{\text{CO}}} \\ L'_{\text{CO}} &= \pi r_{\text{CO}}^2 T_{\text{CO}} \left(\frac{GM_{\text{gas}}}{r_{\text{CO}}} \right)^{1/2} \\ &= \pi r_{\text{CO}}^2 T_{\text{CO}} M_{\text{gas}} \left(\frac{G}{r_{\text{CO}} M_{\text{gas}}} \right)^{1/2} \end{aligned} \quad (125)$$

$$= T_{\text{CO}} M_{\text{gas}} \left(\frac{3\pi G}{4 \times 1.35 \times 2m_{\text{H}} n_{\text{H}_2}} \right). \quad (126)$$

In the last step, we've converted $M_{\text{gas}} = \rho \frac{4}{3} \pi r_{\text{CO}}^3$, and $\rho = 1.35 \times 2m_{\text{H}} \times n_{\text{H}_2}$, where the factor of 1.35 corrects for He.

So now we can implicitly define a *Galactic conversion factor* to get $M_{\text{gas}} = \alpha_{\text{CO}} L'_{\text{CO}}$:

$$\alpha_{\text{CO}} = \left(\frac{1.35 \times 8m_{\text{H}}}{3\pi G} \right)^{1/2} \frac{n_{\text{H}_2}^{1/2}}{T_{\text{CO}}}. \quad (127)$$

The virial theorem tells us that $M \propto (\text{projected size}) \times T \times \Delta v$, e.g., observable L'_{CO} IFF it is (roughly) spherical and $n_{\text{H}_2}^{1/2}/T_{\text{CO}}$ is nearly constant, somehow.

Now let's look at some (more) caveats. CO might *not* be a good tracer of molecular gas if

1. The metallicity of the gas is too low for CO to be present (e.g., in galaxies at early stages of formation). The galactic conversion factor, α_{CO} will underestimate M_{gas}
2. Gas is not confined to gravitationally bound clouds (e.g., in turbulent, high-pressure ISM of a galaxy merger). Then the linewidth Δv will misrepresent the size of the molecular clouds and cause α_{CO} to overestimate M_{gas}

Properties of molecular clouds

For a typical Giant Molecular Cloud (GMC),

- $n_{\text{H}_2} \sim 200 \text{ cm}^{-3}$, H_2 is the dominant collider species
- $r_{\text{CO}} \sim 20 \text{ pc}$
- $M_{\text{gas}} \sim 5 \times 10^5 M_{\odot}$ (can be measured from size and linewidth, since $M_{\text{gas}} \propto r_{\text{CO}}^2 \Delta v_{\text{CO}}$, but only if the linewidth is resolved)
- $T \sim 10 \text{ K}$
- $\Delta v \sim 5 - 10 \text{ km s}^{-1}$, however this significantly exceeds the sound speed in $T = 10 \text{ K}$ gas.²³ This means that the gas motions are supersonic; it must also be turbulent since so much MW gas has not formed stars.

$$^{23} c_s = (\gamma kT / \mu m_H)^{1/2} \simeq 0.2 \text{ km s}^{-1}$$

Molecular gas in the MW (as traced by CO) has:

- a vertical distribution peaked at mid-plane with a scale height of 75 pc
- a radial distribution rising and peaking at $\sim 300 \text{ pc}$ from the Galactic Center (GC), void out to 3.5 kpc, and a ring from 5-8 kpc.
- an azimuthal distribution that predominantly peaks along spiral arms
- a mean molecular gas surface density is $1 M_{\odot} \text{ pc}^{-2}$.

If we balance the rates of atomic \leftrightarrow molecular hydrogen gas, we can find timescales

$$\frac{M_{\text{H}_2}}{\tau_{\text{H}_2}} \simeq \frac{M_{\text{H I} + \text{H II}}}{\tau_{\text{H I} + \text{H II}}},$$

such that $\tau_{\text{H}_2} \sim 10 - 100 \text{ Myr}$.

In the outskirts of GMCs, we have CR heating dominant, and in the inside, photodissociative heating dominates. But CO rotational line cooling dominates the cooling all across the cloud. Equilibrium temperatures are 5 – 10 K in densest cores, and 30 – 40 K in outskirts or in star-forming cores.

Friday, 23 October

The molecular ISM

Once a molecular cloud begins to form stars, we find that there are three additional types of the “molecular ISM”

1. Dense cores: $n_{\text{H}_2} \gg 200 \text{ cm}^{-3}$. N_2H^+ is a popular choice for observing this phase (because it's not depleted onto dust grains).
2. Hot cores: After the protostar begins to shine, the dust is heated to $T \sim 100 \text{ K}$, and often we have $n_{\text{H}_2} > 10^6 \text{ cm}^{-3}$.

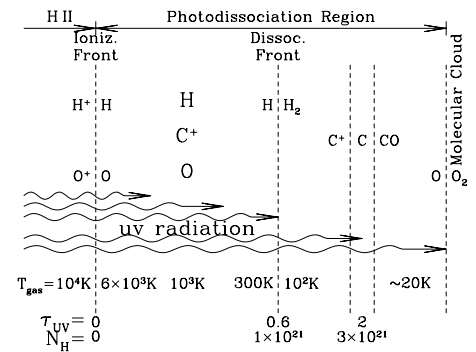
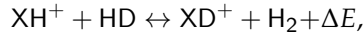


Figure 7: A schematic of a photon-dominated region (Draine Figure 31.2).

For low-temperature gas, charge transfer reactions take place:



yielding abundances

$$n_{\text{XH}^+} \cdot n_{\text{HD}} = n_{\text{XD}^+} \cdot n_{\text{H}_2} \cdot \exp(-\Delta E/kT).$$

The righthand side of the reaction tends to dominate as *deuteration* takes place and creates slightly more stable molecules. So we can measure timescales on the order of thousands of years this ways.

3. Photon-dominated regions (PDRs): UV photons from OB stars dominate chemistry and energy balance cause **fluorescent excitation**²⁴ of H₂ vibrational levels.

Cooling is mostly de to fine structure lines [C II] 158 μm , [O I] 63 μm (the latter of which is strong due to high density).

²⁴ H₂ fluorescence is due to UV photons exciting the electronic state, which then de-excites into a random vibro-rotational mode.

Magnetic fields

There are four approaches to measuring magnetic fields:

First is **synchrotron emission** - if we assume a power-law energy distribution with spectral index p ($N dE \propto E^{-p} dE$), we can measure the B field strength from synchrotron intensity, B field geometry from its linear polarization

$$\Pi = \frac{p+1}{p+7/3}. \quad (128)$$

For example, a supernova remnant with $p = 2$ gives $\Pi \sim 0.7$. But the measured polarization might be far lower, which tells us about how de-coherent the B field direction may be (i.e., geometry).

Secondly, we can use **Faraday rotation**. Let's assume a plane wave propagating parallel to a fixed magnetic field, B_0 , which give us circularly polarized waves with cyclotron frequency $\omega_B = eB_0/m_e c$. The electrons move with different velocities for different polarizations (right vs. left), as given by their dielectric constants:

$$\epsilon_{R,L} = 1 - \frac{\omega_p^2}{\omega(\omega \pm \omega_B)}. \quad (129)$$

$\omega_p = 4\pi n_e e^2 / m_e$ is the plasma frequency, and essentially indicates the electron mobility. Therefore the phase angle, ϕ , will rotate different amounts for R, L polarized emission.

$$\phi_{R,L} = \frac{1}{c} \int_0^L ds \omega \sqrt{\epsilon_{R,L}}. \quad (130)$$

So, for a background, linearly polarized source, the phase angle will rotate by an angle

$$\begin{aligned} \Delta\phi &= \frac{1}{2} \int_0^L ds (k_R - k_L) \\ &\approx \frac{2\pi e^3}{m_e^2 c^2 \omega^2} \int_0^L ds n B_0. \end{aligned} \quad (131)$$

where the last step used a first-order Taylor expansion (because $\omega_p \ll \omega$; $\omega_p \sim \text{kHz}$ and observed $\omega \sim \text{GHz}$). This latter integral is also called the **rotation measure**, or RM.

Keep in mind the following:

- The magnitude of $\Delta\phi \propto \omega^{-2} \propto \lambda^2$. Observationally, we usually use ≥ 3 wavelengths to break $\pm 2\pi$ degeneracies (two to measure $\Delta\phi$, and a third to break degeneracies).
- We've assumed B_0 to be parallel to the propagation direction (so substitute $B_0 \rightarrow B_{||}$). The field can double back on itself, so Faraday rotation underestimates $B_{||}$ strength due to cancellations—particularly on small scales.
- We only measure $\int n_e B_{||} ds$, so we can't break the degeneracy between $B_{||}$ and n_e from Faraday rotation. To get around this degeneracy, we must measure Faraday rotation towards pulsars, where we can *also* get the **dispersion measure** (DM).

