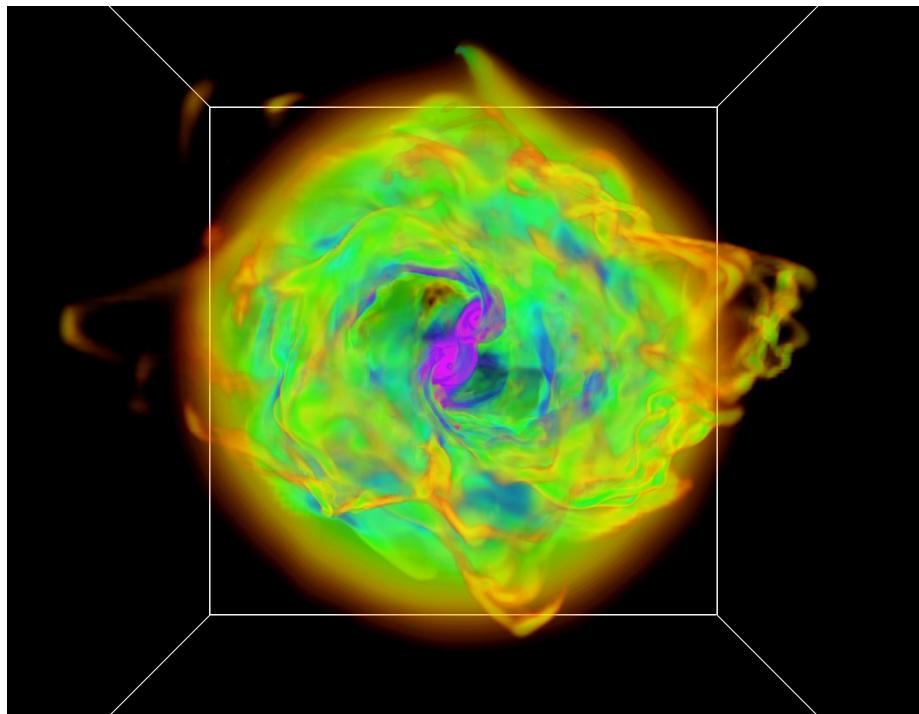


MARK R. KRUMHOLZ

# NOTES ON STAR FORMATION



THE OPEN ASTROPHYSICS BOOKSHELF

Original version: 2015 Mark R. Krumholz

PUBLISHED AS PART OF THE OPEN ASTROPHYSICS BOOKSHELF

<http://open-astrophysics-bookshelf.github.io/>

Licensed under the Creative Commons 1.0 Universal License, <http://creativecommons.org/publicdomain/zero/1.0/>.

# *Contents*

<i>I Introduction and Phenomenology</i>	<b>11</b>
<b>1 Observing the Cold Interstellar Medium</b>	<b>13</b>
<b>1.1 Observing Techniques</b>	<b>13</b>
<b>1.2 Observational Phenomenology</b>	<b>23</b>
<b>2 Observing Young Stars</b>	<b>29</b>
<b>2.1 Individual Stars</b>	<b>29</b>
<b>2.2 Statistics of Resolved Stellar Populations</b>	<b>32</b>
<b>2.3 Unresolved Stellar Populations and Extragalactic Star Formation</b>	<b>35</b>
<b>II Physical Processes</b>	<b>43</b>
<b>3 Chemistry and Thermodynamics</b>	<b>45</b>
<b>3.1 Chemical Processes in the Cold ISM</b>	<b>45</b>
<b>3.2 Thermodynamics of Molecular Gas</b>	<b>52</b>
<b>4 Gas Flows and Turbulence</b>	<b>59</b>
<b>4.1 Characteristic Numbers for Fluid Flow</b>	<b>59</b>
<b>4.2 Modeling Turbulence</b>	<b>62</b>
<b>4.3 Supersonic Turbulence</b>	<b>65</b>

	<i>Problem Set 1</i>	71
5	<i>Magnetic Fields and Magnetized Turbulence</i>	73
	5.1 <i>Observing Magnetic Fields</i>	73
	5.2 <i>Equations and Characteristic Numbers for Magnetized Turbulence</i>	75
	5.3 <i>Non-Ideal Magnetohydrodynamics</i>	78
6	<i>Gravitational Instability and Collapse</i>	85
	6.1 <i>The Virial Theorem</i>	85
	6.2 <i>Stability Conditions</i>	89
	6.3 <i>Pressureless Collapse</i>	97
7	<i>Stellar Feedback</i>	103
	7.1 <i>General Formalism</i>	103
	7.2 <i>Momentum-Driven Feedback Mechanisms</i>	107
	7.3 <i>(Partly) Energy-Driven Feedback Mechanisms</i>	110
	<i>III Star Formation Processes and Problems</i>	119
A	<i>Solutions to Problem Sets</i>	121
	<i>Solutions to Problem Set 1</i>	123
	<i>Bibliography</i>	127

# List of Figures

1.1	H <sub>2</sub> level diagram	13
1.2	Dust absorption opacity	15
1.3	<i>Herschel</i> map of IC 5146	16
1.4	Dust extinction map of the Pipe Nebula	17
1.5	COMPLETE spectra of Ophiuchus and Perseus	22
1.6	Distribution of H I and GMCs in M33	24
1.7	Distribution of CO(1 → 0) emission in M51	25
1.8	<sup>13</sup> CO(2 → 1) maps of Perseus	26
2.1	Outflow in CO(2 → 1)	30
2.2	Sample SEDs of protostellar cores	30
2.3	Bolometric temperatures of protostellar cores	31
2.4	Measured stellar IMFs in a variety of regions	34
2.5	Bolometric luminosity versus stellar population age	37
2.6	Optical spectra of galaxies across the Hubble sequence	38
4.1	Comparison of flows at varying Reynolds numbers	62
4.2	Experimental power spectra for Kolmogorov turbulence	66
4.3	Linewidth versus size in the Polaris Flare cloud	67
4.4	Volume rendering of the density field for supersonic turbulence	68
5.1	Sample Zeeman detection of a magnetic field	74
5.2	Comparison of simulations of Alfvénic and sub-Alfvénic turbulence	78
6.1	Magnetic field strength measurements	97
A.1	Solution to problem set 1, problem 2b	124



*Dedicated to my students, who have contributed  
tremendously to the development of this  
book.*



# *Introduction*

This book is based on a series of lectures given by the author in his graduate class on star formation, taught from 2009 - 2014 at UC Santa Cruz. It is intended for graduate students or advanced undergraduates in astronomy or physics, but does not presume detailed knowledge of particular areas of astrophysics (e.g., the interstellar medium or galactic structure). It is intended to provide a general overview of the field of star formation, at a level that would enable a student to begin independent research in the area.

This course covers the basics of star formation and ending at the transition to planet formation. The structure of the course / book is as follows. Each chapter corresponds roughly to a single lecture. The first two chapters begin with a discussion of observational techniques, and the basic phenomenology they reveal. The goal is to familiarize students with the basic techniques that will be used throughout, and to provide a common vocabulary for the rest of the course. The next five chapters provide a similar review of the basic physical processes that are important for star formation. Again, the goal is to provide a basis for what follows. The remaining chapters discuss star formation over a variety of scales, starting with the galactic scale and working our way down to the scales of individual stars and their disks. The course concludes with the clearing of disks and the transition to planet formation.

The "texts" intended to go with these notes are the review articles "[The Big Problems in Star Formation: the Star Formation Rate, Stellar Clustering, and the Initial Mass Function](#)", Krumholz, M. R., 2014, *Physics Reports*, 539, 49, which provides a snapshot of the theoretical literature as of the most recent time the course was given, and "[Star Formation in the Milky Way and Nearby Galaxies](#)", Kennicutt, R. C., & Evans, N. J., 2012, *Annual Reviews of Astronomy & Astrophysics*, 50, 531, which is more focused on observations. Another extremely useful reference is the series of review chapters from the [Protostars and Planets VI Conference](#), which took place in July 2013. Suggested background readings to accompany most chapters are listed at the chapter beginning. In addition to these background materials, most

chapters also include "suggested literature": papers from the recent literature whose content is relevant to the material covered in that chapter. These readings are included to help students engage with the active research literature, as well as the more general reviews.

In addition to the text and reading, this book contains five problem sets, which are interspersed with the chapters at appropriate locations. Solutions to the problems are included as an Appendix.

## **Part I**

# **Introduction and Phenomenology**



# 1

## *Observing the Cold Interstellar Medium*

This first chapter focuses on observations of interstellar gas. Because the interstellar clouds that form stars are generally cold, most (but not all) of these techniques require on infrared, sub-millimeter, and radio observations. Interpretation of the results is often highly non-trivial. This will naturally lead us to review some of the important radiative transfer physics that we need to keep in mind to understand the observations. With this background complete, we will then discuss the phenomenology of interstellar gas derived from these observations.

### *1.1 Observing Techniques*

#### *1.1.1 The Problem of H<sub>2</sub>*

Before we dive into all the tricks we use to observe the dense ISM, we have to mention why it is necessary to be so clever. Hydrogen is the most abundant element, and when it is in the form of free atomic hydrogen, it is relatively easy to observe. Hydrogen atoms have a hyperfine transition at 21 cm (1.4 GHz), associated with a transition from a state in which the spin of the electron is parallel to that of the proton to a state where it is anti-parallel. The energy associated with this transition is  $\ll 1$  K, so even in cold regions it can be excited. This line is seen in the Milky Way and in many nearby galaxies.

However, at the high densities where stars form, hydrogen tends to be molecular rather than atomic, and H<sub>2</sub> is extremely hard to observe directly. To understand why, we can look at an energy level diagram for rotational levels of H<sub>2</sub> (Figure 1.1). A diatomic molecule like H<sub>2</sub> has three types of excitation: electronic (corresponding to excitations of one or more of the electrons), vibrational (corresponding to vibrational motion of the two nuclei), and rotational (corresponding to rotation of the two nuclei about the center of mass). Generally electronic excitations are highest in energy scale, vibrational are next,

#### Suggested background reading:

- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531, sections 1 – 2

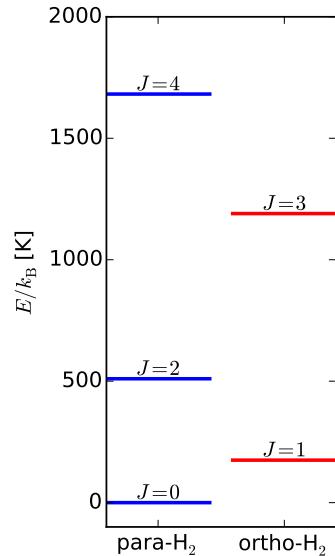


Figure 1.1: Level diagram for the rotational levels of para- and ortho-H<sub>2</sub>, showing the energy of each level. Level data are taken from <http://www.gemini.edu/sciops/instruments/nir/wavecal/h2lines.dat>.

and rotational are the lowest in energy.

For  $\text{H}_2$ , the first thing to notice is that the first excited state, the  $J = 1$  rotational state, is 175 K above the ground state. Since the dense ISM where molecules form is often also cold,  $T \sim 10$  K (as we will see later), almost no molecules will be in this excited state. However, it gets even worse:  $\text{H}_2$  is a homonuclear molecule, and for reasons of symmetry  $\Delta J = 1$  radiative transitions are forbidden in homonuclear molecules. Indeed, there is no electronic process by which a hydrogen molecule with odd  $J$  to turn into one with even  $J$ , and vice versa, because the allowed parity of  $J$  is determined by the spins of the hydrogen nuclei. We refer to the even  $J$  state as para- $\text{H}_2$ , and the odd  $J$  state as ortho- $\text{H}_2$ .

The observational significance of this is that there is no  $J = 1 \rightarrow 0$  emission. Instead, the lowest-lying transition is the  $J = 2 \rightarrow 0$  quadrupole. This is very weak, because it's a quadrupole. More importantly, however, the  $J = 2$  state is 510 K above the ground state. This means that, for a population in equilibrium at a temperature of 10 K, the fraction of molecules in the  $J = 2$  state is  $\sim e^{-510/10} \approx 10^{-22}$ ! In effect, in a molecular cloud there are simply no  $\text{H}_2$  molecules in states capable of emitting. The very high temperature required to excite the  $\text{H}_2$  molecular is its low mass: for a quantum oscillator or rotor the level spacing varies with reduced mass as  $m^{-1/2}$ . It is the low mass of the hydrogen atom that creates our problems.

Given this result, we see that, for the most part, observations of the most abundant species can only be done by proxy. Only in very rare circumstances is it possible to observe  $\text{H}_2$  directly – usually when there is a bright background UV source that allows us to see it in UV absorption rather than in emission. Since these circumstances do not generally prevail, we are forced to consider alternatives.

### 1.1.2 Dust Emission

The most conceptually straightforward proxy technique we use to study star-forming clouds is thermal dust emission. Interstellar gas clouds are always mixed with dust, and the dust grains emit thermal radiation which we can observe. The gas, in contrast, does not emit thermal radiation because it is nowhere near dense enough to reach equilibrium with the radiation field. Instead, gas emission comes primarily in the form of lines, which we will discuss a bit later.

Consider a cloud of gas of mass density  $\rho$  mixed with dust grains at a temperature  $T$ . The gas-dust mixture has an absorption opacity  $\kappa_\nu$  to radiation at frequency  $\nu$ . Although the vast majority of the mass is in gas rather than dust, the opacity will be almost entirely

due to the dust grains except for frequencies that happen to match the resonant absorption frequencies of atoms and molecules in the gas. Here we follow the standard astronomy convention that  $\kappa_\nu$  is the opacity per gram of material, with units of  $\text{cm}^2 \text{ g}^{-1}$ , i.e., we assign the gas an effective cross-sectional area that is blocked per gram of gas. For submillimeter observations, typical values of  $\kappa_\nu$  are  $\sim 0.01 \text{ cm}^2 \text{ g}^{-1}$ . Figure 1.2 shows a typical extinction curve for Milky Way dust.

Since essentially no interstellar cloud has a surface density  $> 100 \text{ g cm}^{-2}$ , absorption of radiation from the back of the cloud by gas in front of it is completely negligible. Thus, we can compute the emitted intensity very easily. The emissivity for gas of opacity  $\kappa_\nu$  is  $j_\nu = \kappa_\nu \rho B_\nu(T)$ , where  $j_\nu$  has units of  $\text{erg s}^{-1} \text{ cm}^{-3} \text{ sr}^{-1} \text{ Hz}^{-1}$ , i.e. it describes the number of ergs emitted in 1 second by 1  $\text{cm}^3$  of gas into a solid angle of 1 sr in a frequency range of 1 Hz and

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_B T} - 1} \quad (1.1)$$

is the Planck function.

Since none of this radiation is absorbed, we can compute the intensity transmitted along a given ray just by integrating the emission:

$$I_\nu = \int j_\nu ds = \Sigma \kappa_\nu B_\nu(T) = \tau_\nu B_\nu(T) \quad (1.2)$$

where  $\Sigma = \int \rho ds$  is the surface density of the cloud and  $\tau_\nu = \Sigma \kappa_\nu$  is the optical depth of the cloud at frequency  $\nu$ . Thus if we observe the intensity of emission from dust grains in a cloud, we determine the product of the optical depth and the Planck function, which is determined solely by the observing frequency and the gas temperature. If we know the temperature and the properties of the dust grains, we can therefore determine the column density of the gas in the cloud in each telescope beam.

Figure 1.3 show an example result using this technique. The advantage of this approach is that it is very straightforward. The major uncertainties are in the dust opacity, which we probably don't know better than a factor of few level, and in the gas temperature, which is also usually uncertain at the factor of  $\sim 2$  level. This produces a corresponding uncertainty in the conversion between dust emission and gas column density. Both of these can be improved substantially by observations that cover a wide variety of wavelengths, since these allow one to simultaneously fit the column density, dust opacity curve, and dust temperature.