So how do we calculate the dispersion measure? Ignoring the B field for a moment, our dielectric constant is given by $\epsilon = 1 - (\omega_p/\omega)^2$, leading to the dispersion relation

$$c^2 k^2 = \omega^2 - \omega_p^2.$$

For frequencies below the plasma frequency, the k wavenumbers are imaginary and those waves are damped exponentially. But they *do* manifest in the group velocity,

$$v_g = \frac{\partial \omega}{\partial k} = c \left(1 - \left(\frac{\omega_p}{\omega} \right)^2 \right)^{1/2}.$$

The arrival of a given pulse, from a pulsar at distance L , is $t_p = \int_0^L ds v_g^{-1}$, where we can Taylor expand (again because $\omega_p \ll \omega$) to estimate

$$t_p \approx \frac{L}{c} + \frac{1}{2c\omega^2} \int_0^L ds \omega_p^2.$$

Then taking the frequency derivative,

$$\frac{\partial t_p}{\partial \omega} = -\frac{4\pi e^2}{m_e c \omega^3} \int_0^L ds n_e, \quad (132)$$

and the dispersion measure, $\text{DM} = \int_0^L ds n_e$. We should be able to plot the pulses from pulsars at different frequencies and find a simple ω^{-3} relation. Then, using equations (131) and (132), we can make an average line-of-sight B field

$$\langle B_{||} \rangle = \frac{\text{RM}}{\text{DM}} = \frac{\int_0^L ds n_e B_{||}}{\int_0^L ds n_e}. \quad (133)$$

The rotation measure is formally defined:

$$\frac{\text{RM}}{\text{rad m}^{-2}} = 8.1 \times 10^5 \int_0^L \frac{ds}{\text{pc}} \left(\frac{B_{||}}{\text{Gauss}} \right) \left(\frac{n_e}{\text{cm}^{-3}} \right).$$

Thirdly, we can use **polarization**. Let's go back and review the **Stoke's parameters**:

$$I = E_1^2 + E_2^2 \quad (134)$$

$$Q = E_1^2 - E_2^2 \quad (135)$$

$$U = 2E_1E_2 \cos(\phi_1 - \phi_2) \quad (136)$$

$$V = 2E_1E_2 \sin(\phi_1 - \phi_2) \quad (137)$$

These apply for *single monochromatic waves* (and $I^2 = Q^2 + U^2 + V^2$).

If the incident waves are *quasi-monochromatic waves*, where the amplitudes/phases can drift slowly, then the time-averaged versions of the Stoke's parameters above are valid, and $I^2 \geq Q^2 + U^2 + V^2$ and we define the polarization

$$\Pi = \frac{\sqrt{Q^2 + U^2 + V^2}}{I}. \quad (138)$$

So we write the **Stoke's vector** as a sum of unpolarized and polarized components:

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I - \sqrt{Q^2 + U^2 + V^2} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \sqrt{Q^2 + U^2 + V^2} \\ Q \\ U \\ V \end{pmatrix}. \quad (139)$$

Observationally, dust causes

- Polarized extinction of background stars (comes from everywhere, low value)
- Polarized emission (dominated by *large* grains in *dense* clouds)

Draine Figure 21.3 shows a map of the Galactic dust polarization. But to actually estimate B , we can follow the method of Chandrasekhar & Fermi (1953): assume that the gas velocity dispersion is driven at the **Alfvén speed**, $v_A = B / \sqrt{4\pi\rho}$. It's the speed at which a transverse wave propagates through a magnetic field with oscillations $y = a \cos k(x - v_A t)$. Then, by differentiating the oscillation displacements with distance and time, we find

$$y' = \frac{dy}{dx} = -ak \sin k(x - v_A t)$$

$$\dot{y} = \frac{dy}{dt} = -akv_A \sin k(x - v_A t),$$

so that

$$v_A^2 y'^2 = \dot{y}^2$$

$$v_A^2 \sigma^2 = \frac{1}{3} v^2,$$

where σ is the rms polarization angle and v is the rms velocity of the gas. Thus, we can finally solve (assuming that the turbulent velocity is driven at the Alfvén speed):

$$B = \frac{v}{\sigma} \left(\frac{4\pi\rho}{3} \right)^{1/2}. \quad (140)$$

Friday, 30 October

Magnetic fields: the Zeeman effect

For a neutral atom, the ground state of $\vec{F} = \vec{J} + \vec{I}$, where F is the total magnetic moment, and I is the nuclear spin. In H I, the ground state has $F = J + 1 = 1$ (0), with corresponding $2J + 1 = 3$ (1) degeneracy. In an external magnetic field, $U \propto -\vec{F} \cdot \vec{B}$, and so for $F = 1$, we see a very small splitting in the 21-cm line, due to the magnetic field breaking the degeneracy.

If the line of sight is parallel to \vec{B} , then the unsplit (middle energy) line will be missing, and split lines will be completely polarized.

If the line of sight is perpendicular, then all three lines are observable, and split lines will be linearly polarized perpendicular \vec{B} .

For a general line of sight, the split lines will have elliptical polarization.

The splitting strength is proportional to $|\vec{B}|$; for H I the two split lines are separated by a frequency

$$\Delta\nu_B = \frac{eB_{||}}{2\pi m_e c} = 2.8 \times 10^6 \left(\frac{B}{\text{Gauss}} \right) \text{ Hz}.$$

For general molecules and atoms, the split levels of identical J is given by

$$\begin{aligned} U &= \frac{\mu_B}{\hbar} (g_L \vec{L} + g_S \vec{S}) \cdot \vec{B} \\ &\approx \frac{\mu_B}{\hbar} (\vec{J} + \vec{S}) \cdot \vec{B}, \end{aligned} \quad (141)$$

where we have assumed $g_L = 1$, $g_S = 2$. In terms of the Landé g factor,

$$g(L, S, J) = \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}, \quad (142)$$

and the J value, we can find the average magnetic energy

$$\langle U \rangle = \frac{1}{2} \left(\frac{e\hbar B_{||}}{m_e c} \right) g M_J. \quad (143)$$

So the frequency of newly split transitions are

$$\omega_{12} = \frac{1}{2} \left(\frac{e\hbar B_{||}}{m_e c} \right) (g_1 M_{J_1} - g_2 M_{J_2}). \quad (144)$$

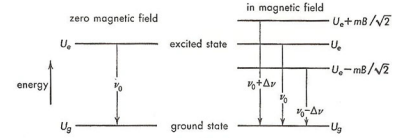


Figure 8: The Zeeman Effect in neutral hydrogen. The higher frequency line emission will be circularly polarized, the original frequency line emission will be linearly polarized parallel to \vec{B} , and the lower frequency line emission will be oppositely circularly polarized.

Observationally, we can use the *polarized* spectra, and take the **difference spectrum** of the right and left circularly polarized emission to quantify the field strength.²⁵

The median field strength as found in difference absorption spectra via Zeeman splitting of the 21-cm line is $B \sim 6 \mu\text{G}$, implying that the magnetic pressure is several times greater than that of the gas pressure!

Star formation: observational constraints

Stars form from dense molecular gas, which is traced by the correlation between HCN (1-0) line and FIR continuum.

Stars form in lots of binaries ($\sim 2/3$ of stars are in binaries).

Stars tend to form in embedded clusters.

Stars do *not* form via free-fall collapse of gas.²⁶

Stars do not form with a Salpeter IMF down to $0.1 M_\odot$. A better modern IMF, $\Phi(m)$ is given by Kroupa et al. (1993) or Chabrier et al. (2003).

The mass function of dense cores has a power-law index ($\propto m^{-2.5}$) similar to the that of the IMF for CO-emitting clouds ($\propto m^{-1.7}$).

The efficiency of star formation (SF) is 1 – 3% of GMC mass into stars, and 10 – 20% of dense gas mass into stars.

Protostars grow from the difference between accretion (inflow) and outflow. We can trace the outflow via SiO emission.

The observational classification of young stars are

Class 0 have more than half of their mass in infalling envelopes, but are too heavily obscured to have optical counterparts

Class I have less than half of their mass infalling, have strong evidence of outflows, and have IR (and sometimes optical) counterparts

Class II have dissipated envelopes, but still have circumstellar disks; are also called **classical T Tauri stars**

Class III have dispersed the inner edge of protostellar disks (due to radiation); are also known as **weak-lined T Tauri stars**

To understand the star formation process better (since it's not as simple as computing the free fall of gas), we'll delve into fluid mechanics

Basic equations of fluid mechanics

We will begin by deriving the equations for conservation of mass, momentum, and energy. First ups is *conservation of mass*. We'll consider a small volume

²⁵ Empirically, we see that $B \propto n_H^{1/2}$ for $n_H \geq 10 \text{ cm}^{-3}$. It also turns out that the Alfvén speed is $v_A \propto B/\sqrt{\rho} \sim 1.85 \text{ km s}^{-1}$, and that this velocity is close to the observed linewidths. Therefore, the motions in molecular clouds seem to be manifest from large-scale MHD waves.