Before the Herschel satellite (launched in 2009) such multi-wavelength observations were rare, because most of the dust emission was in at far-infrared wavelengths of several hundred  $\mu\text{m}$  that

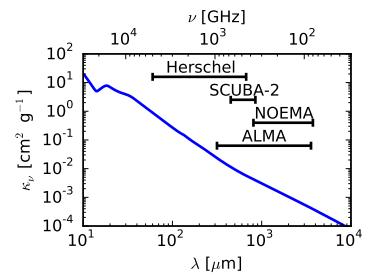


Figure 1.2: Milky Way dust absorption opacities per unit gas mass as a function of wavelength  $\lambda$  and frequency  $\nu$  in the infrared and sub-mm range, together with wavelength coverage of selected observational facilities. Dust opacities are taken from the model of Draine (2003) for  $R_V = 5.5$ .

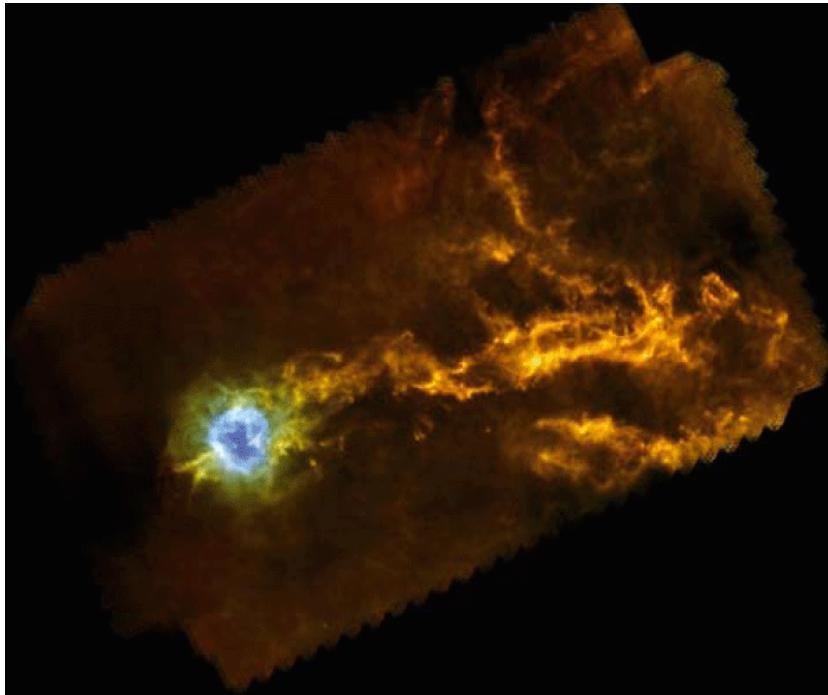


Figure 1.3: Three-color composite image of IC 5146 taken by the SPIRE and PACS instruments aboard *Herschel*. Red is SPIRE  $500\ \mu\text{m}$ , green is SPIRE  $250\ \mu\text{m}$  plus PACS  $160\ \mu\text{m}$ , and blue is PACS  $70\ \mu\text{m}$ . Image taken from Arzoumanian et al. (2011).

are inaccessible from the ground. *Herschel* was specifically targeted at this wavelength range, and has greatly improved our knowledge of cloud properties from dust emission.

### 1.1.3 Dust Absorption

A second related technique is, instead of looking at dust emission, looking at absorption of background starlight by dust, usually in the near infrared. The advantages of this compared to dust thermal emission are: (1) Since stars are bright compared to interstellar dust grains, and the observations are done in the near IR rather than the sub-mm, the available resolution is much, much higher. (2) Since opacity doesn't depend on temperature, the uncertainty in converting what we see into a column density is reduced. (3) We know the dust opacity curve in the infrared considerably better than we know it in the sub-mm, further reducing the uncertainty.

The major disadvantages are: (1) Due to the comparatively higher opacity in the infrared, it is only possible to use this technique for fairly diffuse regions; in denser regions the background stars are completely extinguished. (2) One needs a good, clean field of background stars to get something like a map, and only a few clouds have such favorable geometry. Probably the best example of this technique is the Pipe Nebula (Figure 1.4). In this case the calculation is even sim-

pler. One measures the extinction of the background star and then simply divides by the gas opacity to get a column density.

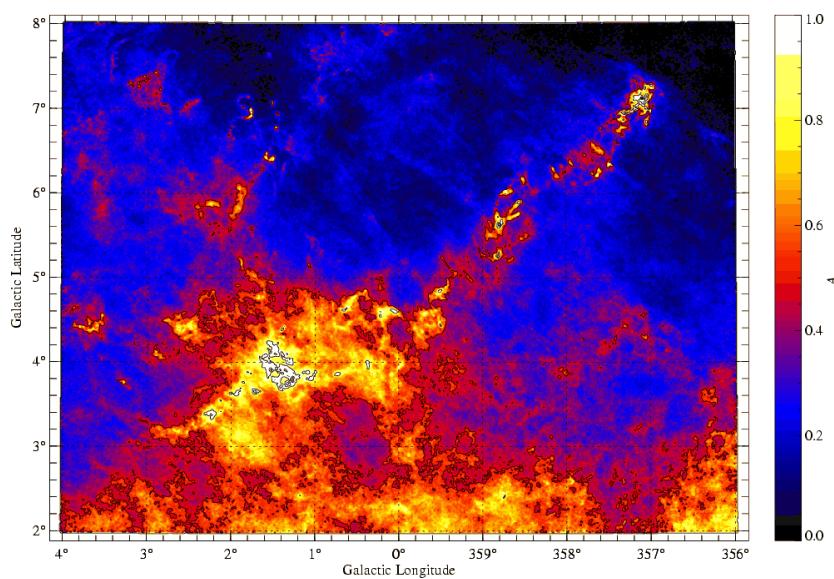


Figure 1.4: Extinction map of the Pipe Nebula (Lombardi et al., 2006).

#### 1.1.4 Molecular Lines

Much of what we know about star forming gas comes from observations of line emission. These are usually the most complex measurements in terms of the modeling and required to understand them. However, they are also by far the richest in terms of the information they provide. They are also among the most sensitive, since the lines can be very bright compared to continuum emission. Indeed, the great majority of what we know about the ISM beyond the local group comes from studying emission in the rotational lines of the CO molecule, because these (plus the C II line found in atomic regions) are by far the easiest types of emission to detect from the cold ISM.

The simplest line-emitting system is an atom or molecule with exactly two energy states, but this example contains most of the concepts we will need. In practice we're usually concerned with molecules rather than atoms, because in the dense parts of the ISM where stars form, most of the gas forms molecules.

*Einstein Coefficients and Collision Rates* Consider an atom or molecule of species X with two states separated by an energy  $E$ . Suppose we have a gas of such particles with number density  $n_X$  at temper-

ature  $T$ . The number density of atoms in the ground state is  $n_0$  and the number density in the excited state is  $n_1$ . At first suppose that this system does not radiate. In this case collisions between the atoms will eventually bring the two energy levels into thermal equilibrium. In that case, it is straightforward to compute  $n_0$  and  $n_1$ . They just follow a Maxwellian distribution, so  $n_1/n_0 = e^{-E/kT}$ , and thus we have  $n_0 = n_X/Z$  and  $n_1 = n_X e^{-E/kT}/Z$ , where  $Z = 1 + e^{-E/kT}$  is the partition function.

Now let us consider radiative transitions between these states. There are three processes: spontaneous emission, stimulated emission, and absorption, which are described by the three Einstein coefficients. In the radiation and diffuse matter classes you will learn about these in detail, but in studying star formation, we can often ignore stimulated emission and absorption, because the ambient radiation field is so weak that these processes occur at negligible rates.

The main exceptions to this are (1) when lines become extremely optically thick, so there are a lot of line photons bouncing around trapped inside a structure; (2) when the frequency of the transition in question is at very low energy, and interactions with CMB photons become significant. However, we will not discuss these cases in detail, and we'll just focus on spontaneous emission.

An atom in the excited state can spontaneously emit a photon and decay to the ground state. The rate at which this happens is described by the Einstein coefficient  $A_{10}$ , which has units of  $\text{s}^{-1}$ . Its meaning is simply that a population of  $n_1$  atoms in the excited state will decay to the ground state by spontaneous emission at a rate

$$\left( \frac{dn_1}{dt} \right)_{\text{spon. emis.}} = -A_{10}n_1 \quad (1.3)$$

atoms per  $\text{cm}^3$  per s, or equivalently that the  $e$ -folding time for decay is  $1/A_{10}$  seconds. For the molecules we'll be spending most of our time talking about, decay times are typically at most a few centuries, which is long compared to pretty much any time scale associated with star formation. Thus if spontaneous emission were the only process at work, all molecules would quickly decay to the ground state and we wouldn't see any emission.

However, in the dense interstellar environments where stars form, collisions occur frequently enough to create a population of excited molecules. Of course collisions involving excited molecules can also cause de-excitation, with the excess energy going into recoil rather than into a photon. Since hydrogen molecules are almost always the most abundant species in the dense regions we're going to think about, with helium second, we can generally only consider collisions

between our two-level atom and those partners. For the purposes of this exercise (and the problem sets), we'll take an even simple and ignore everything but H<sub>2</sub>.

The rate at which collisions cause transitions between states is a horrible quantum mechanical problem. We cannot even confidently calculate the energy levels of single isolated molecules except in the simplest cases, let alone the interactions between two colliding ones at arbitrary velocities and relative orientations. Exact calculations of collision rates are generally impossible. Instead, we either make due with approximations (at worst), or we try to make laboratory measurements. Things are bad enough that, for example, we often assume that the rates for collisions with H<sub>2</sub> molecules and He atoms are related by a constant factor.

Fortunately, as astronomers we generally leave these problems to chemists, and instead do what we always do: hide our ignorance behind a parameter. We let the rate at which collisions between species X and H<sub>2</sub> molecules induce transitions from the ground state to the excited state be

$$\left( \frac{dn_1}{dt} \right)_{\text{coll. exc.}} = k_{01} n_0 n, \quad (1.4)$$

where  $n$  is the number density of H<sub>2</sub> molecules and  $k_{01}$  has units of cm<sup>3</sup> s<sup>-1</sup>. In general  $k_{01}$  will be a function of the gas kinetic temperature  $T$ , but not of  $n$  (unless  $n$  is so high that three-body processes start to become important, which is almost never the case in the ISM).

The corresponding rate coefficient for collisional de-excitation is  $k_{10}$ , and the collisional de-excitation rate is

$$\left( \frac{dn_1}{dt} \right)_{\text{coll. de-exc.}} = -k_{10} n_1 n. \quad (1.5)$$

A little thought will convince you that  $k_{01}$  and  $k_{10}$  must have a specific relationship. Consider an extremely optically thick region where so few photons escape that radiative processes are not significant. If the gas is in equilibrium then we have

$$\frac{dn_1}{dt} = \left( \frac{dn_1}{dt} \right)_{\text{coll. exc.}} + \left( \frac{dn_1}{dt} \right)_{\text{coll. de-exc.}} = 0 \quad (1.6)$$

$$n(k_{01} n_0 - k_{10} n_1) = 0. \quad (1.7)$$

However, we also know that the equilibrium distribution is a Maxwellian, so  $n_1/n_0 = e^{-E/kT}$ . Thus we have

$$nn_0(k_{01} - k_{10}e^{-E/kT}) = 0 \quad (1.8)$$

$$k_{01} = k_{10}e^{-E/kT}. \quad (1.9)$$

This argument applies equally well between a pair of levels even for a complicated molecule with many levels instead of just 2. Thus, we

only need to know the rate of collisional excitation or de-excitation between any two levels to know the reverse rate.

*Critical Density and Density Inference* We are now in a position to write down the full equations of statistical equilibrium for the two-level system. In so doing, we will see that we can immediately use line emission to learn a great deal about the density of gas. In equilibrium we have

$$\frac{dn_1}{dt} = 0 \quad (1.10)$$

$$n_1 A_{10} + nn_1 k_{10} - nn_0 k_{01} = 0 \quad (1.11)$$

$$\frac{n_1}{n_0} (A_{10} + k_{10}n) - k_{01}n = 0 \quad (1.12)$$

$$\frac{n_1}{n_0} = \frac{k_{01}n}{A_{10} + k_{10}n} \quad (1.13)$$

$$= e^{-E/kT} \frac{1}{1 + A_{10}/(k_{10}n)} \quad (1.14)$$

This physical meaning of this expression is clear. If radiation is negligible compared to collisions, i.e.  $A_{10} \ll k_{10}n$ , then the ratio of level populations approaches the Maxwellian ratio  $e^{-E/kT}$ . As radiation becomes more important, i.e.  $A_{10}/(k_{10}n)$  get larger, the fraction in the upper level drops – the level population is sub-thermal. This is because radiative decays remove molecules from the upper state much faster than collisions re-populate it.

Since the collision rate depends on density and the radiative decay rate does not, the balance between these two processes depends on density. This make it convenient to introduce a critical density  $n_{\text{crit}}$ , defined by  $n_{\text{crit}} = A_{10}/k_{10}$ , so that

$$\frac{n_1}{n_0} = e^{-E/kT} \frac{1}{1 + n_{\text{crit}}/n}. \quad (1.15)$$

At densities much larger than  $n_{\text{crit}}$ , we expect the level population to be close to the Maxwellian value, and at densities much smaller than  $n_{\text{crit}}$  we expect the upper state to be under-populated relative to Maxwellian;  $n_{\text{crit}}$  itself is simply the density at which radiative and collisional de-excitations out of the upper state occur at the same rate.

This process of thermalization has important consequences for the line emission we see from molecules. The energy emission rate per molecule from the line is

$$\frac{\mathcal{L}}{n_X} = \frac{EA_{10}n_1}{n_X} \quad (1.16)$$

$$= EA_{10} \frac{n_1}{n_0 + n_1} \quad (1.17)$$

$$= EA_{10} \frac{n_1/n_0}{1 + n_1/n_0} \quad (1.18)$$

$$= EA_{10} \frac{e^{-E/kT}}{1 + e^{-E/kT} + n_{\text{crit}}/n} \quad (1.19)$$

$$= EA_{10} \frac{e^{-E/kT}}{Z + n_{\text{crit}}/n'} \quad (1.20)$$

where again  $Z$  is the partition function.

It is instructive to think about how this behaves in the limiting cases  $n \ll n_{\text{crit}}$  and  $n \gg n_{\text{crit}}$ . In the limit  $n \gg n_{\text{crit}}$ , the partition function  $Z$  dominates the denominator, and we get  $\mathcal{L}/n_X = EA_{10}e^{-E/kT}Z$ . This is just the energy per spontaneous emission times the spontaneous emission rate times the fraction of the population in the upper state when the gas is in statistical equilibrium. This is density-independent, so this means that at high density you just get a fixed amount of emission per molecule of the emitting species. The total luminosity is just proportional to the number of emitting molecules.

For  $n \ll n_{\text{crit}}$ , the second term dominates the denominator, and we get

$$\frac{\mathcal{L}}{n_X} \approx EA_{10}e^{-E/kT} \frac{n}{n_{\text{crit}}} \quad (1.21)$$

Thus at low density each molecule contributes an amount of light that is proportional to the ratio of density to critical density. Note that this is the ratio of collision partners, i.e. of H<sub>2</sub>, rather than the density of emitting molecules. The total luminosity varies as this ratio times the number of emitting molecules.

The practical effect of this is that different molecules tell us about different densities of gas in galaxies. Molecules with low critical densities reach the linear regime at low density, and since most of the mass tends to be at lower density, they probe this widespread, low-density component. Molecules with higher critical densities will have more of their emission contributed by higher density gas, and thus tell us about rarer, higher-density regions. This is all somewhat qualitative, since a transition between  $\mathcal{L}/n_X \propto n$  and  $\mathcal{L}/n_X \sim \text{constant}$  doesn't represent a particularly sharp change in behavior. Nonetheless, the luminosity ratios of lines with different critical densities are a very important diagnostic of the overall density distribution in the ISM.

As a caution, we should note that this is computed for optically thin emission. If the line is optically thick, we can no longer ignore stimulated emission and absorption processes, and not all emitted photons will escape from the cloud. CO is usually optically thick.