²⁶ The free-fall time scale is

$$t_{\text{ff}} = \left(\frac{3\pi}{32G\rho} \right)^{1/2}. \quad (145)$$

In the Milky Way, $t_{\text{ff}} \sim 4.3 \times 10^7 \sqrt{(n_H/\text{cm}^{-3})} \text{ yr}$.

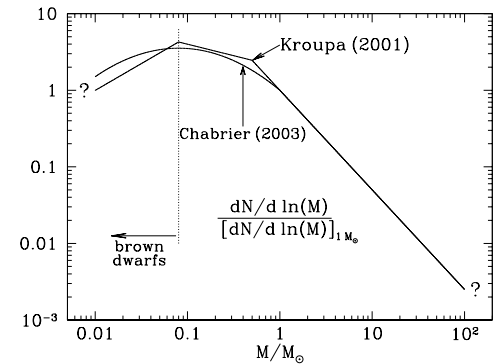


Figure 9: A few forms of the initial mass function (Draine Figure 42.1).

\mathcal{V} bounded by a surface \mathcal{S} , and a change of density within \mathcal{V} as a function of time $\frac{\partial \rho}{\partial t}(\vec{x}, t)$. The mass flux is $\rho \vec{v}$ for a velocity field $\vec{v}(\vec{x}, t)$.

$$\begin{aligned}\frac{\partial}{\partial t} \int_{\mathcal{V}} dV \rho + \int_{\mathcal{S}} d\Sigma \rho \vec{v} \cdot \hat{n} &= 0 \\ \int_{\mathcal{V}} dV \left(\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) \right) &= 0 \\ \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) &= 0.\end{aligned}\quad (146)$$

This is also called the *continuity equation*. We can also write this as a **convective** (or **Lagrangian**) **derivative**:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \vec{v} \cdot \vec{\nabla} \rho = -\rho \vec{\nabla} \cdot \vec{v}, \quad (147)$$

where $\vec{v} \cdot \vec{\nabla} \rho$.²⁷

²⁷ The Lagrangian derivative is evaluated by following a fluid element, whereas the Eulerian derivative is evaluated at a fixed location.

Tuesday, 3 November

Basic equations of fluid mechanics (continued)

The momentum flux crossing a differential surface element, $d\Sigma$ is $(\rho \vec{v} \cdot \hat{n} d\Sigma) \vec{v}$. Then, the equation of momentum conservation can be found (similarly to last time):

$$\begin{aligned}\frac{\partial}{\partial t} \int_{\mathcal{V}} dV \rho \vec{v} + \int_{\mathcal{S}} d\Sigma (\hat{n} \cdot \rho \vec{v}) \vec{v} &= 0 \\ \frac{\partial}{\partial t} \int_{\mathcal{V}} [dV \rho \vec{v} + \vec{\nabla} \cdot (\rho \vec{v} \otimes \vec{v})] &= 0 \\ \frac{\partial(\rho \vec{v})}{\partial t} + \vec{\nabla} \cdot \mathbf{T}_{\text{mo}} &= 0,\end{aligned}\quad (148)$$

where $\mathbf{T}_{\text{mo}} = \rho \vec{v} \otimes \vec{v}$ is the *momentum flux in tensor form*.

However, now we can add surface stress via a term $\mathbf{T}_{\text{surf}}(i, j)$, which is the force per area in direction i across surface area normal to direction j . Using Gauss's law again,

$$\int_{\mathcal{S}} d\Sigma_j (\hat{n} \cdot \mathbf{T}_{\text{surf}}(i, j)) = \int_{\mathcal{V}} dV (\vec{\nabla}_j \cdot \mathbf{T}_{\text{surf}}(i, j)). \quad (149)$$

So then, our pressure term can be found

$$-\vec{\nabla} P = \vec{\nabla} \cdot \mathbf{T}_{\text{surf}}(i, j),$$

and we can additionally include a force per volume $-\rho \vec{\nabla} \Phi$,²⁸ and end up with the full equation of momentum conservation

$$\frac{\partial(\rho \vec{v})}{\partial t} + \vec{\nabla} \cdot \mathbf{T}_{\text{mo}} = -\vec{\nabla} P - \rho \vec{\nabla} \Phi. \quad (150)$$

Were we to set the left-hand side of equation (150) to zero, then we have hydrostatic equilibrium (i.e., pressure is balanced by gravity), and the fluid

²⁸ This might be from a potential Φ —which might be gravity, for example. We're ignoring things like viscosity, or magnetic fields.

remains motionless. Now we add a body force (i.e., a magnetic field, \vec{B}) in a nonzero charged particle density (i.e., current, \vec{j}),²⁹ and find in terms of the metric, \mathbf{g} , the **Navier-Stokes equation** for zero viscosity:

²⁹ Recall that $\vec{\nabla} \cdot \vec{B} = 0$ and $\vec{j} = \frac{1}{4\pi} \vec{\nabla} \times \vec{B}$.

$$\frac{\partial(\rho\vec{v})}{\partial t} + \vec{\nabla} \cdot (\mathbf{T}_{\text{mo}} + \mathbf{T}_{\text{mag}} + \mathbf{g}P) = -\rho\vec{\nabla}\Phi. \quad (151)$$

The magnetic tensor is simply defined

$$\mathbf{T}_{\text{mag}} = \frac{B^2}{8\pi} \mathbf{g} - \frac{\vec{B} \otimes \vec{B}}{4\pi}.$$

For *conservation of energy density*, we multiply equation (146) by $v^2/2$ and take the dot product of equation (150). Additionally, we define the internal energy u (e.g., heat), the entropy s , and enthalpy $h = (u + P)/\rho$:

$$\frac{\partial}{\partial t} \left[\rho \left(\frac{1}{2} v^2 + u + \Phi \right) + \frac{B^2}{8\pi} \right] + \vec{\nabla} \cdot \left[\rho \vec{v} \left(\frac{1}{2} v^2 + h + \Phi \right) \right] + \frac{1}{4\pi} \vec{E} \times \vec{B} = \rho T \frac{ds}{dt}. \quad (152)$$

The Jeans instability

Firstly, let's assume $\vec{B} = 0$, and we can treat the density and velocities as *perturbations* to their equilibrium quantities: $\rho = \rho_0 + \rho_1$ and $\vec{v} = \vec{v}_0 + \vec{v}_1$. Using the continuity equation (equations 14-15 from the Lecture 18 equation sheet), we take the difference between the total and the equilibrium quantities of $\partial\rho/\partial t$ in order to find equation (16). Equation (17) is found via perturbing the potential, $\Phi = \Phi_0 + \Phi_1$ and substituted $dP \simeq c_s^2 d\rho$. Using Poisson's law ($\nabla^2 \Phi_1 = 4\pi G \rho_1$), and assuming *constant* ρ_0 and zero initial velocity ($v_0 = 0$), we find perturbed quantities

It will be useful to invoke the vector identity

$$\vec{\nabla} \cdot (\rho \vec{v}) = \vec{v} \cdot \vec{\nabla} \rho + \rho (\vec{\nabla} \cdot \vec{v}). \quad (153)$$

$$\begin{aligned} \frac{\partial \rho_1}{\partial t} &= -\rho_0 \vec{\nabla} \cdot \vec{v}_1 \\ \frac{\partial \vec{v}_1}{\partial t} &= -\vec{\nabla} \Phi_1 - \frac{c_s^2}{\rho_0} \vec{\nabla} \rho_1. \end{aligned}$$

We take the divergence of the latter equation,

$$\frac{\partial(\vec{\nabla} \cdot \vec{v}_1)}{\partial t} = -4\pi G \rho_1 - \frac{c_s^2}{\rho_0} \nabla^2 \rho_1, \quad (154)$$

and recognize that the term $\vec{\nabla} \cdot \vec{v}_1$ can be substituted with the former equation:

$$\frac{\partial^2 \rho_1}{\partial t^2} = -4\pi G \rho_0 \rho_1 - c_s^2 \nabla^2 \rho_1. \quad (155)$$

If the perturbation propagates as a plane wave ($\rho_1 \propto e^{-i(\omega t - kx)}$), then it is subject to a dispersion relation

$$\omega^2 = c_s^2 \left[k^2 - \left(\sqrt{4\pi G \rho_0} / c_s \right)^2 \right]. \quad (156)$$

When $k < \sqrt{4\pi G \rho_0}/c_s$, ω is imaginary, and the perturbations will have an exponentially growing and damping mode. Throwing out the damped solution, we see how our instability grows exponentially. The critical wavenumber, k_* , corresponds to a length scale (inversely); this is called the **Jean's length**:³⁰

$$L_J = \frac{2\pi}{k_*} = \sqrt{\frac{\pi \gamma k_B T}{\mu m_H G \rho_0}}, \quad (157)$$

or an equivalent **Jean's mass**:

$$M_J = \rho_0 L_J^3 = \left(\frac{\pi \gamma k_B T}{\mu m_H G} \right)^{3/2} \rho_0^{-1/2}. \quad (158)$$

This is expressed as a *thermal* Jean's mass, at which gravitational energy exceeds thermal energy.