*Velocity and Temperature Inference* We can also use molecular lines to infer the velocity and temperature structure of gas if the line in

question is optically thin. For an optically thin line, the width of the line is determined primarily by the velocity distribution of the emitting molecules. The physics here is extremely simple. Suppose we have gas along our line of sight with a velocity distribution  $\psi(v)$ , i.e. the fraction of gas with velocities between  $v$  and  $v + dv$  is  $\psi(v)dv$ , and  $\int_{-\infty}^{\infty} \psi(v) dv = 0$ .

For an optically thin line, in the limit where natural and pressure-broadening of lines is negligible, we can think of emission producing a delta function in frequency in the rest frame of the gas. There is a one-to-one mapping between velocity and frequency. Thus emission from gas moving at a frequency  $v$  relative to us along our line of sight produces emission at a frequency  $\nu \approx \nu_0(1 - v/c)$ , where  $\nu_0$  is the central frequency of the line in the molecule's rest frame, and we assume  $v/c \ll 1$ . In this case the line profile is described trivially by  $\phi(\nu) = \psi(c(1 - \nu/\nu_0))$ .

We can measure  $\phi(\nu)$  directly, and this immediately tells us the velocity distribution  $\psi(v)$ . In general the velocity distribution of the gas  $\psi(v)$  is produced by a combination of thermal and non-thermal motions. Thermal motions arise from the Maxwellian velocity distribution of the gas, and produce a Maxwellian profile  $\phi(\nu) \propto e^{-(\nu - \nu_{\text{cen}})^2/\sigma_{\nu}^2}$ . Here  $\nu_{\text{cen}}$  is the central frequency of the line, which is  $\nu_{\text{cen}} = \nu_0(1 - \bar{v}/c)$ , where  $\bar{v}$  is the mean velocity of the gas along our line of sight. The width is  $\sigma_{\nu} = \sqrt{kT/\mu}/c$ , where  $T$  is the gas temperature and  $\mu$  is the mean mass of the emitting molecule. This is just the 1D Maxwellian distribution.

Non-thermal motions involve bulk flows of the gas, and can produce a variety of velocity distributions depending how the cloud is moving. Unfortunately even complicated motions often produce distributions that look something like Maxwellian distributions, just because of the central limit theorem: if you throw together a lot of random junk, the result is usually a Gaussian / Maxwellian distribution. Figure 1.5 shows an example of velocity distributions measured in two nearby star-forming clouds.

Determining whether a given line profile reflects predominantly thermal or non-thermal motion requires that we have a way of estimating the temperature independently. This can often be done by observing multiple lines of the same species. Our expression

$$\frac{\mathcal{L}}{n_X} = EA_{10} \frac{e^{-E/kT}}{Z + n_{\text{crit}}/n} \quad (1.22)$$

shows that the luminosity of a particular optically thin line is a function of the temperature  $T$ , the density  $n$ , and the number density of emitting molecules  $n_X$ . If we observe three transitions of the same molecule, then we have three equations in three unknowns and we

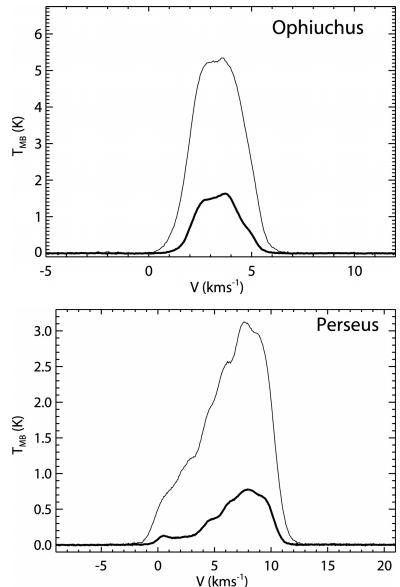


Figure 1.5: Position-integrated velocity distributions of  $^{12}\text{CO}$  (thin lines) and  $^{13}\text{CO}$  (thick lines) for the Ophiuchus and Perseus clouds, measured the COMPLETE survey (Ridge et al., 2006). The y axis shows the beam temperature.

can solve for  $n$ ,  $n_X$ , and  $T$  independently. Certain molecules, because of their level structures, make this technique particularly clean. The most famous example of this is ammonia,  $\text{NH}_3$ .

*Complications* Before moving on it is worth mentioning some complications that make it harder to interpret molecular line data. The first is optical depth: for many of the strongest lines and most abundant species, the line becomes optically thick. As a result observations in the line show only the surface a given cloud; emission from the back side of the cloud is absorbed by the front side. One can still obtain useful information from optically thick lines, but it requires a bit more thought. We'll discuss this a bit more in a few weeks when we discuss the large-scale distribution of giant molecular clouds.

The second complication is chemistry and abundances. The formation and destruction of molecules in the ISM is a complicated problem, and in general the abundance of any given species depends on the density, temperature, and radiation environment of the gas. At the edges of clouds, certain molecules may not be present because they are dissociated by the interstellar UV field. At high densities and low temperatures, many species freeze out onto the surfaces of dust grains. This is true for example of CO. One often sees that peaks in density found in dust emission maps correspond to local minima of CO emission. This is because in the densest parts of clouds CO goes out of the gas phase and forms CO ice on the surfaces of dust grains. Thus one must always be careful to investigate whether changes in molecular line emission are due to changes in gas bulk properties (e.g. density, temperature) or due to changes in the abundance of the emitting species.

## 1.2 Observational Phenomenology

### 1.2.1 Giant Molecular Clouds

As we just discussed, we usually can't observe  $\text{H}_2$  directly, so we are forced to do so by proxy. The most common proxy is the rotational lines of CO. These are useful because (1) CO is the single most abundant molecule in the ISM after  $\text{H}_2$ , (2) CO tends to be found in the same places as  $\text{H}_2$ , and (3) the CO molecule has a number of transitions that can be excited at the low temperatures found in molecular clouds – for example the CO  $J = 1$  state is only 5.5 K above the ground state. Indeed, the CO molecule is the primary coolant of molecular gas, so its excitation in effect sets the molecular gas temperature.

Later in the class we will discuss how one infers masses from CO

emission, and for now we'll just take it for granted that we can do so. By mass the Milky Way's ISM inside the solar circle is roughly 70% H I and 30% H<sub>2</sub>. The molecular fraction rises sharply toward the galactic center, reaching near unity in the molecular ring at  $\sim 3$  kpc, then falling to  $\sim 10\%$  our where we are. In other nearby galaxies the proportions vary from nearly all H I to nearly all H<sub>2</sub>.

In galaxies that are predominantly H I, like ours, the atomic gas tends to show a filamentary structure, with small clouds of molecular gas sitting on top of peaks in the H I distribution. In galaxies with large-scale spiral structure, the molecular gas closely tracks the optical spiral arms. Figures 1.6 and 1.7 show examples of the former and the latter, respectively. The physical reasons for the associations between molecular gas and H I, and between molecular clouds and spiral arms, are an interesting point that we will discuss later.

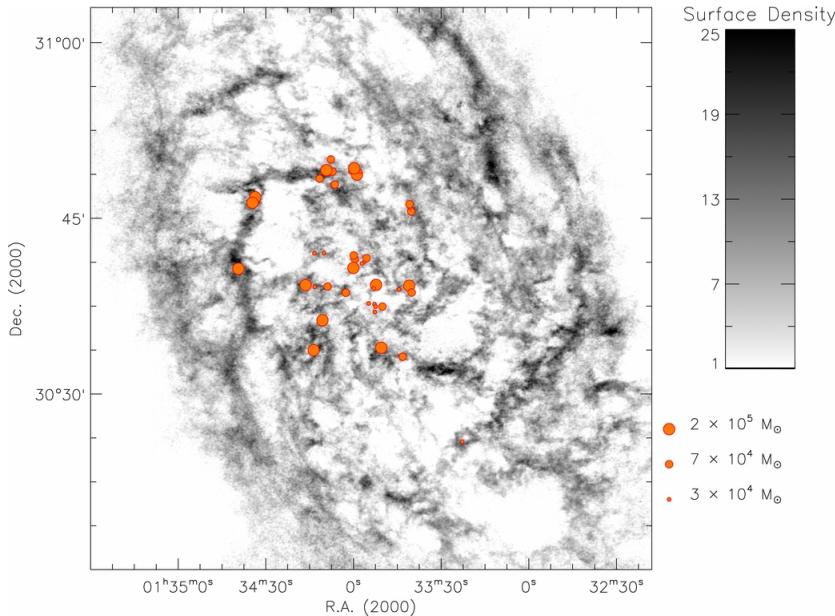


Figure 1.6: Map of H I in M33 (grayscale), with giant molecular clouds detected in CO(1  $\rightarrow$  0) overlaid (circles, sized by GMC mass) (Imara et al., 2011).

As the images show, molecular gas in galaxies that are predominantly atomic tends to be organized into discrete clouds, called giant molecular clouds (GMCs). These can have a range of masses; in the Milky Way the most massive are a few million  $M_{\odot}$ , but there is a spectrum that seems to continue down to at least  $10^4 M_{\odot}$ . This organization into GMCs is clearest where the gas is predominantly atomic. In regions where molecules make up most of the mass, the clouds can begin to run together into a predominantly molecular overall ISM.

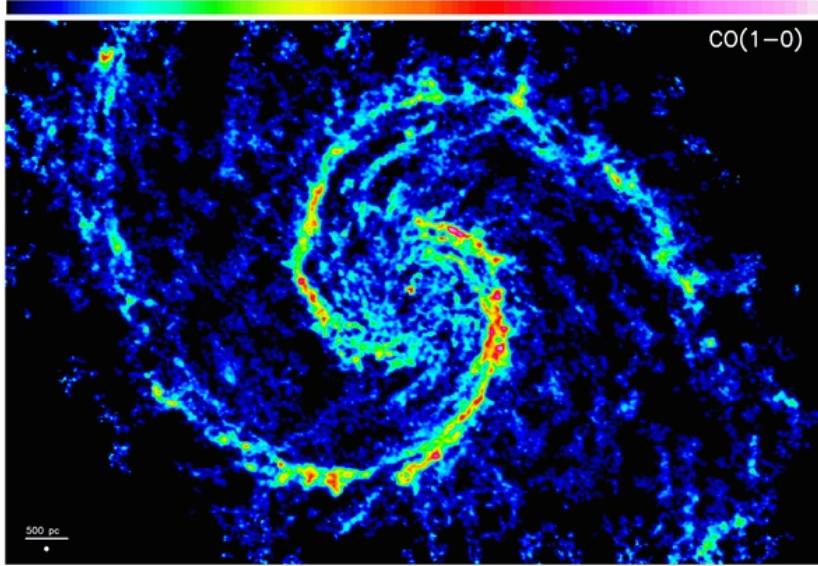


Figure 1.7: Map of CO( $1 \rightarrow 0$ ) emission in M51, as measured by the PdBI Arcsecond Whirlpool Survey (PAWS) project (Schinnerer et al., 2013).

### 1.2.2 Internal structure of GMCs

Giant molecular clouds are not spheres. They have complex internal structures. They tend to be highly filamentary and clumpy, with most of the mass in low density structures and only a little bit in very dense parts. However, if one computes a mean density by dividing the total mass by the rough volume occupied by the  $^{12}\text{CO}$  gas, the result is  $\sim 100 \text{ cm}^{-3}$ . Typical size scales for GMCs are tens of pc – the Perseus cloud shown is a small one by Galactic standards, but the most massive ones are found predominantly in the molecular ring, so our high resolution images are all of nearby small ones.

This complex structure on the sky is matched by a complex velocity structure. GMCs typically have velocity spreads that are much larger than the thermal sound speed of  $\sim 0.2 \text{ km s}^{-1}$  appropriate to 10 K gas. One can use different tracers to explore the distributions of gas at different densities in position-position-velocity space – at every position one obtains a spectrum that can be translated into a velocity distribution along that line of sight. The data can be slides into different velocities.

One can also get a sense of density and velocity structure by combining different molecular tracers. For example, the data set from COMPLETE (see Figure 1.5) consists of three-dimensional cubes of  $^{12}\text{CO}$  and  $^{13}\text{CO}$  emission in position-position-velocity space, and from this one can draw isosurfaces. Generally the  $^{12}\text{CO}$  isosurfaces contain the  $^{13}\text{CO}$  ones, as expected since the  $^{12}\text{CO}$  traces less dense gas and the  $^{13}\text{CO}$  traces more dense gas. The density increases as one

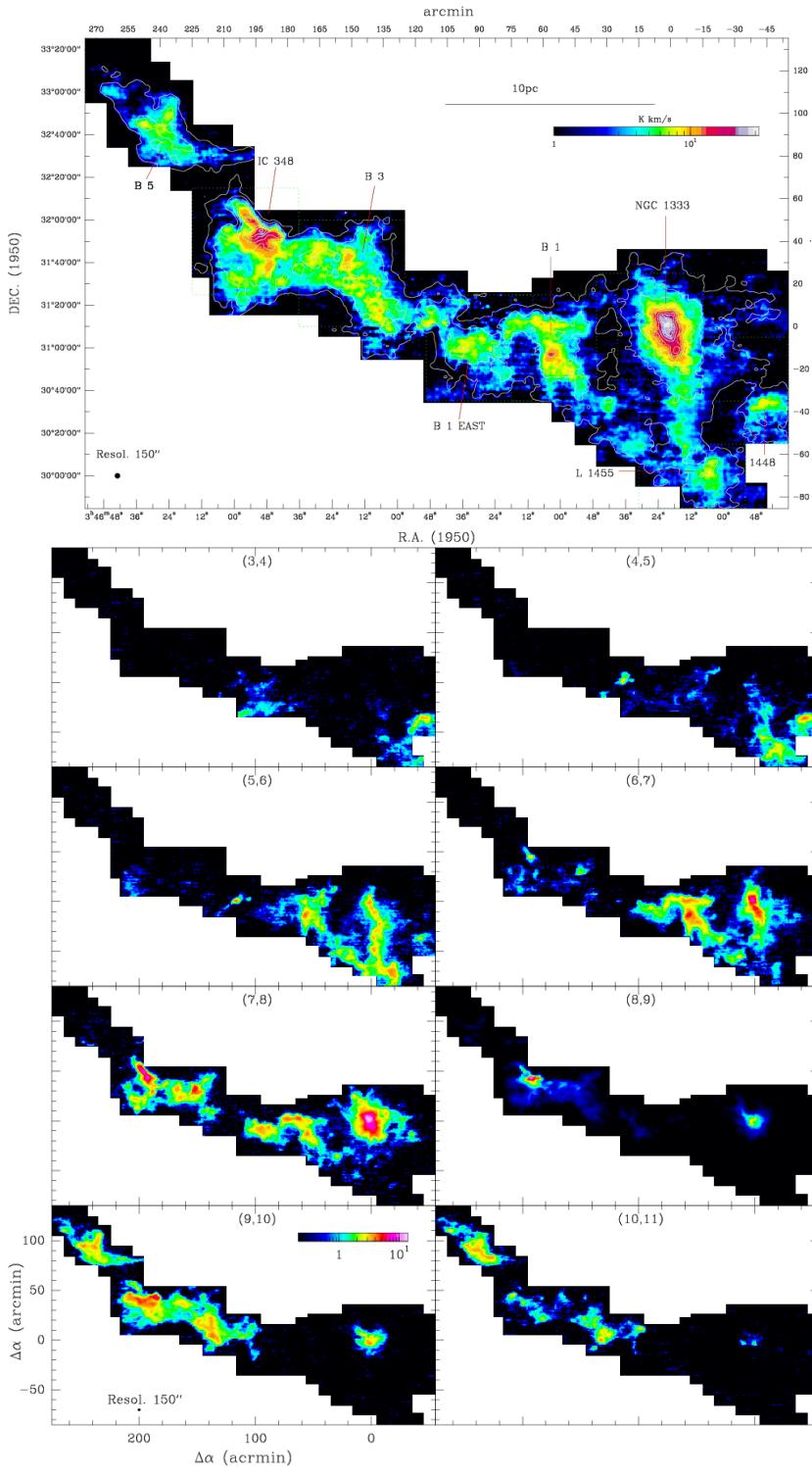


Figure 1.8: Map of the Perseus cloud in  $^{13}\text{CO}(2 \rightarrow 1)$  from Sun et al. (2006). The top panel shows the emission integrated over all velocities, while the bottom panel shows maps integrated over different velocity channels. In each sub-panel in the bottom, the numbers at the top indicate the velocity range (in  $\text{km s}^{-1}$ ) of the emission shown.

moves toward the cloud "center" in both position and velocity, but the morphology is not simple.