Using typical values ($\mu \simeq 2$, $\gamma = 3/5$), we find

$$\frac{M_J}{M_\odot} = 1.2 \left(\frac{T}{10 \text{ K}} \right)^{3/2} \left(\frac{n_0}{10^3 \text{ cm}^{-3}} \right)^{-1/2},$$

suggestive of the importance of the Jean's criterion in star formation, seeing as this mass is close to that of our Sun.

Magnetic fields and star formation

The free-fall timescale is too short to be physically plausible, because it's solely based on the thermal pressure. We will invoke the magnetic pressure to help slow down gravitational free fall.

For a uniform sphere, the gravitational binding energy is $U_G = 3GM^2/5R$, and the magnetic energy content is $U_B = \int dV B^2/8\pi = B^2 R^3/6$. If we set these equal, the back-of-the-envelope critical magnetic mass turns out to be

$$M_{\text{crit}} = \left(\frac{5}{18G} \right)^{1/2} B R^2 \simeq 0.17 \frac{\Phi_B}{\sqrt{G}},$$

where the magnetic flux is $\Phi_B = \pi B R^2$. For typical systems in the ISM, this critical mass is *larger* than the Jeans mass—so that magnetic fields actually provide support against collapse, rather than thermal pressure.

For subcritical cloud masses, **ambipolar diffusion**—where neutral and ionized material slip past each other and the B field diffuses outwards—allows less massive stars to collapse. The ambipolar diffusion timescale is $t_{\text{ambi}} \sim 10^6 \text{ yr}$ cf. $t_{\text{ff}} \sim 10^4 \text{ yr}$.

Friday, 6 November

Magnetic fields and star formation (continued)

As protostars collapse, angular momentum conservation leads to rotational support against gravitational instability. The rotating gas cloud has $E_{\text{grav}} \sim$

³⁰ We've snuck in the sound speed $c_s^2 = \gamma k_B T / \mu m_H$.

$-GM^2/R$ and $E_{\text{rot}} \sim J^2/2I = 5J^2/4MR^2$ (for a constant density sphere). Contraction will stop when these energies roughly balance each other, at

$$R_{\text{min}} \sim \frac{5}{4} \frac{J^2}{GM^3}.$$

This can cause the protostellar disk to shrink by one or maybe two orders of magnitude, but nowhere near the many orders of magnitude needed to actually create a star. The catalyst for collapse is the magnetic field. The cloud will collapse along (e.g., parallel to) the B field lines to make a pancake shape, and twist the B field lines to corotate in the ambient gas. This is described in Draine §41.5, and can be quantified with a **magnetic spin-down timescale** of $t_B \sim 8000$ yr. The magnetic spin-down timescale is always shorter than the contraction timescale, which means that the B fields can robustly transfer angular momentum from the system and allow for collapse.

Shocks

A shock is a pressure-driven disturbances travel faster than the signal speed, or **magnetosonic speed**³¹ and increase the entropy of the fluid through which they pass.

There are two ways to produce shocks. The first way is by the steepening of a (finite-amplitude) sound wave, if $\gamma = c_P/c_V > 1$. Another way is via supersonic compressive perturbations (e.g., like a supersonic jet or the crack of a whip). Below, in Fig 10, some relevant quantities are plotted against position. Note that we work in the shock frame, so that the shock is moving to the right.

³¹ On Earth, we think about sonic booms and such, but if a magnetic field allows for signal propagation, then we should care about the magnetosonic speed. The magnetosonic speed is simply the sum of the sound speed and Alfvén speed:

$$c_{\text{ms}}^2 = \frac{\gamma P}{\rho} + \frac{B^2}{4\pi\rho}. \quad (159)$$

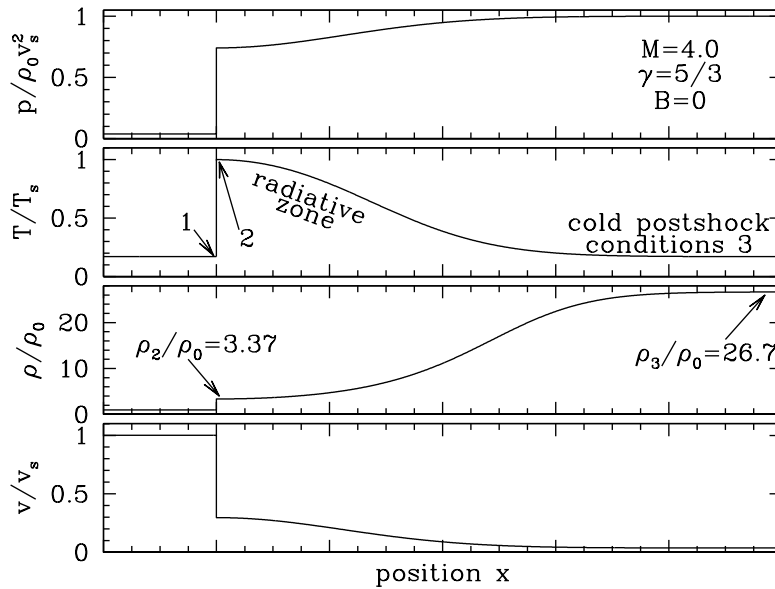


Figure 10: The pressure, temperature, density, and fluid velocity (in the shock frame) of a J -type shock (Draine Figure 36.1).

In the ISM, shocks can be generated by SN explosions, stellar winds, gas cloud collisions, pressure-driven expansion of an H II region, infall of material onto a cloud, or dynamical cloud perturbations (e.g., a spiral shock)—to give a few examples.

There are two categories of shocks, which are shown in the figure to the right (Draine figure 36.4):

1. **J-type (jump)**: The gas stops suddenly, is heated to high temperatures, and there is very little relaxation. The shock front is much thinner than the post-shock relaxation layer.
2. **C-type (continuous)**: These (weaker) shocks tend to occur in magnetized gas that have a low ionization fraction. Heating and cooling are balanced and determine the temperature in the C-shock.

Within molecular clouds, shocks with $v < 40 \text{ km s}^{-1}$ are likely C-type shocks, and with $v > 40 \text{ km s}^{-1}$ are likely J-type shocks.

To model our shock, we will work in a shock frame, and assume an infinite plane geometry. The fluid moves in the \hat{x} direction. Quantities with a 0 subscript are pre-shock, and quantities with a 1 subscript are post-shock.

We will calculate the **jump conditions** across a J-type shock. Starting with the continuity equation (or Equation 146), we note that the shock is stationary in the shock frame, so that

$$\frac{\partial \rho}{\partial t} = 0, \text{ and that } \vec{\nabla} \cdot (\rho \vec{v}) = \frac{\partial}{\partial x} (\rho v_x) = 0.$$

We then integrate from pre- to post-shock, finding that

$$0 = \int_{\text{pre}}^{\text{post}} dx \frac{\partial}{\partial x} (\rho v_x) = (\rho v_x)_1 - (\rho v_x)_0,$$

$$\rho_0 v_0 = \rho_1 v_1. \quad (160)$$

This is the mass conservation across the shock. The momentum flux is conserved by examining

$$\rho \frac{d\vec{v}}{dt} = -\vec{\nabla} \left(P + \frac{B^2}{8\pi} \right) + \frac{1}{4\pi} (\vec{v} \cdot \vec{\nabla}) \vec{B} - \rho \Phi,$$

which ignores the viscosity on both sides of the shock. Details are shown in equations (10-14) of the Lecture 19 equation sheet—note that we've used the product rule and that $\frac{\partial}{\partial x} (\rho v_x) = 0$ to derive equation (13). If we assume a small (negligible) change in gravity across the shock, then the right-hand side of equation (12) is zero, so that

$$\rho_0 v_0^2 + P_0 + \frac{1}{8\pi} (B_{y0}^2 + B_{z0}^2) = \rho_1 v_1^2 + P_1 + \frac{1}{8\pi} (B_{y1}^2 + B_{z1}^2). \quad (161)$$

Finally, to conserve energy, we start with equation (152), and ignore $\frac{\partial \Phi}{\partial x}$ and radiative heating and cooling, heat conduction, and substitute

$$u = \frac{1}{\gamma - 1} P \quad h = \frac{\gamma}{\gamma - 1} P,$$

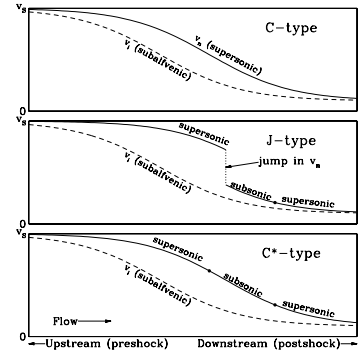


Figure 11: The velocity gradient over C-type (and C*-) and J-type shocks (Draine figure 36.4).

to derive (with no derivation—sorry!)

$$\frac{1}{2}\rho_0 v_0^3 + \frac{\gamma}{\gamma-1}v_0 P_0 + v_0 \frac{B_0^2}{8\pi} = \frac{1}{2}\rho_1 v_1^3 + \frac{\gamma}{\gamma-1}v_1 P_1 + v_1 \frac{B_1^2}{8\pi}. \quad (162)$$

These three are the Rankine-Hugoniot jump conditions. Additionally, we have an equation for magnetic flux conservation derived from conserving

$$\frac{\partial \vec{B}}{\partial t} = -\vec{\nabla} \times (\vec{v} \times \vec{B}),$$

so that

$$\vec{v}_0 \times \vec{B}_0 = \vec{v}_1 \times \vec{B}_1. \quad (163)$$

If the magnetic field is *perpendicular to* the propagation of the shock, then clearly (recalling again that we're in the shock front) we derive that $v_0 B_0 = v_1 B_1$.

To consolidate, we have our four quantities ρ, B, P, v , and we have these four equations of conservation. Our shock equation solution is nontrivial if $v_0^2 > c_{\text{ms}}^2$, of course. Let's define a compression factor

$$\xi \equiv \frac{\rho_1}{\rho_0} = \frac{B_1}{B_0} = \frac{v_0}{v_1}, \quad (164)$$

where we're collapsing the magnetic flux conservation and continuity equations to get this definition.

We are left with the remaining two equations:

$$\rho_0 v_0^2 + P_0 + \frac{B_0^2}{8\pi} = \rho_0 v_0^2 \frac{1}{\xi} + P_1 + \frac{B_0^2}{8\pi} \xi^2$$

$$\frac{1}{2}\rho_0 v_0^3 + \frac{\gamma}{\gamma-1}v_0 P_0 + v_0 \frac{B_0^2}{8\pi} = \frac{1}{2}\rho_0 v_0^3 \frac{1}{\xi^2} + \frac{\gamma}{\gamma-1}v_0 P_1 \frac{1}{\xi} + v_0 \frac{B_0^2}{8\pi} \xi.$$

We can define a **dimensionless Mach number**, \mathcal{M} ,

$$\mathcal{M}^2 \equiv \frac{v_0}{(\gamma P_0 / \rho_0) + (B_0^2 / 4\pi \rho_0)}. \quad (165)$$

In the limit of $B \rightarrow 0$, $\mathcal{M} \rightarrow \mathcal{M}_s = v_0 / c_s$, and in the limit of $P \rightarrow 0$, $\mathcal{M} \rightarrow \mathcal{M}_A = v_0 / v_A$.

In the limit of $B \rightarrow 0$:

$$\frac{P_1}{P_0} = \frac{2\gamma}{\gamma+1} \mathcal{M}_s^2 - \frac{\gamma-1}{\gamma+1}$$

$$\frac{\rho_0}{\rho_1} = \frac{\gamma-1}{\gamma+1} + \frac{2}{\gamma+1} \frac{1}{\mathcal{M}_s^2}.$$

For a very strong shock, $\mathcal{M}_s \gg 1$. Then (assuming an atomic gas with $\gamma = 5/3$), we can find that $\xi = 4$ (follow equations 32-35 on the lecture notes).

Friday, 13 November

Supernova remnants

The evolution of a supernova remnant occurs in three phases:

1. The **free expansion phase** when the swept up mass is negligible compared to the ejected mass ($M_{\text{ejecta}} \sim 1 - 5 M_{\odot}$): $M_{\text{swept}} \ll M_{\text{ejecta}}$. We then get a blastwave moving outward at a constant velocity:

$$v_0 = \left(\frac{2E_0}{M_{\text{ejecta}}} \right) = 10^4 \text{ km s}^{-1} \left(\frac{E_0}{10^{51} \text{ erg}} \right)^{1/2} \left(\frac{M_{\text{ejecta}}}{M_{\odot}} \right)^{-1/2}.$$

These speeds are highly supersonic—resulting in a jump shock—but there's so little material for it to do much at early stages. The free expansion phase ends when $M_{\text{ejecta}} \sim M_{\text{swept}}$, at a radius

$$R_1 \simeq \left(\frac{3M_{\text{ejecta}}}{4\pi\rho_0} \right)^{1/3} = 2.4 \text{ pc} \left(\frac{M_{\text{ejecta}}}{M_{\odot}} \right)^{1/3} \left(\frac{n_H}{0.5 \text{ cm}^{-3}} \right)^{-1/3}.$$

The thickness of the swept-up shell is estimated by using conservation of mass, or

$$\begin{aligned} \frac{4\pi}{3} R_1^3 \rho_0 &= 4\pi R_1 \Delta R_1 \rho_1 \\ \frac{\Delta R_1}{R_1} &= \frac{\rho_0}{3\rho_1} = \frac{1}{12}, \end{aligned}$$

where we recall that $\xi = 4$ for jump shocks and $\gamma = 5/3$. The time at which R_1 is reached is simply $t_1 = R_1/v_0 \approx 240 \text{ yr}$.

2. Next is the **adiabatic expansion**³² or **Sedov-Taylor phase**. Material can't cool efficiently (via radiation) because of the hot temperatures within the shock, so that the kinetic + internal thermal energy of gas is conserved.

The kinetic energy during the Sedov-Taylor phase is

$$E_{\text{kin}} = \frac{1}{2} \int_0^{R_s} dR 4\pi R^2 \rho v^2 = \frac{2\pi}{3} R_s^3 \rho_0 v_s^2 \phi_{\text{kin}}, \quad (166)$$

where obviously these quantities in the integral are not spatially constant, so that we introduce a fudge factor, ϕ_{kin} . Also, the swept-up mass is becoming increasingly important compared to M_{ejecta} . The internal energy is

$$\begin{aligned} E_{\text{int}} &= \int_0^{R_s} dR 4\pi R^2 \frac{P}{\gamma - 1} \\ &= \frac{4\pi}{3} R_s^3 v_s^2 \frac{P_1}{\gamma - 1} \phi_{\text{int}} \\ &= \frac{4\pi}{3} R_s^3 \rho_0 v_s^2 \frac{2}{\gamma^2 - 1} \phi_{\text{int}}, \end{aligned} \quad (167)$$

³² The shock is rather abrupt, and is irreversible (entropy increases), so that “adiabatic” is a poor description of this phase.

where we've introduced another fudge factor, and $P_1 = 2\rho_0 v_s^2 / (\gamma + 1)$. The ratio of kinetic-to-internal energy is $E_{\text{kin}}/E_{\text{int}} = (\gamma^2 - 1)\phi_{\text{kin}}/4\phi_{\text{int}}$, and the conservation of total energy results in

$$R_s^3 v_s^2 = \frac{2E_0}{2\pi\rho_0} \left(\phi_{\text{kin}} + \frac{4}{\gamma^2 - 1} \phi_{\text{int}} \right)^{-1}.$$

But because $v_s = \dot{R}_s$, let's define a term

$$\xi_0 \equiv \frac{75}{8\pi} \left(\phi_{\text{kin}} + \frac{4}{\gamma^2 - 1} \phi_{\text{int}} \right)^{-1},$$

to derive

$$R_s^3 \dot{R}_s^2 = \left(\frac{2}{5} \right)^2 \xi_0 \frac{E_0}{\rho_0}$$

$$R_s(t) = t^{2/5} \left(\frac{\xi_0 E_0}{\rho_0} \right)^{1/5} \quad (168)$$

$$v_s(t) = \frac{2}{5} t^{-3/5} \left(\frac{\xi_0 E_0}{\rho_0} \right)^{1/5}. \quad (169)$$

So now we have our shock speed evolving over time.

To get the temperature inside the bubble, we can apply the ideal gas law ($P = \rho kT / \mu m_H$) to the post-shock gas:

$$T_1 = \frac{\mu m_H}{k} \frac{2(\gamma - 1)}{(\gamma + 1)^2} v_0^2 \rightarrow \frac{\mu m_H}{k} \frac{3}{16} v_s^2. \quad (170)$$

Note that we had been using v_0 as pre-shock velocity, but the shock is always moving at v_s .