### 1.2.3 Cores

As we zoom into yet smaller scales, the density rises to  $10^5 - 10^7$  cm $^{-3}$  or more, while the mass decreases to a few  $M_\odot$ . These regions, called cores, tend to be strung out along filaments of lower density gas. Morphologically cores tend to be closer to round than the lower-density material around them. These objects are thought to be the progenitors of single stars or star systems. Cores are distinguished not just by simple, roundish density structures, but by similarly simple velocity structures. Unlike in GMCs, where the velocity dispersion is highly supersonic, in cores it tends to be subsonic. This is indicated by a thermal broadening that is comparable to what one would expect from purely thermal motion.



# 2

## *Observing Young Stars*

Having discussed how we observe interstellar gas that is forming stars, we now turn to the phenomenology of the young stars themselves. In this chapter we will first discuss individual young stars, then resolved young stellar populations, and then end by discussing unresolved stellar populations in the Milky Way and nearby galaxies.

### *2.1 Individual Stars*

Since we think star formation begins with a core that is purely gas, we expect to begin with a cloud that is cold and lacks a central point source. Once a protostar forms, it will begin gradually heating up the cloud, while the gas in the cloud collapses onto the protostar, reducing the opacity. Eventually enough material accretes from the envelope to render it transparent in the near infrared and eventually the optical, and we begin to be able to see the star directly for the first time. The star is left with an accretion disk, which gradually accretes and is then dispersed. Eventually the star contracts onto the main sequence.

This theoretical cartoon has been formalized into a system of classification of young stars based on observational diagnostics. At one end of this sequence lies purely gaseous sources where there is no evidence at all for the presence of a star, and at the other end lies ordinary main sequence stars. In between, objects are classified based on their emission in the infrared and sub-mm parts of the spectrum. These classifications probably give more of an impression of discrete evolutionary stages than is really warranted, but they nonetheless serve as a useful rough guide to the evolutionary state of a forming star.

Consider a core of mass  $\sim 1 M_{\odot}$ , seen in dust or molecular line emission. When a star first forms at its center, the star will be very low mass and very low luminosity, and will heat up only the dust nearest to it, and only by a very small amount. Thus the total light

### **Suggested background reading:**

- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531, section 3
- Krumholz, M. R. 2014, *Phys. Rep.*, 539, 49, section 2

output will still be dominated by the thermal emission of the dust at its equilibrium temperature. The spectral energy distribution of the source will therefore look just like that which prevailed before the star formed.

However, there might be other indicators that a star has formed. For example, the density distribution might show a very sharp, unresolved peak. Another sign that a star has formed might be the presence of an outflow, which, as we discuss below, all protostars seem to generate. Outflows coming from the center of a core can be detected in a few ways. Most directly, one can see bipolar, high velocity structures in molecular emission (Figure 2.1).

One can also detect indirect evidence of an outflow, from the presence of highly excited molecular line emission that is produced by shocks at hundreds of  $\text{km s}^{-1}$ . One example of such a line is SiO(2 → 1) line, which is generally seen in gas moving at several tens of  $\text{km s}^{-1}$  with temperatures of several hundred K – this is taken to be indication that emission in this line is produced in warm shocks. Since we know of no processes other than formation of a compact object with a  $\gtrsim 100 \text{ km s}^{-1}$  escape velocity that can accelerate gas in molecular clouds to such speeds, the presence of such an outflow is taken to indicate that a compact object has formed.

These are the earliest indications of star formation we have available to us. We call objects that show one of these signs, and do not fall into one of the other categories, class 0 sources. The dividing line between class 0 and class 1 is that the star begins to produce enough heating that there is non-trivial infrared emission. Before the advent of *Spitzer* and *Herschel*, the dividing line between class 0 and 1 was taken to be a non-detection in the IR, but as more sensitive IR telescopes became available, the detection limit went down of course. Today, the dividing line is taken to be a luminosity cut. A source is said to be class 0 if more than 0.5% of its total bolometric output emerges at wavelengths longer than 350  $\mu\text{m}$ , i.e., if  $L_{\text{smm}}/L_{\text{bol}} > 0.5\%$ , where  $L_{\text{smm}}$  is defined as the luminosity considering only wavelengths of 350  $\mu\text{m}$  and longer (Figure 2.2).

In practice, measuring  $L_{\text{smm}}$  can be tricky because it can be hard to get absolute luminosities (as opposed to relative ones) correct in the sub-mm, so it is also common to define the class 0-1 divide in terms of another quantity: the bolometric temperature  $T_{\text{bol}}$ . This is defined as the temperature of a blackbody that has the same flux-weighted mean frequency as the observed SED. That is, if  $F_\nu$  is the flux as a function of frequency from the observed source, then we define  $T_{\text{bol}}$  by the implicit equation

$$\frac{\int \nu B_\nu(T_{\text{bol}}) d\nu}{\int B_\nu(T_{\text{bol}}) d\nu} = \frac{\int \nu F_\nu d\nu}{\int F_\nu d\nu} \quad (2.1)$$

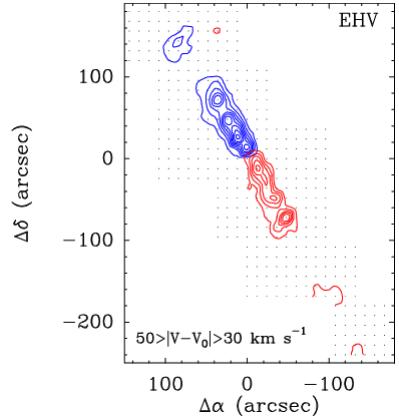


Figure 2.1: An integrated intensity map in CO(2 → 1), showing material at velocities between  $\pm 30 - 50 \text{ km s}^{-1}$  (blue and red contours, respectively) relative to the mean (Tafalla et al., 2004). Contours are spaced at intensities of  $1 \text{ K km s}^{-1}$ . The outflow shown is in the Taurus star-forming region.

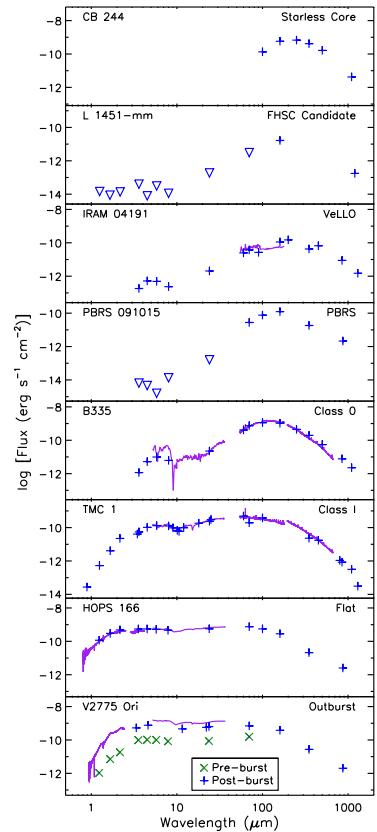


Figure 2.2: Sample spectral energy distributions (SEDs) of protostellar cores, together with the assigned class, as collected by Dunham et al. (2014).

The class 0-1 dividing line is also sometimes taken to be  $T_{\text{bol}} = 70$  K. In cases where  $L_{\text{smm}}$  is accurately measured,  $T_{\text{bol}}$  is observed to be a reasonably good proxy for  $L_{\text{smm}}/L_{\text{bol}}$  (Figure 2.3).

Once protostars reach class I, their evolution into further classes is defined in terms of the infrared spectral energy distribution. The motivating cartoon is as follows. At early times, the envelope of dust around the protostar is very optically thick at visible and even near infrared wavelengths. As a result, we don't get to see the stellar photosphere at all. All the radiation is absorbed by the envelope. The dust is in thermal equilibrium, so it re-radiates that energy. Since the radius of the sphere of dust is much larger than that of the star, and the luminosity radiated by the dust must ultimately be equal to that of the star, this emission must be at lower temperature and thus longer wavelengths. Thus as the radiation propagates outward through the dust it is shifted to longer and longer wavelengths. However, dust opacity decreases with wavelength (for reasons that will be / were discussed in the ISM class), and thus eventually the radiation is shifted to wavelengths where the remaining dust is optically thin, and it escapes. What we observe is therefore not a stellar photosphere, but a "dust photosphere".

Given this picture, the greater the column density of the dust around the star, the further it will have to diffuse in wavelength in order to escape. Thus the wavelength at which the emission peaks, or, roughly equivalently, the slope of the spectrum at a fixed wavelength, is a good diagnostic for the amount of circumstellar dust. Objects whose SEDs peak closer to the visible are presumed to be more evolved, because they have lost more of their envelopes.

More formally, this classification scheme was based on fluxes as measured by the IRAS satellite. We define

$$\alpha_{\text{IR}} = \frac{d \log(\lambda F_{\lambda})}{d \log \lambda}, \quad (2.2)$$

as the infrared spectral index, and in practice we measure  $\alpha_{\text{IR}}$  using two points from the IRAS SED: 2.2  $\mu\text{m}$  and 10 – 25  $\mu\text{m}$ . More positive values of  $\alpha_{\text{IR}}$  indicate SEDs that peak at longer wavelengths, further into the IR, while more negative values indicate SEDs that peak closer to visible. We define sources with  $\alpha_{\text{IR}} \geq 0.0$ , i.e. rising at longer wavelengths from 2 to 25  $\mu\text{m}$ , as class I sources. Alternately, in terms of bolometric temperature, the class I to class II transition is generally taken to be at 650 K (Figure 2.2).

As more of the envelope accretes, it eventually becomes optically thin at the peak emitting wavelengths of the stellar photosphere. In this case we see the stellar blackbody spectrum, but there is also excess infrared emission coming from the disk of warm, dusty gas that

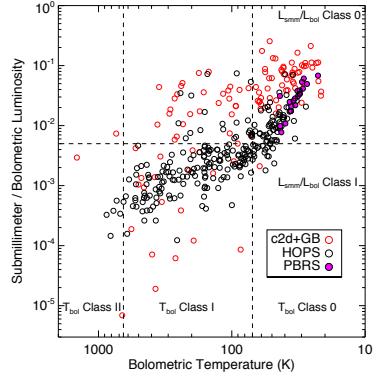


Figure 2.3: Bolometric temperatures of protostellar cores as compared to sub-mm to bolometric luminosity ratios (Dunham et al., 2014). The samples shown are from three different surveys as indicated in the legend.

still surrounds the star. Thus the SED looks like a stellar blackbody plus some extra emission at near- or mid-infrared wavelengths. Stars in this class are also known as classical T Tauri stars, named for the first object of the class, although the observational definition of a T Tauri star is somewhat different than the IR classification scheme, so the alignment may not be perfect.

In terms of  $\alpha_{\text{IR}}$ , these stars have indices in the range  $-1.6 < \alpha_{\text{IR}} < 0$ . (Depending on the author, the breakpoint may be placed at  $-1.5$  instead of  $-1.6$ . Some authors also introduce an intermediate classification between o and I, which they take to be  $-0.3 < \alpha_{\text{IR}} < 0.3$ .) A slope of around  $-1.6$  is what we expect for a bare stellar photosphere without any excess infrared emission coming from circumstellar material. Since the class II phase is the last one during which there is a disk of any significant mass, this is also presumably the phase where planet formation must occur.

The final class is class III, which in terms of SED have  $\alpha_{\text{IR}} < -1.6$ . Stars in this class correspond to weak line T Tauri stars. The SEDs of these stars look like bare stellar photospheres in the optical through the mid-infrared. If there is any IR excess at all, it is in the very far IR, indicating that the emitting circumstellar material is cool and located far from the star. The idea here is that the disk around them has begun to dissipate, and is either now optically thin at IR wavelengths or completely dissipated, so there is no strong IR excess.

However, these stars are still not mature main sequence stars. First of all, their temperatures and luminosities do not correspond to those of main sequence stars. Instead, they are still puffed up to larger radii, so they tend to have either lower effective temperatures or higher bolometric luminosities (or both) than main sequence stars of the same mass. Second, they show extremely high levels of magnetic activity compared to main sequence stars, producing high levels of x-ray emission. Third, they show lithium absorption lines in their atmospheres. This is significant because lithium is easily destroyed by nuclear reactions at high temperatures, and no main sequence stars with convective photospheres show Li absorption. Young stars show it only because there has not yet been time for all the Li to burn.

## 2.2 *Statistics of Resolved Stellar Populations*

Young stars tend to be born in the presence of other stars, rather than by themselves. This is not surprising: the gas cores from which they form are very small fragments,  $\sim 1 M_{\odot}$ , inside much larger,  $\sim 10^6 M_{\odot}$  clouds. It would be surprising if only one tiny fragment formed. In this next part of the class, we'll pull back to somewhat larger scales to look at the formation of stars in groups.

### 2.2.1 Multiplicity

The smallest scale we can look at beyond a single star is multiple systems. When we do so, we find that a significant fraction of stars are members of multiple systems – usually binaries, but also some triples, quadruples, and larger. The multiplicity is a strong function of stellar mass. The vast majority of B and earlier stars are multiples, while the majority of G, K, and M stars are singles. This means that most stars are single, but that most massive stars are multiples. The distribution of binary periods is extremely broad, ranging from hours to Myr. The origin of the distribution of periods, and of the mass-dependence of the multiplicity fraction, is a significant area of research in star formation theory.

### 2.2.2 The Initial Mass Function

If we observe a cluster of stars, the simplest thing to do is simply count up how many of them there are as a function of mass. The result is one of the most important objects in astrophysics, the initial mass function (IMF). This requires a bit of modeling, since of course what we can actually measure is a luminosity function, not a mass function. The problem of determining the IMF can be tackled in two ways: either by looking at stars in the solar neighborhood, or by looking at individual star clusters.

Looking at stars in the Solar neighborhood has the advantage that there are a lot of them compared to what you see in a clusters, so one gets a lot of statistical power. One also don't have to worry about two things that a major headache for studies of young clusters. First, young clusters usually have remaining bits of gas and dust around them, and this creates reddening that can vary with position and has to be modeled. Second, for clusters younger than  $\sim 10$  Myr, the stars are not on the main sequence yet. Since young stars are brighter than main sequence stars of the same mass, this produces an age-mass degeneracy that you have to break by obtaining more information than just luminosities (usually temperatures or colors), and then making pre-main sequence evolutionary models.

On the other hand, if we want to talk about the IMF of massive stars, we're pretty much stuck looking at young clusters. The same is also true for brown dwarfs. Since these fade with time, it is hard to find a large number of them outside of young clusters. An additional advantage of star clusters is that they are chemically homogenous, so you don't have to worry about chemical variations masquerading as mass variations.

A big problem for either method is correction for unresolved binaries, particularly at the low mass end, where the companions of

Protostellar evolution is covered in Chapter ??.

brighter stars are very hard to see. When one does all this, one gets results that look like this pretty much no matter where one looks (Figure 2.4). The basic features we see are a break peak centered around a few tenths of  $M_{\odot}$ , with a fairly steep fall off at higher masses that comes to resemble a powerlaw. There is also a fall-off at lower masses, although some authors argue for a second peak in the brown dwarf regime – this is still controversial, both because brown dwarfs are hard to find, and because their evolutionary tracks are less secure than those for more massive stars.

There have been recent claims of IMF variation extragalactically, but we'll get to that later.

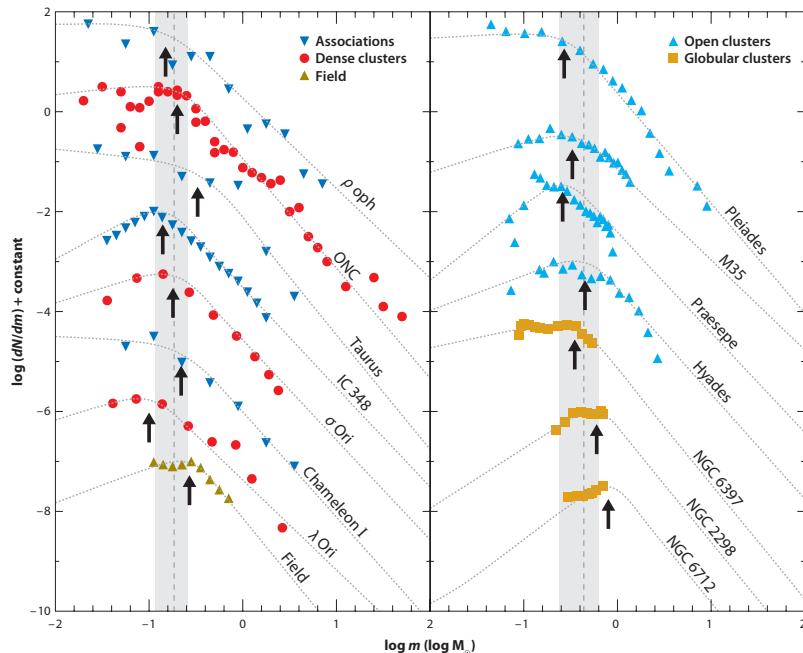


Figure 2.4: Stellar initial mass functions inferred for a wide variety of regions in the Milky Way, with the type of region as indicated in the legend. The dashed lines represent powerlaw fits to the observations in each region, with the black arrows indicating the best fit turnover mass.

The basic features illustrated in Figure 2.4 are a break peak centered around a few tenths of  $M_{\odot}$ , with a fairly steep fall off at higher masses that can be reasonably well-fit by a powerlaw. There is also a fall-off at lower masses, although some authors argue for a second peak in the brown dwarf regime – this is still controversial, both because brown dwarfs are hard to find, and because their evolutionary tracks are less secure than those for more massive stars.

The functional form shown in Figure 2.4 has been parameterized in a number of ways. Two of the most popular are from Kroupa (2001, 2002) and Chabrier (2003, 2005). Both of these fit the field star data, and the data individual clusters, within the error bars. The

See the reviews by Bastian et al. (2010) and Offner et al. (2014) for a thorough listing of alternate parameterizations.

functional form for Chabrier is

$$\frac{dn}{d \log m} \propto \begin{cases} \exp \left[ -\frac{(\log m - \log 0.22)^2}{2 \times 0.57^2} \right], & m < 1 \\ \exp \left[ -\frac{(-\log 0.22)^2}{2 \times 0.57^2} \right] m^{-1.35}, & m \geq 1 \end{cases}, \quad (2.3)$$

while the functional form for Kroupa is

$$\frac{dn}{d \log m} \propto \begin{cases} \left( \frac{m}{m_0} \right)^{-\alpha_0}, & m_0 < m < m_1 \\ \left( \frac{m_1}{m_0} \right)^{-\alpha_0} \left( \frac{m}{m_1} \right)^{-\alpha_1}, & m_1 < m < m_2 \\ \left[ \prod_{i=1}^{n-1} n \left( \frac{m_i}{m_{i-1}} \right)^{-\alpha_i} \right] \left( \frac{m}{m_n} \right)^{-\alpha_n}, & m_{n-1} < m < m_n \end{cases}, \quad (2.4)$$

with

$$\begin{aligned} \alpha_0 &= -0.7 \pm 0.7, & 0.01 < m/M_\odot < 0.08 \\ \alpha_1 &= 0.3 \pm 0.5, & 0.08 < m/M_\odot < 0.5 \\ \alpha_2 &= 1.3 \pm 0.3, & 0.5 < m/M_\odot < 1 \\ \alpha_3 &= 1.3 \pm 0.7, & 1 < m/M_\odot \end{aligned}. \quad (2.5)$$

In both equations (2.3) and (2.4) the mass  $m$  is in units of  $M_\odot$ .

### 2.3 Unresolved Stellar Populations and Extragalactic Star Formation

What about cases where we can't resolve the stellar population, as is usually the case for extragalactic work? What can we learn about star formation in that case? The answer turns out to be that the thing we can most directly measure is the star formation rate, and that doing so yields some very interesting results.

#### 2.3.1 Measuring the Star Formation Rate: General Theory

The first issue here is how we distinguish young stars beyond the Galaxy, since we can't obtain spectra, or even colors, for individual protostars as we can in the Milky Way. The answer is pretty much always the same: we exploit the fact that massive stars have short lifetimes, so if we measure the total number of massive stars in a galaxy, or some patch of a galaxy, then we are effectively measuring how many such stars formed there over some relatively short period.

We can formalize this theory a bit as follows.

Consider stars born with an initial mass function  $dn/dm$ . The mean stellar mass for this IMF is  $\bar{m}$ . A time  $t$  after a star is born, the star has a luminosity  $L(m, t)$ , where the luminosity can be bolometric, or integrated over some particular filter or wavelength range. First consider the simplest possible case, of a population of stars all born at the same instant at time 0. A time  $t$  later, the luminosity of the stars is

$$L(t) = N_* \int_0^\infty dm L(m, t) \frac{dn}{dm}, \quad (2.6)$$

where  $N_*$  is the total number of stars, and we have normalized the IMF so that  $\int (dn/dm)dm = 1$ . That is, we simply integrate the luminosity per star at time  $t$  over the mass distribution of stars.

Now consider a region, e.g., a galaxy, forming stars at a rate  $\dot{M}_*(t)$ ; in terms of number, the star formation rate is  $\dot{N}_*(t) = \dot{M}_*(t)/\bar{m}$ . To find the luminosity of the stellar population that is present today, we simply take the expression we just derived and integrate over all the possible stellar ages. Thus we have

$$L = \int_0^\infty dt \frac{\dot{M}_*(t)}{\bar{m}} \int_0^\infty dm L(m, t) \frac{dn}{dm}. \quad (2.7)$$

By itself this doesn't do us much good. The right hand side depends on the full star formation history  $\dot{M}_*(t)$ . However, let us assume that  $\dot{M}_*$  is constant in time. The integral still converges as long as  $L(m, t)$  reaches 0 after a finite time. In this case the integrals over  $m$  and  $t$  are separable, and we can rearrange them to

$$L = \frac{\dot{M}_*}{\bar{m}} \int_0^\infty dm \frac{dn}{dm} \int_0^\infty dt L(m, t) \equiv \frac{\dot{M}_*}{\bar{m}} \int_0^\infty dm \frac{dn}{dm} \langle Lt_{\text{life}} \rangle_m \quad (2.8)$$

In the final step we defined a new quantity  $\langle Lt_{\text{life}} \rangle_m$ , which has a simple physical meaning. It is simply the total amount of radiant energy that a star of mass  $m$  puts out over its lifetime.

Notice the expression on the right depends only on the constant star formation rate  $\dot{M}_*$ , the energy output  $\langle Lt_{\text{life}} \rangle_m$ , which we can generally calculate from stellar structure and evolution theory, and the IMF  $dn/dm$ . Thus if we measure  $L$  and use the "known" values of  $\langle Lt_{\text{life}} \rangle_m$  and  $dn/dm$ , we can measure the star formation rate. The physical picture to have here is that we're looking at a stellar population where there is equilibrium between new stars forming and old stars dying, so the total number of stars present and contributing to the light at any time is proportional to the rate at which they are forming. Thus a measurement of the light tells us about the star formation rate.

Is our assumption that  $\dot{M}_*$  is constant reasonable? That depends on the system we're looking at. If we're examining an entire galaxy that is forming stars quiescently and has not been externally perturbed, it is probably reasonable to assume that  $\dot{M}_*$  cannot vary on timescales much shorter than the dynamical time of the galaxy, which is  $\sim 200$  Myr for a galaxy like the Milky Way. If we choose to observe the luminosity at a wavelength where the light is coming mostly from stars with lifetimes shorter than this, so that  $L(m, t)$  reaches 0 (at least to good approximation) at times much less than 200 Myr, then assuming constant  $\dot{M}_*$  is quite reasonable.

However, it is always important to keep this constraint in mind – we can only measure the star formation rate as long as we believe

it to be constant on timescales long compared to the lifetimes of the stars responsible for generating the luminosity we're measuring. One can actually see how the ratio of luminosity to star formation rate behaves in systems that do not satisfy the constraint by generating synthetic stellar populations. In the simple case of star formation that turns on at some specific time, the luminosity just increases linearly in time until the first stars star evolving off the main sequence, and only becomes constant after  $\sim 4$  Myr.

The need to satisfy this constraint generally drives us to look for luminosities that are dominated by very massive stars, because these have very short lifetimes. Thus we will begin by discussing what luminosities we can measure that are particularly good at picking out massive stars. I should mention that this is far from an exhaustive list – people have come up with huge numbers of ways to infer star formation rates for galaxies at different redshifts. The accuracy of these techniques is highly variable, and in some cases is based on little more than a purely empirical calibration. We will just focus on the most reliable and widely used techniques which we can apply to relatively nearby galaxies.

### 2.3.2 Recombination Lines

Probably the most common technique, and the only one that can be used from the ground for most galaxies, is hydrogen recombination lines. To illustrate why this is useful, it is helpful to look at some galaxy spectra (Figure 2.6). As we move from quiescent E4 and SB galaxies to actively star-forming Sc and Sm/Im galaxies, there is a striking difference in the prominence of emission lines.

In the example optical spectra, the most prominent lines are the H $\alpha$  line at 6563 Å and the H $\beta$  at 4861 Å. These are lines produced by the  $3 \rightarrow 2$  and  $4 \rightarrow 2$ , respectively, electronic transitions in hydrogen atoms. In the infrared (not shown in the figure) are the Paschen  $\alpha$  and  $\beta$  lines at 1.87 and 1.28  $\mu\text{m}$ , and the Bracket  $\alpha$  and  $\gamma$  lines at 4.05 and 2.17  $\mu\text{m}$ . These come from the  $4 \rightarrow 3$ ,  $5 \rightarrow 3$ ,  $5 \rightarrow 4$ , and  $7 \rightarrow 4$  transitions.

Why are these related to star formation? The reason is that these lines come from H II regions: regions of ionized gas produced primarily by the ionizing radiation of young stars. Since only massive stars (larger than  $10 - 20 M_\odot$ ) produce significant ionizing fluxes), these lines indicate the presence of young stars. Within these ionized regions, one gets hydrogen line emission because atoms sometimes recombine to excited states rather than to the ground state. These excited atoms then radiatively decay down to the ground state, producing line emission in the process.

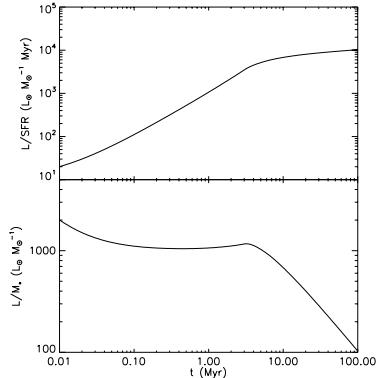


Figure 2.5: Bolometric luminosity versus time for stellar populations as a function of population age. The top panel shows the luminosity normalized by the star formation rate, while the bottom shows the luminosity normalized by the total stellar mass. Figure taken from Krumholz & Tan (2007).

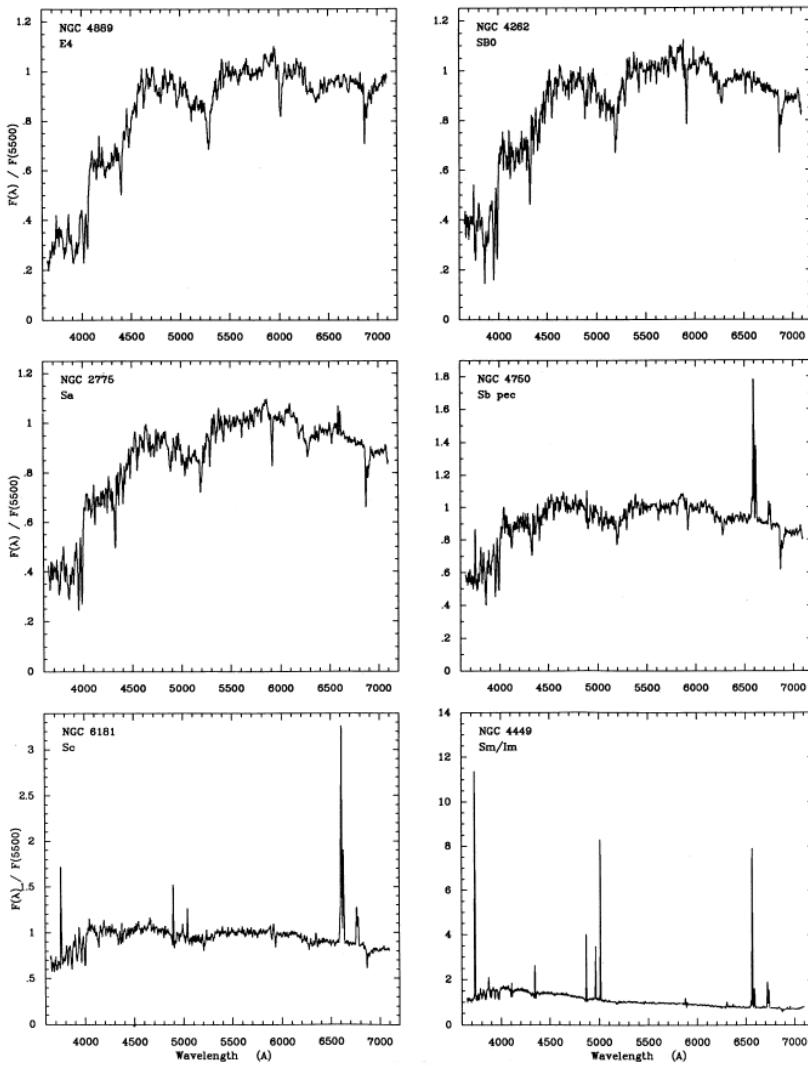


Figure 2.6: Example spectra of galaxies of varying Hubble type, from the atlas of Kennicutt (1992). In each panel, the galaxy name and Hubble type are listed.

Obtaining a numerical conversion between the observed luminosity in one of these lines and the star formation rate is a four-step process. First, we do the quantum statistical mechanics calculation to compute the yield of photons in the various lines per recombination. This can be done very precisely from first principles, using methods discussed in the diffuse matter class. Second, we equate the total recombination rate to the total ionization rate, and use this to determine the total rate of emission for the line in question per ionizing photon injected into the nebula. Third, we use stellar models to compute  $\langle L_{\text{ion}} t_{\text{life}} \rangle_m$ , the total ionizing photon production by a star of mass  $m$  over its lifetime. Four, we evaluate the integral over the IMF given by equation (2.8) to obtain the numerical conversion between star formation rate and luminosity. As of this writing, the most up-to-date resource for the results of such calculations is [Kennicutt & Evans \(2012\)](#).

Note that there are significant uncertainties in these numbers, the biggest one of which is the IMF. The reason the IMF matters so much is that the light is completely dominated by the massive stars, while the mass is all on the low mass stars we're not observing directly. To give an example, for a Chabrier IMF stars more massive than  $15 M_\odot$  contribute 99% of the total ionizing flux for a stellar population, but constitute less than 0.3% of the mass. Thus we are extrapolating by at least a factor of 30 in mass, and small changes in the IMF can produce large changes in the resulting ionizing luminosity to mass conversion.

Another complication is that some of the line emission is likely to be absorbed by dust grains within the source galaxy, and some of the ionizing photons are absorbed by dust grains rather than hydrogen atoms. Thus, one must make an extinction correction to the luminosities.

### 2.3.3 Radio Free-Free

A closely related method for measuring massive stars is to use radio free-free emission. An H II region emits not only optical lines from transition between energy levels of hydrogen and other atoms, it also emits free-free radiation in the radio. This is radiation produced by bremsstrahlung: free electrons scattering off ions, and emitting because accelerating charges emit.