Combining the last two equations together, we find that

$$T_1 = \frac{\mu m_H}{k} \frac{3}{100} t^{-6/5} \left(\frac{\xi_0 E_0}{\rho_0} \right)^{2/5}. \quad (171)$$

Were we to do a more detailed analysis, we'd find that $\phi_{\text{kin}} = 0.417$, $\phi_{\text{int}} = 0.470$, and then $\xi_0 = 2.026$, resulting in the quantities

$$R_s(t) = 14.5 \text{ pc} \left(\frac{E_0}{10^{51} \text{ erg}} \right)^{1/5} \left(\frac{n_H}{0.5 \text{ cm}^{-3}} \right)^{-1/5} \left(\frac{t}{10^4 \text{ yr}} \right)^{2/5}$$

$$v_s(t) = 570 \text{ km s}^{-1} \left(\frac{E_0}{10^{51} \text{ erg}} \right)^{1/5} \left(\frac{n_H}{0.5 \text{ cm}^{-3}} \right)^{-1/5} \left(\frac{t}{10^4 \text{ yr}} \right)^{-3/5}$$

$$T_1(t) = 8 \times 10^6 \text{ K} \left(\frac{E_0}{10^{51} \text{ erg}} \right)^{2/5} \left(\frac{n_H}{0.5 \text{ cm}^{-3}} \right)^{-2/5} \left(\frac{t}{10^4 \text{ yr}} \right)^{-6/5}.$$

The temperature in a shock drops below 10^6 K at $t \sim 4 \times 10^4 \text{ yr}$, with $R_s \simeq 25 \text{ pc}$, $v_s \simeq 250 \text{ km s}^{-1}$, and the *swept-up mass* is an astounding $M_{\text{swept}} \simeq 1100 M_\odot$.

3. The final phase is called the **radiative expansion phase**, also known as the **snowplow phase**. Pre- and post- shock temperatures will be nearly

identical as radiation cools the post-shock medium. Momentum is conserved, slowing down the shock during this phase: $\frac{d}{dt}(Mv) = 0$. The Sedov-Taylor phase mass and velocity, M_{ST} and v_{ST} become:

$$M_{ST}v_{ST} = M(t)v_s(t) \propto R_s^3(t)v_s(t),$$

but since the left-hand side is constant, so $v_s(t) \propto R_s^{-3}$, and $R_s^{-3} \propto t^{-1/4}$.

However, the full solution to this problem requires inclusion of the thermal pressure from the interior bubble (see Draine chapter 39 for the full derivation).

The SN remnant basically “ends” after its motions become comparable to that of the ambient medium ($\sim 10 \text{ km s}^{-1}$). The final size for a typical SN remnant is $\sim 75 \text{ pc}$.

Tuesday, 17 November

The hot ionized medium

The **hot ionized medium** (HIM) is the dominant phase of the local ISM by volume.

Let's now consider the thermal energy density ($\frac{3}{2}nkT$) changed by a volumetric cooling rate,

$$\frac{d}{dt} \left(\frac{3}{2} n_H kT \right) = \Lambda,$$

which can be expressed in terms of a **cooling curve**, $F(T)$:

$$F(T) = \frac{\Lambda}{n_H^2} = \frac{3k}{2n_H} \frac{dT}{dt}. \quad (172)$$

This cooling curve is shown in the figure to the right.

How long does it take to cool from $T_0 \rightarrow T_0/2$?

$$\begin{aligned} t &= \frac{3k}{2n_H} \int_{T_0/2}^{T_0} \frac{dT}{F(T)} \\ &\sim \frac{3kT_0}{8n_H} \left[\frac{1}{F(T_0)} + \frac{1}{F(T_0/2)} \right]. \end{aligned}$$

In the last step, we've approximated $\int dT/F(T) \sim \Delta T \cdot \langle 1/F(T) \rangle$. These gas cooling timescales are shown in Lecture 21 Equation 8 (which assumes that $x = 0.1$ and n_H in units of cm^{-3}). There is a rapid drop in $F(T)$ at $T < 10^4 \text{ K}$, causing gas to “hang up” at $\sim 10^4 \text{ K}$ and not cool efficiently above that.

At this point, we can make a nice big table of phases of the ISM (sorta like we did on 2 October):

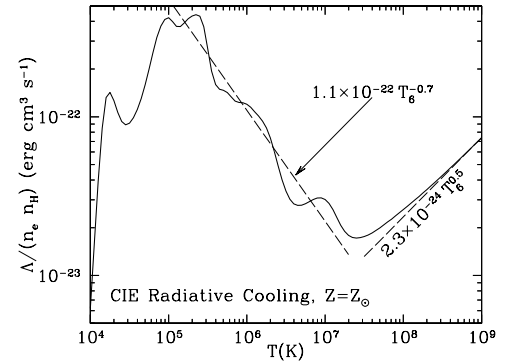


Figure 12: The cooling curve as a function of temperature (Draine Figure 34.1). Note that cooling is inefficient at $T \sim 10^7 \text{ K}$.

	n_H [cm ⁻³]	T [K]	nT [K cm ⁻³]	Fractional volume	Scale height [pc]	Surface density $M_\odot \text{ pc}^{-2}$
HIM	0.004	$\geq 3 \times 10^5$	≥ 1200	~ 0.5	3000	0.3
WIM	$0.3 - 10^4$	10^4	≥ 3000	0.1	900	1.1
WNM	0.6	5000	3000	0.4	220 – 400	2.9
CNM	30	100	3000	0.01	94	2.3
GMCs	$10^3 - 10^6$	10-50	> 2000	0.0001	75	1.0

As we can see, the *volume* is dominated by warmer, diffuse phases, whereas the *mass* is dominated by cooler, dense phases. All phases are basically in pressure equilibrium, except the self-gravitating GMCs.

We now introduce the **three-phase ISM**, which is introduced by McKee & Ostriker (1977). At first, the three phases were *clouds* (CNM), the *cloud surfaces* (WNM & WIM), and *inter-cloud gas* (HIM); they produced this model when contemplating the topology of the ISM when heated by SNe.

We can estimate the volumetric rate of supernova explosions in the Galactic disk to be $S \sim 10^{-13} \text{ pc}^{-3} \text{ yr}^{-1}$. The probability that a second supernova will go off within the volume of the a first's remnant is $P_{\text{overlap}} = \frac{4\pi}{3} R_{\text{fade}}^3 S t_{\text{fade}} \sim 0.3$.³³ It's evident that there will be overlap, and that the topology of the HIM won't be trivial (isolated bubbles). New SNe can rapidly drive pressure through other ionized bubbles, or "tunnel" along them.

³³ We know that $t_{\text{fade}} \sim 1.7 \times 10^6 \text{ yr}$.

The McKee & Ostriker model of the ISM includes:

1. Three phases of the ISM that are in pressure equilibrium
2. Pressure maintained by supernova remnants, with self-regulated pressure equilibrium
3. Mass equilibrium: cycle of material through the three phases of ISM:
 - some HIM becomes a cool, dense shell swept up at the edge of a SNR
 - shell fragments into cool cloud cores (CNM and eventually GMCs)
 - cores heated and ionized by contact with HIM, producing WNM & WIM
 - tenuous WNM & WIM material on outer edges of clouds *evaporating* back into HIM

Regarding the last point about **thermal evaporation**, we can calculate the *heat flow* into a cloud via $K|\nabla T|4\pi r^2 \simeq 4\pi r K T$ (which is in terms of the conductivity $K(T) \simeq 6.1 \times 10^{-7} (T/\text{K})^{-5/2} \text{ erg s}^{-1} \text{ K}^{-1} \text{ cm}^{-1}$). The energy deposited into the gas flowing away from the cloud causes evaporation at a rate

$$\dot{M} = \frac{16\pi}{25} \frac{\mu}{k} r K(T), \quad (173)$$

where μ is the average mass of the gas particles.

McKee & Ostriker predict a spectrum of cloud masses given by $dN \propto R^{-4} dR \propto M^{-2} dM$, such that small clouds dominate the surface area and

evaporation like $d\dot{M} \propto R dN \propto R^{-3} dR$. This model does pretty well, except that it underpredicts the mass in the WNM.

Friday, 20 November

Global star formation rates

Regarding this section title, “global” refers to the SFR of entire galaxies. Much of this content comes from Kennicutt & Evans (2012). What are good SFR estimators?