Bremsstrahlung is a topic for the radiative processes class, and its application to H II regions is covered in the diffuse matter class, but the relevant point for us is that the flux from the H II region at radio wavelengths is proportional to  $n_e n_i$ , i.e. the product of the electron and ion densities. Since the recombination rate is also proportional

to  $n_e n_{\text{H}^+}$ , and the recombination rate just equals the ionization rate, this means that the free-free flux is directly proportional to the rate at which ionizing photons are injected into the H II region. Thus one can also do the same trick as above with a radio atlas of HII regions. One converts between free-free emission rate and ionization rate based on the physics of H II regions, and then converts between ionization rate and star formation rate using stellar structure and the IMF.

The free-free method is great in that radio emission is not obscured by dust, so one of the dust absorption corrections goes away. (The correction for absorption of ionizing photons by dust grains within the H II region remains, but this is generally only a few tens of percent.) This makes it more reliable than recombination lines, and also makes it the only technique we can use in the Milky Way, where obscuration is a big problem.

The downside is that the free-free emission is quite weak, and separating free-free from other sources of radio emission requires the ability to resolve individual H II regions. Thus this technique is really only feasible for the Milky Way and a few other nearby galaxies, since those are the only places where we can detect and resolve individual H II regions.

### 2.3.4 *Infrared*

The recombination line methods work well for galaxies that are like the Milky Way, but considerably less well for galaxies that are dustier and have higher star formation rates. This is because the dust extinction problem becomes severe, so that the vast majority of the Balmer emission may be absorbed. The Paschen and Brackett emission is much less sensitive to this, since those lines are in the IR, but even they can be extinguished in very dusty galaxies, and they are also much harder to use than H $\alpha$  and H $\beta$  because they are 1 – 2 orders of magnitude less bright intrinsically.

Instead, for dusty sources the tracer of choice is far infrared. The idea here is that, in a sufficiently dusty galaxy, essentially all stellar light will eventually be absorbed by dust grains. These grains will then re-emit the light in the infrared. As a result, the SED peaks in the IR. In this case one can simply use the total IR output of the galaxy as a sort of calorimeter, measuring the total bolometric power of the stars in that galaxy. In galaxies or regions of galaxies with high star formation rates, which tend to be where H $\alpha$  and other recombination line techniques fail, this bolometric power tends to be completely dominated by young stars. Since these stars die quickly, the total number present at any given time is simply proportional to the star formation rate.

The derivation of the conversion in this case is very straightforward – the  $L(m, t)$  that is required is just the total bolometric output of the stars. Again, results are given in Kennicutt & Evans (2012), and Problem Set 1 includes a calculation to be performed by the student. Of course IR emission has its problems too. First of all, it misses all the optical and UV radiation from young stars that is not absorbed within the galaxy. This is most if the output light for galaxies like the Milky Way, so IR is not a good indicator for it, although it can be good indicator for particular small regions that are much more opaque.

A second problem is that if the SFR is low, then young stars may not dominate the bolometric output, so the IR indicator can give an artificially high SFR. A more common problem for the dusty galaxies where IR tends to be used most is AGN contamination. If an AGN contributes significantly to the bolometric output of a galaxy, then that can masquerade as star formation. This can be hard to detect in a very dusty galaxy where most of the AGN light, light most of the starlight, is absorbed and reprocessed by dust.

### 2.3.5 *Ultraviolet*

Yet another way of measuring star formation rates is by the broad-band UV flux at wavelengths that are longer than 912 Å (corresponding to 13.6 eV, the energy required to ionized hydrogen) but shorter than where old stars put out most of their light. This range is roughly 1250 – 2500 Å. This light does not ionize hydrogen, so unlike shorter wavelengths it can get out of a galaxy.

For galaxies in the right redshift range this light gets redshifted into the visible, so we can see it from the ground. However, for local galaxies these wavelengths are only accessible from space (or least a balloon or rocket). For this reason this band was not used much until the launch of the GALEX satellite, which has detectors operating at 1300-1800 and 1800-2800 Å (the FUV and NUV bands, respectively). Sadly, the FUV detector on GALEX died, so, while it produced a great data set, we won't be getting any more of that any time soon.

Emission in these bands is dominated by stars with masses  $\sim 5 M_{\odot}$  and up, which have lifetimes of  $\sim 50$  Myr, so the total FUV light measures the star formation rate integrated over this time scale.

UV suffers from the same problems with dust extinction as H $\alpha$ , and they are perhaps even more severe, since opacity increases as frequency does. On the other hand, the UV is less sensitive to the IMF than H $\alpha$ , because ionizing photons come from hotter and thus more massive stars than UV ones. One more potential concern with UV is that 50 Myr is getting uncomfortably close to the typical orbital

periods of galaxies, and so one can legitimately worry about whether the SFR has really been constant over the required timescale. This problem becomes even worse if one looks at small-subregions of galaxies, rather than galaxies as a whole. One also has to worry about stars moving from their birth locations over such long timescales.

### 2.3.6 *Combined Estimators*

As one might guess from the discussion thus far, none of the indicators by itself is particularly good. Recombination lines and UV get into trouble in dusty galaxies because they miss light from young stars that is obscured by dust, while IR gets into trouble because it misses light from young stars that is not dust-obscured. This suggests that the best way to proceed is to combine one or more estimators, and this is indeed the current state of the art. A number of combined indicators are suggested in [Kennicutt & Evans \(2012\)](#).

## **Part II**

# **Physical Processes**



# 3

## *Chemistry and Thermodynamics*

Having completed our whirlwind tour of the observational phenomenology, we will now devote the next four chapters to understanding the physical processes that govern the behavior of the star-forming ISM and its transformation into stars. The goal here is to develop physical intuition for how this gas behaves, and to develop some analytic tools that we can use through the remainder of the course. This part begins with a discussion of the microphysics of the cold ISM.

### *3.1 Chemical Processes in the Cold ISM*

We will begin our discussion of the microphysics of the cold ISM with the goal of understanding something important that should be clear from the observational discussion: the parts of the ISM associated with star formation are overwhelmingly molecular gas. This is in contrast to the bulk of the ISM, at least in the Milky Way and similar galaxies, where the bulk of interstellar matter is composed of atomic or ionized gas with few or no molecules. So why does the ISM in some places turn molecular, and how is this transition associated with star formation? We will focus this discussion on the most important atoms / molecules in the ISM: hydrogen / H<sub>2</sub> and carbon / oxygen / CO.

#### *3.1.1 Hydrogen Chemistry*

Molecular hydrogen is a lower energy state than atomic hydrogen, so an isolated box of hydrogen left for an infinite amount of time will eventually become predominantly molecular. In interstellar space, though, the atomic versus molecular fraction in a gas is determined by a balance between formation and destruction processes.

Atomic hydrogen can turn into molecular hydrogen in the gas phase, but this process is extremely slow. This is ultimately due to

#### **Suggested background reading:**

- Krumholz, M. R. 2014, *Phys. Rep.*, 539, 49, sections 3.1 – 3.2

#### **Suggested literature:**

- Glover, S. C. O., Federrath, C., Mac Low, M.-M., & Klessen, R. S. 2010, *MNRAS*, 404, 2

the symmetry of the hydrogen molecule. To form an  $H_2$  molecule, two H atoms must collide and then undergo a radiative transition that removes enough energy to leave the resulting pair of atoms in a bound state. However, two H atoms that are both in the ground state constitute a symmetric system, as does an  $H_2$  molecule in its ground state. Because both the initial and final states are symmetric, one can immediately show from symmetry considerations that the system cannot emit dipole radiation. Formally, in semi-classical theory, the rate of transitions from a starting state  $\langle \psi_{2H} |$  to a final state  $|\psi_{H_2} \rangle$  is proportional to a matrix element of the form

$$\langle \psi_{2H} | \mathcal{E}\mathbf{r} | \psi_{H_2} \rangle, \quad (3.1)$$

where  $\mathcal{E}\mathbf{r}$  is the dipole radiation operator. However, one can immediately see that if  $\psi_{2H}$  and  $\psi_{H_2}$  are both symmetric, then the inner product is anti-symmetric, and its integral over all space is therefore zero, yielding a transition rate of zero. Transitions are possible only if one continues to the next order of expansion of the radiation field (which, in quantum field theory, constitutes thinking about multi-photon processes), or if one if one considers either starting or final states there are not symmetric (say because one of the H atoms is in an excited state, or the final  $H_2$  molecule is in an excited state). Neither of these routes leads to an appreciable transition rate either: multi-photon processes are suppressed compared to single-photon ones by high powers of the fine structure constant, and the lowest-lying energy states of the  $H_2$  molecule are energetic enough that only a negligible fraction of collisions have enough energy to produce them.

Due to this limitation, the dominant formation process is instead formation on the surfaces of dust grains. In this case the excess energy released by forming the molecule is transferred into vibrations in the dust grain lattice, and there is no need for forbidden photon emission. The rate of  $H_2$  formation by surface catalysis is given by

$$\frac{1}{2} S(T, T_{\text{gr}}) \eta(T_{\text{gr}}) n_{\text{gr}} n_H \sigma_{\text{gr}} v_H. \quad (3.2)$$

Here  $S$  is the probability that a hydrogen molecule that hits a dust grain will stick, which is a function of both the gas temperature and the grain temperature.  $\eta$  is the probability that a grain which sticks will migrate across the grain surface and find another H atom before it is evaporated off the grain surface  $n_{\text{gr}}$  and  $n_H$  are the number densities of grains and hydrogen atoms,  $\sigma_{\text{gr}}$  is the mean cross section for a dust grain, and  $v_H$  is the thermal velocity of the hydrogen atoms.

The last three factors can be estimated reasonably well from observations of dust extinction and gas velocity dispersions, while

the former two have to be determined by laboratory measurements and/or theoretical chemistry calculations. Since I am not a chemist, and the literature on this problem is large and heavily dominated by experimental atomic beam chemists, I will simply report the result: for conditions appropriate to the edges of giant molecular clouds, the formation rate is roughly

$$\mathcal{R}nn_H, \quad (3.3)$$

where  $n_H$  and  $n$  are the number densities of H atoms and H nuclei (in atomic or molecular form), respectively, and  $\mathcal{R} \approx 3 \times 10^{-17}$   $\text{cm}^3 \text{s}^{-1}$  is the rate coefficient. It may be a factor of a few lower in warmer regions where the sticking probability is reduced. This is for Milky Way dust content. If we go to a galaxy with less dust, the rate coefficient is presumably reduced proportionally.

The reverse process, destruction, is mostly due to photo-destruction. As with  $\text{H}_2$  formation, things are somewhat complicated by the symmetry of the  $\text{H}_2$  system. The binding energy of  $\text{H}_2$  in the ground state is only 4.5 eV, but this doesn't mean that 4.5 eV photons can destroy it. A reaction of the form



is forbidden by symmetry for exactly the same reason as its inverse. The reaction can only occur if the  $\text{H}_2$  molecule is in an excited state that thus asymmetric (almost never the case at molecular cloud temperatures), or unless one of the H atoms is left in an excited state, which would require an energy of 14.5 eV. Photons with an energy that high are not generally available, because they can ionize neutral hydrogen and thus all get absorbed before propagating very far.

Instead, the main  $\text{H}_2$  destruction process proceeds in two stages. Hydrogen molecules have a series of excited electronic states with energies of 11.2 – 13.6 eV (corresponding to 912 – 1100 Å) above the ground state, which produce absorption features known as the Lyman and Werner bands. Since these energies exceed the binding energy of the  $\text{H}_2$  molecule (4.5 eV), absorptions into them undergo radiative decay to a ground electronic state that can be unbound. This happens roughly 10–15% of the time, depending on exactly which excited state the absorption is into.

Photons in the LW energy range are produced by hot stars, and the Galaxy is saturated with them, which is why most of the Galaxy's volume is filled with atomic or ionized rather than molecular gas. (There are some galaxies that are mostly molecular, for reasons we will see in a moment.)

If  $E_\nu^*$  is the number density of photons as a function of frequency

$\nu$ , then the destruction rate of  $\text{H}_2$  is

$$\int n_{\text{H}_2} \sigma_{\text{H}_2, \nu} c E_\nu^* f_{\text{diss}, \circ} d\nu, \quad (3.5)$$

where  $n_{\text{H}_2}$  is the molecular hydrogen number density,  $\sigma_{\text{H}_2, \nu}$  is the absorption cross-section at frequency  $\nu$ , and  $f_{\text{diss}, \circ}$  is the dissociation probability when a photon of frequency  $\nu$  is absorbed. The expression inside the integral is just the number of hydrogen molecule targets times the cross-section per target times the number of photons times the relative velocities of the photons and molecules ( $= c$ ) times the probability of dissociation per collision. The integral in frequency goes over the entire LW band, from 912 – 1100 Å.

To understand the circumstances under which  $\text{H}_2$  can form, we can take a simple example. Suppose we have some cloud of gas, which we will treat as a uniform slab, which has a beam of UV radiation shining on its surface. The number density of hydrogen nuclei in the cloud is  $n$ , and the UV radiation field shining on the surface has a photon number density  $E_0^*$ . The photon flux is  $F^* = cE_0^*$ .

As a result of this radiation field, the outer parts of the cloud are atomic hydrogen. However, when a hydrogen molecule absorbs a photon and then re-emits that energy, the energy generally comes out in the form of multiple photons of lower energy, which are no longer able to excite resonant LW transitions. Thus photons are being absorbed as hydrogen forms, and the number of photons penetrating the cloud decreases as one moves further and further into it. Eventually the number of photons drops to near zero, and the gas becomes mostly molecular. This process is known as self-shielding.

We can get a rough estimate of when self-shielding is important by writing down two equations to describe this process. First, let us equate the rates of  $\text{H}_2$  formation and destruction, i.e. assume the cloud is in chemical equilibrium. (This is generally true because the reaction rates go as  $n^2$ , so as long as turbulence produces high density regions, there will be places where the reaction occurs quite fast.) This gives

$$n_H n \mathcal{R} = \int n_{\text{H}_2} \sigma_{\text{H}_2, \nu} c E_\nu^* f_{\text{diss}, \circ} d\nu \approx f_{\text{diss}} \int n_{\text{H}_2} \sigma_{\text{H}_2, \nu} c E_\nu^* d\nu. \quad (3.6)$$

In the second step we have made the approximation that  $f_{\text{diss}}$  is roughly frequency-independent, which is true, since it only varies by factors of less than order unity.

Second, let us write down the equation for photon conservation. This just says that the change in photon number density as we move into the cloud is given by the rate at which collisions with  $\text{H}_2$

molecules remove photons.

$$\frac{dF_\nu^*}{dx} = c \frac{dE_\nu^*}{dx} = -n_{H_2} \sigma_{H_2,\nu} c E_\nu^* \quad (3.7)$$

In principle there should be a creation term at lower frequencies, representing photons absorbed and re-emitted, but we're going to focus on the higher LW frequencies, where there is only photon removal. The term on the right hand side is just the collision rate we calculated before.

Now we can integrate the second equation over frequency over the LW band. This gives

$$\frac{dE^*}{dx} = - \int n_{H_2} \sigma_{H_2,\nu} E_\nu^* d\nu, \quad (3.8)$$

where  $E^*$  is the frequency-integrated photon number density. If we combine this equation with the chemical balance equation, we get

$$\frac{dE^*}{dx} = -\frac{n_H n \mathcal{R}}{c f_{\text{diss}}} \quad (3.9)$$

This just says that the rate at which photons are taken out of the beam is equal to the recombination rate, increased by a factor of  $1/f_{\text{diss}}$  because only  $\sim 1$  in  $10$  absorptions actually have to be balanced by a recombination.