1. *Stellar UV continuum emission* (from stars with masses $M_* > 8 M_\odot$) estimates the SFR averaged over the past ~ 100 Myr.³⁴ We have to assume an IMF to convert from massive star formation ($\dot{M}_{>8 M_\odot}$) to total star formation ($\dot{M}_{>0.08 M_\odot}$). However, this will need to be corrected for dust.
2. *Recombination line emission* (usually the Balmer series, sometimes $\text{Pa}\alpha$, etc.) is dominated by stars with $M_* \geq 10 M_\odot$ and lifetimes ≤ 20 Myr. $\text{H}\alpha$ is considered to be the “gold standard” (Kennicutt).
3. *Collisional line emission*, e.g., the $[\text{O III}]$ 3727 Å doublet, can also be a SFR indicator. The $[\text{O III}]$ is a nebular emission line and can be seen at optical wavelengths out to $z = 1.5$.
4. *FIR dust continuum emission* at 40 – 120 μm, or *total IR emission* from 8 – 1000 μm, which is due to dust heating—mostly from OB stellar light—and traces SFR in highly dust obscured or very actively star-forming galaxies. To a small extent, it also processes general radiation from old stars.
5. *Total inventory of SF*, determined by summing the $L_{\text{UV}} + L_{\text{IR}}$ gets you the total SFR. Note that upwards of half the UV radiation can be extinguished by dust.³⁵
6. *X-ray luminosity* (excluding AGN) is dominated by thermal *bremsstrahlung* high-mass X-ray binaries (HMXBs) from massive stars. We need to correct for low-mass X-ray binaries (LMXBs), which trace stellar mass (but contribute only a bit of L_X), and also correct for obscuration by atoms such as Fe.
7. *Radio continuum emission*, whereby O stars ($M_* > 10 M_\odot$) ionize the gas and lead to free-free emission, and also explode in Type II SNe, whose remnants generate synchrotron emission. We also need to exclude AGN in this case.

³⁴ This 100 Myr timescale is considered “instantaneous.”

³⁵ B This isn’t always constant, as $L_{\text{IR}}/L_{\text{UV}} \sim 0.1 - 10^5$, which is a huge range! In very dusty galaxies, using just L_{IR} is reasonable.

Global SFR and gas content

Stars from molecular gas; we often will use the CO luminosity as a proxy for gas mass (or sometimes, HCN as a proxy for dense gas mass), and often

There is a remarkable result linking points 4 and 7 called the **radio-FIR correlation**. Radio emission at $\nu = 1.4$ GHz and FIR emission are tightly correlated over *five orders of magnitude* of SFR. The FIR emission traces OB stars whereas radio continuum emission traces mostly just O stars, so it’s odd to have such a strong correlation.

use the IR luminosity as a proxy for SFR. We can see that $L_{\text{IR}}/L'_{\text{CO}} \sim \text{SFR}/M_{\text{gas}}$, which is the inverse of the gas depletion time, or a *star formation efficiency* (SFE).³⁶

$$L_{\text{IR}}/L'_{\text{CO}} \propto (L'_{\text{CO}})^x \propto (L_{\text{IR}})^{\frac{x}{1+x}},$$

where $x \simeq 0.25 - 0.44$ (Gao & Solomon 2004a,b). Modulo the brightest IR galaxies (especially those with dust-obscured AGN), this relation should work well. Additionally, HCN ($1 - 0$) luminosity correlations more strongly and linearly with L_{IR} , with $L_{\text{IR}} \propto (L'_{\text{HCN}})^{1+x}$, where $x \simeq 0.0 - 0.05$. This makes sense, because a higher fraction of the densest gas is being turned into stars.

The Schmidt law

Schmidt (1959) found a power-law between SFR and gas density: $\text{SFR} \propto \rho^{1.5}$ (where the exponent is based on modern data). But it's very hard to figure out the density of a disk, so we often use 2-D surface densities, called the Schmidt-Kennicutt law (Kennicutt 1998):

$$\frac{\Sigma_{\text{SFR}}}{M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}} = (2.5 \pm 0.7) \times 10^{-4} \left(\frac{\Sigma_{\text{gas}}}{M_{\odot} \text{ pc}^{-2}} \right)^{1.4 \pm 0.15}. \quad (174)$$

There are competing interpretations for the basis of the Schmidt-Kennicutt law:

1. The SFR surface density is tied to the free-fall time scale ($t_{\text{ff}} \sim 1/\sqrt{G\rho}$), so that

$$\Sigma_{\text{SFR}} \propto \frac{\rho_{\text{gas}}}{t_{\text{ff}}} \propto \rho_{\text{gas}}^{1.5}.$$

A constant scale height is assumed, so that $\rho_{\text{gas}} \propto \Sigma_{\text{gas}}$ and we recover the Schmidt-Kennicutt law. But this requires some fine tuning since SF doesn't occur through free-fall, and personally I think this is rubbish.

2. If we try to invoke some relation between SFR and orbital timescales by using $\Omega_{\text{dyn}} = 2\pi/t_{\text{dyn}}$, which is the angular velocity at half-maximum radius of disk. Kennicutt (1998) finds

$$\frac{\Sigma_{\text{SFR}}}{M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}} = 0.017 \left(\frac{\Sigma_{\text{gas}}}{M_{\odot} \text{ pc}^{-2}} \right) \left(\frac{\Omega_{\text{dyn}}}{\text{yr}^{-1}} \right)$$

to be a reasonable fit, requiring a 10% conversion of gas into stars per orbit. This last point is also rather fine-tuned (across many galaxies).

3. Wong & Blitz (2002) note that star formation is a two-stage process, in which molecular clouds, and then stars are formed. They trace the power-

³⁶ Note that specific SFR (sSFR) has the same dimensions and is related to SFE, but is not the same thing as SFE.

law index, n :

$$\begin{aligned}
 n &= \frac{d \ln \Sigma_{\text{SFR}}}{d \ln \Sigma_{\text{gas}}} \\
 &= \frac{d \ln \Sigma_{\text{SFR}}}{d \ln \Sigma_{\text{H}_2}} \frac{d \ln \Sigma_{\text{H}_2}}{d \ln \Sigma_{\text{gas}}} \\
 &= n_{\text{mol}} \left(1 + \frac{d \ln f_{\text{mol}}}{d \ln \Sigma_{\text{gas}}} \right), \quad (175)
 \end{aligned}$$

where in the last step, n_{mol} is the index of a new Schmidt-like relation between SFR surface density and *molecular* gas surface density, $\Sigma_{\text{H}_2} = f_{\text{mol}} \Sigma_{\text{gas}}$ is the fraction of . As a result of the greater total gas surface density, the mid-plane pressure, P_{midplane} is higher but still in hydrostatic equilibrium:

$$P_{\text{midplane}} \simeq \frac{\pi}{2} G \Sigma_{\text{gas}} \left(\Sigma_{\text{gas}} + \frac{\sigma_{\text{gas}}^2}{\sigma_*^2} \Sigma_* \right) \quad (176)$$

where σ_{gas} , σ_* are the velocity dispersions of the gas and stars. This relation naturally explains why the $\text{SFR} \rightarrow 0$ at $\Sigma_{\text{gas}} \leq 5 - 10 M_{\odot} \text{pc}^{-2}$ is purely atomic gas.

Tuesday, 1 December

Spatially resolved star formation

Spiral arms of galaxies are **density waves** in which:

- gas enters the arm from behind (concave side)
- molecular clouds are compressed, creating massive molecular associations.
- pre-existing magnetic fields are compressed, resulting in a local enhancement of synchrotron emission
- massive star formation occurs (after a time delay for collapse)
- young stars continue to move towards the front, emerging from their birth clouds on the front side of the arms
- young stars that are blue but less massive (e.g., A stars) occupy a broader arm and are spread farther ahead

The phase difference between the optical and UV/radio synchrotron spiral arms can reveal the dynamics.

A **starburst** is a relatively compact region in a galaxy in which:

- large quantities of gas are available for star formation
- SFR in the region is $1 - 100 M_{\odot} \text{yr}^{-1}$
- SFR surface density $\Sigma_{\text{SFR}} \sim 10 M_{\odot} \text{yr}^{-1} \text{kpc}^2$ (distinctive condition of starbursts)
- the burst timescale is short ($\sim 10^8 \text{yr}$)
- a large fraction of stars can be formed in **super star clusters** (SSCs)

Examples can be found in the archetypical dwarf starburst, M82, and archetypical merging galaxy starburst, Arp 220.

Observational evidence for feedback

Feedback is the injection of energy and momentum back into the ISM by hot stars, supernovae, and active galactic nuclei. But why is it important? Indirect arguments that feedback “solves” some problems:

- Lower-mass galaxies have lower Z (metallicity) and lower effective yields (M_Z / M_*), because it’s easier to blow enriched gas out of the low-mass galaxies via SNe. These feedback mechanisms are called **galactic winds**.
- cD³⁷ galaxies at centers of galaxy clusters do not show the expected SFR, given rates at which the intracluster medium (ICM) should be cooling in local instabilities (the “cooling flow problem”). **AGN feedback** solves this problem by periodically injecting energy and momentum and shut off star formation.
- Models of galaxy evolution struggle to make star formation in the most massive galaxies shut off rapidly enough to explain their red colors (the “quenching problem”).³⁸ AGN feedback is also the likely culprit for shutting off star formation.

Large-scale galaxy winds are also known as **superwinds** (see, e.g., the review by Veilleux et al. 2005). Observational evidence for superwinds include:

1. ionized species (H α reveal edge-brightened features with (bi)conical geometry, suggestive of outflows
2. hot ($T \sim 10^6$ K) X-ray emitting gas, due to free-free emission; also edge-on galaxies show large radio halos
3. outflowing neutral gas—a key line of investigation for winds; locally we can study the Na D doublet at 5890, 5896 Å in absorption to measure outflows³⁹. Dwarf galaxies drive $0.001 - 1 M_\odot \text{ yr}^{-1}$, and ULIRGs drive $10 - 1000 M_\odot \text{ yr}^{-1}$.
4. outflowing molecular gas and dust
5. existence of a hot halo in Milky Way, which implies a “Galactic fountain.”⁴⁰

Driving mechanisms for galactic winds

It turns out that there are two possibilities for driving winds: **energy-driven** vs. **momentum-driven**. All winds take the path of least resistance, by the way, so they will be launched perpendicular to the disk. Let’s consider \dot{M}_W , the rate of mass loss, and v_∞ , the terminal velocity of a wind (after the material is far away enough and acceleration ceases). The energy-driven winds would have a constant rate of energy injection, $\dot{E} \approx \frac{1}{2} \dot{M}_W v_\infty^2$, and momentum-driven winds require a constant rate of momentum injection, $\dot{p} \approx \dot{M}_W v_\infty$.⁴¹

Energy-driven winds are also known as *thermally-driven* or *pressure-driven* (the latter of which might be misleading). A key parameter is the *thermalization efficiency*, ξ , which is the fraction of SN energy that heats the ISM.

³⁷ A backformation of cD is “central dominant,” although it was originally part of the Yerkes galaxy classification scheme.

³⁸ There is actually a low level of star formation in ellipticals too, but whatever.

³⁹ We can break the degeneracy between optical depth, τ , and covering factor, the fraction of background source covered by absorbing cloud, but measuring line ratios.

⁴⁰ This refers to gas being heated and blown out by feedback, cools far away, condenses into neutral clouds that “rain” back down onto the MW disk.

⁴¹ Check out Murray et al. (2005) for more on the theoretical underpinnings and justifications.

Another useful term (though not used in our crude analysis) is the *mass loading factor*, Λ , which is the ratio between total mass swept up in a wind and mass released directly by SNe, stars, or AGN.

Suppose we have a SFR of $1 M_\odot \text{ yr}^{-1}$ causes 1 SN every 100 yr (frequency ν_{SN}), each SN injects $10^{51} \cdot \xi$ erg into the ISM. The net heating rate is then

$$\dot{E}_{\text{SN}} = \xi E_{\text{SN}} \nu_{\text{SN}} \sim 3 \times 10^{40} \left(\frac{\xi}{0.1} \right) \left(\frac{\dot{M}_\star}{M_\odot \text{ yr}^{-1}} \right) \text{ erg s}^{-1}. \quad (177)$$

The net heating rate, after adding in massive stars, is $\dot{E}_{\text{mech}} \sim 2 \times \dot{E}_{\text{SN}}$ (obviously this is silly and fairly inaccurate). Were we to relate the luminosity to SFR by defining efficiency, ϵ :⁴²

⁴² SB stands for “starburst.”

$$L_{\text{SB}} = \epsilon \dot{M}_\star c^2 = 6 \times 10^{43} \left(\frac{\epsilon}{0.001} \right) \left(\frac{\dot{M}_\star}{M_\odot \text{ yr}^{-1}} \right) \left(\frac{v_\infty}{300 \text{ km s}^{-1}} \right)^{-2}. \quad (178)$$

A luminous O star releases $\sim 10^{53}$ erg over its lifetime, whereas the SN releases 10^{51} erg all at once. So $\dot{E}_{\text{mech}} = 0.01 \xi L_{\text{SB}}$, and we solve for the mass loss:

$$\dot{M}_W \approx \left(\frac{\xi}{0.1} \right) \left(\frac{\epsilon}{0.001} \right) \left(\frac{\dot{M}_\star}{M_\odot \text{ yr}^{-1}} \right) \left(\frac{v_\infty}{300 \text{ km s}^{-1}} \right)^{-2} M_\odot \text{ yr}^{-1}. \quad (179)$$

To unbind the gas, the energy injected in a dynamical time, $t_{\text{dyn}} \sim r/\sigma$, must exceed the gravitational energy (assuming, in this analysis, an isothermal sphere— $M = 2\sigma^2 r/G$):

$$\begin{aligned} \dot{E}_{\text{mech}} t_{\text{dyn}} &\geq \frac{G M M_{\text{gas}}}{r} \\ &\geq \frac{G \sigma}{r^2} f_{\text{gas}} M^2 = \frac{4 f_{\text{gas}} \sigma^5}{G} \\ L_{\text{SB}} = \frac{\dot{E}_{\text{mech}}}{0.01 \xi} &\geq 2 \times 10^{46} \left(\frac{f_{\text{gas}}}{0.1} \right) \left(\frac{\xi}{0.1} \right)^{-1} \left(\frac{\sigma}{200 \text{ km s}^{-1}} \right)^5 \text{ erg s}^{-1}. \end{aligned} \quad (180)$$

But... how do you get these winds without *ionizing* everything instantaneously? That becomes the trouble, even if the numbers might work out well.

Momentum-driven winds allow the gas to somehow be pushed out without the harsh ionization effects. A single SN drives $\sim 10 M_\odot$ of material at $\sim 3000 \text{ km s}^{-1}$. So the rate of momentum injection is

$$\dot{p}_{\text{SN}} \sim 2 \times 10^{33} \left(\frac{\dot{M}_\star}{M_\odot \text{ yr}^{-1}} \right) \text{ g cm s}^{-2}.$$

Again, we handwave-ly double that to account for winds, $\dot{p}_{\text{mech}} \sim 2 \times \dot{p}_{\text{SN}}$, add radiation pressure with optical depth τ (and therefore momentum flux $\tau L_{\text{SB}}/c$) to find

$$\dot{p}_{\text{rad}} \sim 2 \times 10^{33} \left(\frac{\tau}{1} \right) \left(\frac{\epsilon}{0.001} \right) \left(\frac{\dot{M}_\star}{M_\odot \text{ yr}^{-1}} \right) \text{ g cm s}^{-2}. \quad (181)$$

An order of magnitude estimate suggests that $\dot{M}_W v_\infty \simeq 3 \times L_{\text{SB}}/c$, so that

$$\dot{M}_W \approx \left(\frac{\epsilon}{0.001} \right) \left(\frac{\dot{M}_\star}{M_\odot \text{ yr}^{-1}} \right) \left(\frac{v_\infty}{300 \text{ km s}^{-1}} \right)^{-1} M_\odot \text{ yr}^{-1}. \quad (182)$$

It's similar to the energy-driven case in Equation (179), except that there's no dependence on ζ , and that now it's proportional to v_∞^{-1} .

But obviously, this whole time we should have known that *both* energy *and* momentum are important. We should count both, but since we're lazy, we just want to pick which one dominates.

Gas blowout will occur when radiation pressure + mechanical momentum flux exceeds gravity:

$$\begin{aligned} \frac{dp}{dt} &= -\frac{GMM_{\text{gas}}}{r^2} + \frac{3L_{\text{SB}}}{c} \\ &= -\frac{4f_{\text{gas}}\sigma^4}{G} + \frac{3L_{\text{SB}}}{c}, \end{aligned}$$

and so $\dot{p} \geq 0$ when

$$L_{\text{SB}} \geq \frac{4f_{\text{gas}}c\sigma^4}{3G} = 10^{46} \left(\frac{f_{\text{gas}}}{0.1} \right) \left(\frac{\sigma}{200 \text{ km s}^{-1}} \right)^4 \text{ erg s}^{-1}. \quad (183)$$

Below $\sigma \simeq 100 \cdot (\zeta/0.1 \text{ km s}^{-1})$, the energy-driven wind threshold is smaller, so that low-mass galaxies lose mass to energy-driven winds. High-mass galaxies would thus be more affected by momentum-driven winds. Intriguing and maybe correct speculations:

- It'd be useful, but difficult, to test whether winds from massive galaxies are momentum-driven.
- The Faber-Jackson relation for bulges/ellipticals ($L \propto \sigma^4$) might be explained?
- The $M_{\text{BH}} - \sigma$ relation is $M_{\text{BH}} \propto \sigma^4$.