If we make the further approximation that the transition from atomic to molecular hydrogen is sharp, so that  $n_H \approx n$  throughout the atomic layer, and we assume that  $\mathcal{R}$  does not vary with position, then the equation is trivial to integrate. At any depth  $x$  inside the slab,

$$E^*(x) = E_0^* - \frac{n^2 \mathcal{R}}{c f_{\text{diss}}} x. \quad (3.10)$$

The transition to molecular hydrogen occurs where  $E^*$  reaches zero, which is at  $x_{H_2} = c f_{\text{diss}} E_0^* / (n^2 \mathcal{R})$ . The total column of atomic hydrogen is

$$N_H = n x_{H_2} = \frac{c f_{\text{diss}} E_0^*}{n \mathcal{R}} \quad (3.11)$$

It is helpful at this point to put in some numbers. In the Milky Way, the observed interstellar UV field is  $E_0^* = 7.5 \times 10^{-4}$  LW photons  $\text{cm}^{-3}$ , and we'll take  $n = 100 \text{ cm}^{-3}$  as a typical number density in a region where molecules might form. Plugging these in with  $f_{\text{diss}} = 0.1$  and  $\mathcal{R} = 3 \times 10^{-17} \text{ cm}^{-3} \text{ s}^{-1}$  gives  $N_H = 7.5 \times 10^{20}$ , or in terms of mass, a column of  $\Sigma = 8.4 M_\odot \text{ pc}^{-2}$ . More precise calculations give numbers closer to  $2 \times 10^{20} \text{ cm}^{-2}$  for the depth of the shielding layer on one side of a GMC. (Of course a comparable column is required on the other side, too.) Every molecular cloud must be surrounded by an envelope of atomic gas with roughly this column density.

This has important implications. First, this means that molecular clouds with column densities of  $100 M_{\odot} \text{ pc}^{-2}$  in molecules must have  $\sim 10\%$  of their total mass in the form of an atomic shield around them. Second, it explains why most of the Milky Way's ISM in the solar vicinity is not molecular. In the regions outside of molecular clouds, the mean column density is a bit under  $10^{21} \text{ cm}^{-2}$ , so the required shielding column is comparable to the mean column density of the entire atomic disk. Only when the gas clumps together can molecular regions form.

This also explains why other galaxies which have higher column densities also have higher molecular fractions. To take an extreme example, the starburst galaxy Arp 220 has a surface density of a few  $\times 10^4 M_{\odot} \text{ pc}^{-2}$  in its nucleus, and the molecular fraction there is at least 90%, probably more.

### 3.1.2 Carbon / Oxygen Chemistry

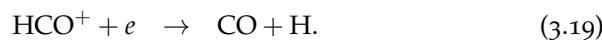
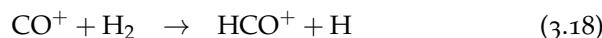
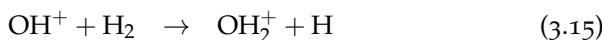
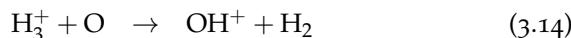
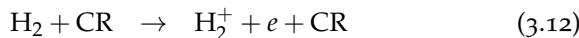
$\text{H}_2$  is the dominant species in molecular regions, but it is very hard to observe directly for the reasons discussed in Chapter 1 – the temperatures are too low for it to be excited. Moreover, as we will discuss shortly,  $\text{H}_2$  is also not the dominant coolant for the same reason. Instead, that role falls to the CO molecule.

Why is CO so important? The main reason is abundances: the most abundant elements in the universe after H and He are O, C, and N, and CO is the simplest (and, under ISM conditions, most energetically favorable) molecule that can be made from them. Moreover, CO can be excited at very low temperatures because its mass is much greater than that of  $\text{H}_2$ , and its dipole moment is weak but non-zero. (A weak dipole moment lowers the energy of radiation emitted, which in turn lowers the temperature needed for excitation.)

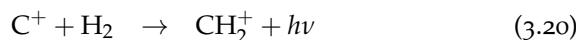
Just as in the bulk of the ISM, hydrogen is mostly H, in the bulk of the ISM the oxygen is mostly O and the carbon is mostly  $\text{C}^+$ . It's  $\text{C}^+$  rather than C because the ionization potential of carbon is less than that of hydrogen, and as a result it tends to be ionized by starlight. (This was / will be covered in the ISM class.) So how do we get from  $\text{C}^+$  and O to CO?

The formation of CO is substantially different than that of  $\text{H}_2$  in that it is dominated by gas-phase rather than grain-surface reactions. Since the temperatures in regions where this reaction is taking place tend to be low, the key processes involve ion-neutral reactions. As those who have taken the diffuse matter class will know (and those who have not yet will learn), these are important because the rate at which they occur is to good approximation independent of temperature, while neutral-neutral reactions.

There are two main pathways to CO. One passes through the OH molecule, and involves a reaction chain that looks like



Here CR indicates cosmic ray. There are also a number of possible variants (e.g., the  $\text{OH}_2^+$  could form  $\text{OH}_3^+$  before proceeding to OH. The second main route is through the CH molecule, where reaction chains tend to follow the general pattern



The rate at which the first reaction chain manufactures CO is limited by the supply of cosmic rays that initiate the production of  $\text{H}_2^+$ , while the rate at which the second reaction chain proceeds is limited by the rate of the final neutral-neutral reaction. Which chain dominates depends on the cosmic ray ionization rate, density, temperature, and similar details. Note that both of these reaction chains require the presence of  $\text{H}_2$ .

CO is destroyed via radiative excitation followed by dissociation in essentially the same manner as  $\text{H}_2$ . The shielding process for CO is slightly different however. As with  $\text{H}_2$ , photons that dissociate CO can be absorbed both by dust grains and by CO molecules. However, due to the much lower abundance of CO compared to  $\text{H}_2$ , the balance between these two processes is quite different than it is for hydrogen, with dust shielding generally the more important of the two. Moreover, there is non-trivial overlap between the resonance lines of CO and those of  $\text{H}_2$ , and thus there can be cross-shielding of CO by  $\text{H}_2$ .

At this point the problem is sufficiently complex that one generally resorts to numerical modeling. The net result is that clouds tend to have a layered structure. In poorly-shielded regions where the FUV has not yet been attenuated, H I and  $\text{C}^+$  dominate. Further in, where the FUV has been partly attenuated,  $\text{H}_2$  and  $\text{C}^+$  dominate. Finally a

transition to H<sub>2</sub> and CO as the dominant chemical states occurs at the center.

For typical Milky Way conditions, the result is that the gas will be mostly CO once the V-band extinction  $A_V$  exceeds 1 – 2 mag. This corresponds to a column density of a few  $\times 10^{21}$  cm<sup>-2</sup>, or  $\sim 20 M_\odot$  pc<sup>-2</sup>, for Milky Way dust. In comparison, recall that typical GMC column densities are  $\sim 10^{22}$  cm<sup>-2</sup>, or  $\sim 100 M_\odot$  pc<sup>-2</sup>. This means that there is a layer of gas where the hydrogen is mostly H<sub>2</sub> and the carbon is still C<sup>+</sup>, but it constitutes no more than a few tens of percent of the mass. However, in galaxies with lower dust to gas ratios, the layer where H<sub>2</sub> dominates but the carbon is not yet mostly CO can be much larger.

### 3.2 Thermodynamics of Molecular Gas

Having discussed the chemistry of molecular gas, we now turn to the problem of its thermodynamics. What controls the temperature of molecular gas? We have already seen that observations imply temperatures that are extremely low,  $\sim 10$  K or even a bit less. How are such cold temperatures achieved? To answer this question, we must investigate what processes heat and cool the molecular ISM.

#### 3.2.1 Heating Processes

The dominant heating process in the atomic ISM is the grain photoelectric effect: photons from stars with energies of  $\sim 8$  – 13.6 eV hit dust grains and eject fast electrons via the photoelectric effect. The fast electrons then thermalize and deposit their energy at heat in the gas. The rate per H nucleus at which this process deposits energy can be written approximately (see the ISM class for justification) as

$$\Gamma_{\text{PE}} \approx 4.0 \times 10^{-26} \chi_{\text{FUV}} Z'_d e^{-\tau_d} \text{ erg s}^{-1} \quad (3.23)$$

where  $\chi_{\text{FUV}}$  is the intensity of the FUV radiation field scaled to its value in the Solar neighborhood,  $Z'_d$  is the dust abundance scaled to the Solar neighborhood value, and  $\tau_d$  is the dust optical depth to FUV photons. The result is, not surprisingly, proportional to the radiation field strength (and thus the number of photons available for heating), the dust abundance (and thus the number of targets for those photons), and the  $e^{-\tau_d}$  factor by which the radiation field is attenuated.

At FUV wavelengths, typical dust opacities are  $\kappa_d \approx 500 \text{ cm}^2 \text{ g}^{-1}$ , so at a typical molecular cloud surface density  $\Sigma \approx 50$  – 100  $M_\odot$  pc<sup>-2</sup>,  $\tau_d \approx 5$  – 10, and thus  $e^{-\tau_d} \approx 10^{-3}$ . Thus in the interiors of molecular clouds, photoelectric heating is strongly suppressed simply because

the FUV photons cannot get in. Typical photoelectric heating rates are therefore of order a few  $\times 10^{-29}$  erg s $^{-1}$  per H atom deep in cloud interiors, though they can obviously be much larger at cloud surfaces or in regions with stronger radiation fields.

We must therefore consider another heating process: cosmic rays. The great advantage of cosmic rays over FUV photons is that, because they are relativistic particles, they have much lower interaction cross sections, and thus are able to penetrate into regions where light cannot. The process of cosmic ray heating works as follows. The first step is the interaction of a cosmic ray with an electron, which knocks the electron off a molecule:



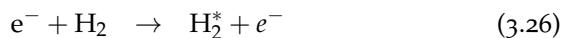
The free electron's energy depends only weakly on the CR's energy, and is typically  $\sim 30$  eV.

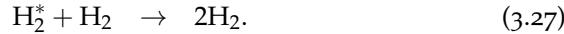
The electron cannot easily transfer its energy to other particles in the gas directly, because its tiny mass guarantees that most collisions are elastic and transfer no energy to the impacted particle. However, the electron also has enough energy to ionize or dissociate other hydrogen molecules, which provides an inelastic reaction that can convert some of its 30 eV to heat. Secondary ionizations do indeed occur, but in this case almost all the energy goes into ionizing the molecule (15.4 eV), and the resulting electron has the same problem as the first one: it cannot effectively transfer energy to the much more massive protons.

Instead, there are a number of other channels that allow electrons to dump their energy into motion of protons, and the problem is deeply messy. The most up to date work on this is Goldsmith et al. (2012, ApJ, 756, 157), and we can very briefly summarize it here. A free electron can turn its energy into heat through three channels. The first is dissociation heating, in which the electron strikes an H<sub>2</sub> molecule and dissociates it:

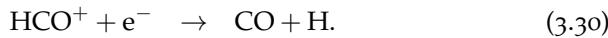
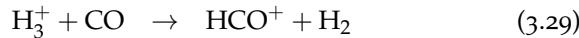


In this reaction any excess energy in the electron beyond what is needed to dissociate the molecule (4.5 eV) goes into kinetic energy of the two recoiling hydrogen atoms, and the atoms, since they are massive, can then efficiently share that energy with the rest of the gas. A second pathway is that an electron can hit a hydrogen molecule and excite it without dissociating it. The hydrogen molecule then collides with another hydrogen molecule and collisionally de-excites, and the excess energy again goes into recoil, where it is efficiently shared. The reaction is





Finally, there is chemical heating, in which the  $\text{H}_2^+$  ion that is created by the cosmic ray undergoes chemical reactions with other molecules that release heat. There are a large number of possible exothermic reaction chains, for example



Each of these reactions is exothermic, and results in heavy ions recoiling at high speed that can efficiently share their energy via collisions. Computing the total energy release requires summing over all these possible reaction chains, which is why the problem is ugly. The final result is that the energy yield per primary cosmic ray ionization is in the range  $\sim 13$  eV under typical molecular cloud conditions, but that it can be several eV higher or lower depending on the local density, electron abundance, and similar variables.

Combining this with the primary ionization rate for cosmic rays in the Milky Way, which is observationally-estimated to be about  $\sim 10^{-16}$  s<sup>-1</sup> per H nucleus in molecular clouds, this gives a total heating rate per H nucleus

$$\Gamma_{\text{CR}} \sim 2 \times 10^{-27} \text{ erg s}^{-1}. \quad (3.31)$$

The heating rate per unit volume is  $\Gamma_{\text{CR}} n$ , where  $n$  is the number density of H nuclei (= 2× the density of H molecules). This is sufficient that, in the interiors of molecular clouds, it generally dominates over the photoelectric heating rate.

### 3.2.2 Cooling Processes

In molecular clouds there are two main cooling processes: molecular lines and dust radiation. Dust can cool the gas efficiently because dust grains are solids, so they are thermal emitters. However, dust is only able to cool the gas if collisions between dust grains and hydrogen molecules occur often enough to keep them thermally well-coupled. Otherwise the grains cool off, but the gas stays hot. The density at which grains and gas become well-coupled is around  $10^4 - 10^5$  cm<sup>-3</sup>, which is higher than the typical density in a GMC, so we won't consider dust cooling further at this point. We'll return to it later when we discuss collapsing objects, where the densities do get high enough for dust cooling to be important.

The remaining cooling process is line emission, and by far the most important molecule for this purpose is CO, for the reasons

stated earlier. The physics is fairly simple. CO molecules are excited by inelastic collisions with hydrogen molecules, and such collisions convert kinetic energy to potential energy within the molecule. If the molecule de-excites radiatively, and the resulting photon escapes the cloud, the cloud loses energy and cools.

Let us make a rough attempt to compute the cooling rate via this process. A diatomic molecule like CO can be excited rotationally, vibrationally, or electronically. At the low temperatures found in molecular clouds, usually only the rotational levels are important. These are characterized by an angular momentum quantum number  $J$ , and each level  $J$  has a single allowed radiative transition to level  $J - 1$ . Larger  $\Delta J$  transitions are strongly suppressed because they require emission of multiple photons to conserve angular momentum.

Unfortunately the CO cooling rate is quite difficult to calculate, because the lower CO lines are all optically thick. A photon emitted from a CO molecule in the  $J = 1$  state is likely to be absorbed by another one in the  $J = 0$  state before it escapes the cloud, and if this happens that emission just moves energy around within the cloud and provides no net cooling. The cooling rate is therefore a complicated function of position within the cloud – near the surface the photons are much more likely to escape, so the cooling rate is much higher than deep in the interior. The velocity dispersion of the cloud also plays a role, since large velocity dispersions Doppler shift the emission over a wider range of frequencies, reducing the probability that any given photon will be resonantly re-absorbed before escaping.

In practice this means that CO cooling rates usually have to be computed numerically, and will depend on the cloud geometry if we want accuracy to better than a factor of  $\sim 2$ . However, we can get a rough idea of the cooling rate from some general considerations. The high  $J$  levels of CO are optically thin, since there are few CO molecules in the  $J - 1$  states capable of absorbing them, so photons they emit can escape from anywhere within the cloud. However, the temperatures required to excite these levels are generally high compared to those found in molecular clouds, so there are few molecules in them, and thus the line emission is weak. Moreover, the high  $J$  levels also have high critical densities, so they tend to be sub-thermally populated, further weakening the emission.

On other hand, low  $J$  levels of CO are the most highly populated, and thus have the highest optical depths. Molecules in these levels produce cooling only if they are within one optical depth the cloud surface. Since this restricts cooling to a small fraction of the cloud volume (typical CO optical depths are many tens for the  $1 \rightarrow 0$  line), this strongly suppresses cooling.

The net effect of combining the suppression of low  $J$  transitions by optical depth effects and of high  $J$  transitions by excitation effects is that cooling tends to be dominated by the single line produced by the lowest  $J$  level for which the line is not optically thick. This line is marginally optically thin, but is kept close to LTE by the interaction of lower levels with the radiation field. Which line this depends on the column density and velocity dispersion of the cloud, but typical peak  $J$  values in Milky Way-like galaxies range from  $J = 2 \rightarrow 1$  to  $J = 5 \rightarrow 4$ .

For an optically thin transition of a quantum rotor where the population is in LTE, the rate of energy emission per H nucleus from transitions between angular momentum quantum numbers  $J$  and  $J - 1$  is given by

$$\Lambda_{J,J-1} = x_{\text{em}} \frac{(2J+1)e^{-E_J/k_B T}}{Z(T)} A_{J,J-1} (E_J - E_{J-1}) \quad (3.32)$$

$$E_J = hB(J+1) \quad (3.33)$$

$$A_{J,J-1} = \frac{512\pi^4 B^3 \mu^2}{3hc^3} \frac{J^4}{2J+1}. \quad (3.34)$$

Here  $x_{\text{em}}$  is the abundance of the emitting species per H nucleus,  $T$  is the gas temperature,  $Z(T)$  is the partition function,  $A_{J,J-1}$  is the Einstein  $A$  coefficient from transitions from state  $J$  to state  $J - 1$ ,  $E_J$  is the energy of state  $J$ ,  $B$  is the rotation constant for the emitting molecule, and  $\mu$  is the electric dipole moment of the emitting molecule. The first equation is simply the statement that the energy loss rate is given by the abundance of emitters multiplied by the fraction of emitters in the  $J$  state in question times the spontaneous emission rate for this state times the energy emitted per transition. Note that there is no explicit density dependence as a result of our assumption that the level with which we are concerned is in LTE. The latter two equations are general results for quantum rotors.

The CO molecule has  $B = 57$  GHz and  $\mu = 0.112$  Debye, and at Solar metallicity its abundance in regions where CO dominates the carbon budget is  $x_{\text{CO}} \approx 1.1 \times 10^{-4}$ . Plugging in these two values, and evaluating for  $J$  in the range 2 – 5, typical cooling rates are of order  $10^{-27} - 10^{-26}$  erg cm $^{-3}$  when the temperature is  $\sim 10$  K. This is why the equilibrium temperatures of molecular clouds are  $\sim 10$  K.

### 3.2.3 Implications

The calculation we have just performed has two critical implications that strongly affect the dynamics of molecular clouds. First, the temperature will be relatively insensitive to variations in the local heating rate. The cosmic ray and photoelectric heating rates are to

good approximation temperature-independent, but the cooling rate is extremely temperature sensitive because, for the dominant cooling lines of CO have level energies are large compared to  $k_B T$ . Examining equation (3.32) would seem to suggest that the cooling rate is exponentially sensitive to temperature. In practice the sensitivity is not quite that great, because which  $J$  dominates changes with temperature, but numerical calculations still show that  $\Lambda_{\text{CO}}$  varies with  $T$  to a power of  $p \sim 2 - 3$ . This means that a factor  $f$  increase in the local heating rate will only change the temperature by a factor  $\sim f^{1/p}$ . Thus we expect molecular clouds to be pretty close to isothermal, except near extremely strong local heating sources.

A second important point is the timescales involved. The gas thermal energy per H nucleus is

$$e \approx \frac{1}{2} \left( \frac{3}{2} kT \right) = 10^{-15} \left( \frac{T}{10 \text{ K}} \right) \text{ erg} \quad (3.35)$$

The factor of 1/2 comes from 2 H nuclei per H<sub>2</sub> molecule, and the equation is only approximate because this neglects quantum mechanical effects that are non-negligible at these low temperatures. However, a correct accounting for these only leads to order unity changes in the result.

The characteristic cooling time is  $t_{\text{cool}} = e / \Lambda_{\text{CO}}$ . Suppose we have gas that is mildly out of equilibrium, say  $T = 20 \text{ K}$  instead of  $T = 10 \text{ K}$ . The heating and cooling are far out of balance, so we can ignore heating completely compared to cooling. At the cooling rate of  $\Lambda_{\text{CO}} = 4 \times 10^{-26} \text{ erg s}^{-1}$  for 20 K gas,  $t_{\text{cool}} = 1.6 \text{ kyr}$ . In contrast, the crossing time for a molecular cloud is  $t_{\text{cr}} = L / \sigma \sim 7 \text{ Myr}$  for  $L = 30 \text{ pc}$  and  $\sigma = 4 \text{ km s}^{-1}$ . The conclusion of this analysis is that radiative effects happen on time scales *much* shorter than mechanical ones. Mechanical effects, such as the heating caused by shocks, simply cannot push the gas any significant way out of radiative equilibrium.



# 4

## Gas Flows and Turbulence

This chapter covers the physics of turbulence in the cold interstellar medium. This will be something of a whirlwind tour, since turbulence is an entire research discipline unto itself. Our goal is to understand the basic statistical techniques used to describe and model interstellar turbulence, so that we will be prepared to apply them in the context of star formation.

### 4.1 Characteristic Numbers for Fluid Flow

#### 4.1.1 The Conservation Equations

To understand the origins of turbulence, both in the ISM and more generally, we start by examining the equations of fluid dynamics and the characteristic numbers that they define. Although the ISM is magnetized, we will first start with the simpler case of an unmagnetized fluid. Fluids are governed by a series of conservation laws. The most basic one is conservation of mass:

$$\frac{\partial}{\partial t} \rho = -\nabla \cdot (\rho \mathbf{v}). \quad (4.1)$$

This equation asserts that the change in mass density at a fixed point is equal to minus the divergence of density times velocity at that point. Physically, this is very intuitive: density at a point changes at a rate that is simply equal to the rate at which mass flows into or out of an infinitesimal volume around that point.

We can write a similar equation for conservation of momentum:

$$\frac{\partial}{\partial t} (\rho \mathbf{v}) = -\nabla \cdot (\rho \mathbf{v} \mathbf{v}) - \nabla P + \rho v \nabla^2 \mathbf{v}. \quad (4.2)$$

Note that the term  $\mathbf{v}\mathbf{v}$  here is a tensor product. This is perhaps more clear if we write things out in index notation:

$$\frac{\partial}{\partial t} (\rho v_i) = -\frac{\partial}{\partial x_j} (\rho v_i v_j) - \frac{\partial}{\partial x_i} P + \rho v \frac{\partial}{\partial x_j} \left( \frac{\partial}{\partial x_j} v_i \right) \quad (4.3)$$

#### Suggested background reading:

- Krumholz, M. R. 2014, *Phys. Rep.*, 539, 49, section 3.3

#### Suggested literature:

- Federrath, C. 2013, *MNRAS*, 436, 1245

The intuitive meaning of this equation can be understood by examining the terms one by one. The term  $\rho\mathbf{v}$  is the density of momentum at a point. The term  $\nabla \cdot (\rho\mathbf{v}\mathbf{v})$  is, in analogy to the equivalent term in the conservation of mass equation, the rate at which momentum is advected into or out of that point by the flow. The term  $\nabla P$  is the rate at which pressure forces acting on the fluid change its momentum. Finally, the last term,  $\rho\nu\nabla^2\mathbf{v}$ , is the rate at which viscosity redistributes momentum; the quantity  $\nu$  is called the kinematic viscosity.

The last term, the viscosity one, requires a bit more discussion. All the other terms in the momentum equation are completely analogous to Newton's second law for single particles. The viscous term, on the other hand, is unique to fluids, and does not have an analog for single particles. It describes the change in fluid momentum due to the diffusion of momentum from adjacent fluid elements. We can understand this intuitively: a fluid is composed of particles moving with random velocities in addition to their overall coherent velocity. If we pick a particular fluid element to follow, we will notice that these random velocities cause some of the particles that make it up diffuse across its boundary to the neighboring element, and some particles from the neighboring element enter. The particles that wander across the boundaries of our fluid element carry momentum with them, and this changes the momentum of the element we're following. The result is that momentum diffuses across the fluid, and this momentum diffusion is called viscosity.

Viscosity is interesting and important because it's the only term in the equation that converts coherent, bulk motion into random, disordered motion. That is to say, the viscosity term is the only one that is dissipative, or that causes the fluid entropy to change.

#### 4.1.2 The Reynolds Number and the Mach Number

To understand the relative importance of terms in the momentum equation, it is helpful to make order of magnitude estimates of their sizes. Let us consider a system of characteristic size  $L$  and characteristic velocity  $V$ ; if we're examining a molecular cloud, we might have  $L \sim 10$  pc and  $V \sim 5$  km s<sup>-1</sup>. The natural time scale for flows in the system is  $L/V$ , so we expect time derivative terms to be of order the thing being differentiated divided by  $L/V$ . Similarly, the natural length scale for spatial derivatives is  $L$ , so we expect spatial derivative terms to be order the quantity being differentiated divided by  $L$ . If we apply these scalings to the momentum equation, we expect the various terms to scale as follows:

$$\frac{\rho V^2}{L} \sim \frac{\rho V^2}{L} + \frac{\rho c_s^2}{L} + \rho \nu \frac{V}{L^2}, \quad (4.4)$$

where  $c_s$  is the gas sound speed, and we have written the pressure as  $P = \rho c_s^2$ . Canceling the common factors, we get

$$1 \sim 1 + \frac{c_s^2}{V^2} + \frac{\nu}{VL} \quad (4.5)$$

From this exercise, we can derive two dimensionless numbers that are going to control the behavior of the equation. We define the Mach number and the Reynolds number as

$$\mathcal{M} \sim \frac{V}{c_s} \quad (4.6)$$

$$\text{Re} \sim \frac{LV}{\nu}. \quad (4.7)$$

The meanings of these dimensionless numbers are fairly clear from the equations. If  $\mathcal{M} \ll 1$ , then  $c_s^2/V^2 \gg 1$ , and this means that the pressure term is important in determining how the fluid evolves. In contrast, if  $\mathcal{M} \gg 1$ , then the pressure term is unimportant for the behavior of the fluid. In a molecular cloud,

$$c_s = \sqrt{\frac{k_B T}{\mu m_H}} = 0.18(T/10\text{ K})^{1/2}, \quad (4.8)$$

where  $\mu = 2.33$  is the mean mass per particle in a gas composed of molecular hydrogen and helium in the usual cosmic abundance ratio of  $1\text{ He}$  per  $10\text{ H}$  atoms. Thus  $\mathcal{M}V/c_s \sim 20$ , and we learn that pressure forces are unimportant.

The Reynolds number is a measure of how important viscous forces are. Viscous forces are significant for  $\text{Re} \sim 1$  or less, and are unimportant of  $\text{Re} \gg 1$ . We can think of the Reynolds number as describing a characteristic length scale  $L \sim \nu/V$  in the flow. This is the length scale on which diffusion causes the flow to dissipate energy. Larger scale motions are effectively dissipationless, while smaller scales ones are damped out by viscosity.

To estimate the Reynolds number in the molecular ISM, we must know the viscosity. For an ideal gas, the kinematic viscosity is  $\nu = 2\bar{u}\lambda$ , where  $\bar{u}$  is the RMS molecular speed (which is of order  $c_s$ ) and  $\lambda$  is the particle mean free-path. The mean free path is of order the inverse of cross-section times density,  $\lambda \sim 1/(\sigma n) \sim [(1\text{ nm})^2(100\text{ cm}^{-3})]^{-1} \sim 10^{12}\text{ cm}$ . Plugging this in then gives  $\nu \sim 10^{16}\text{ cm}^2\text{ s}^{-1}$  and  $\text{Re} \sim 10^9$ . Viscous forces are clearly unimportant in molecular clouds.

The extremely large value of the Reynolds number immediately yields a critical conclusion: molecular clouds must be highly turbulent, because flows with  $\text{Re}$  of more than  $\sim 10^3 - 10^4$  are invariable. Figure 4.1 illustrates this graphically from laboratory experiments.

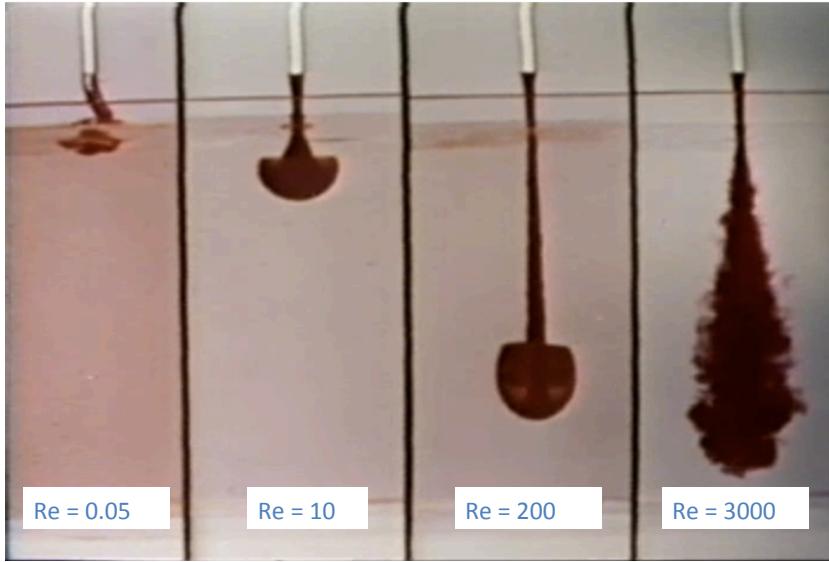


Figure 4.1: Flows at varying Reynolds number  $Re$ . In each panel, a fluid that has been dyed red is injected from the top into the clear fluid on the bottom. The fluids are a glycerin-water mixture, for which the viscosity can be changed by altering the glycerin to water ratio. By changing the viscosity and the injection speed, it is possible to alter the Reynolds number of the injected flow. The frames show how the flow develops as the Reynolds number is varied. This image is a still from the [National Committee for Fluid Mechanics Film Series \(Taylor, 1964\)](#), which, once you get past the distinctly 1960s production values, are a wonderful resource for everything related to fluids.

## 4.2 Modeling Turbulence

We have remarkably little understanding of how turbulence actually works. However, we have developed some simple models and tools to describe it, and we will next explore those.

### 4.2.1 Velocity Statistics

One quantity of interest in a turbulent medium is the structure of the velocity field. How does the velocity change from point to point? In a turbulent medium velocity fluctuates in time and space, and so the best way to proceed is to study those fluctuations statistically. Many statistical tools exist to characterize turbulent motions, and many are used in astrophysics, but we will stick to a few of the simpler ones.

We will also make two simplifying assumptions. First we assume that the turbulence is homogenous, in the sense that the turbulent motions do vary only randomly, and not systematically, with position in the fluid. Second, we assume that it is isotropic, so that turbulent motions do not have a preferred directions. Neither of these are likely to be strictly true in a molecular cloud, particularly the second, since large-scale magnetic fields provide a preferred direction, but we will start with these assumptions and relax them later.

Let  $\mathbf{v}(\mathbf{x})$  be the velocity at position  $\mathbf{x}$ . To characterize how this varies with position, we define the autocorrelation function

$$A(\mathbf{r}) \equiv \frac{1}{V} \int \mathbf{v}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x} + \mathbf{r}) dV \equiv \langle \mathbf{v}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x} + \mathbf{r}) \rangle, \quad (4.9)$$

where the angle brackets indicate an average over all positions  $\mathbf{x}$ . Here,  $A(0) = \langle |\mathbf{v}|^2 \rangle$  is just the RMS velocity in the fluid. If the velocity field is isotropic, then clearly  $A(\mathbf{r})$  cannot depend on the direction, and thus must depend only on  $r = |\mathbf{r}|$ . Thus  $A(r)$  tells us how similar or different the velocity is at some scale  $r$ .

It is often more convenient to think about this in Fourier space than in real space, so rather than the autocorrelation function we often instead think about its Fourier transform. We define the Fourier transform of the velocity field in the usual way, i.e.

$$\tilde{\mathbf{v}}(\mathbf{k}) = \frac{1}{\sqrt{2\pi}} \int \mathbf{v}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}. \quad (4.10)$$

We then define the power spectrum

$$\Psi(\mathbf{k}) \equiv |\tilde{\mathbf{v}}(\mathbf{k})|^2. \quad (4.11)$$

Again, for isotropic turbulence, the power spectrum depends only on the magnitude of the wave number,  $k = |\mathbf{k}|$ , not its direction, so it is more common to talk about the power per unit radius in  $k$ -space,

$$P(k) = 4\pi k^2 \Psi(k). \quad (4.12)$$

This is just the total power integrated over some shell from  $k$  to  $k + dk$  in  $k$ -space. Note that Parseval's theorem tells us that

$$\int P(k) dk = \int |\tilde{\mathbf{v}}(\mathbf{k})|^2 d^3\mathbf{k} = \int \mathbf{v}(\mathbf{x})^2 d^3\mathbf{x}, \quad (4.13)$$

i.e. the integral of the power spectral density over all wavenumbers is equal to the integral of the square of the velocity over all space, so for a flow with constant density (an incompressible flow) the integral of the power spectrum just tells us how much kinetic energy per unit mass there is in the flow. The Wiener-Khinchin theorem also tells us that  $P(\mathbf{k})$  is just the Fourier transform of the autocorrelation function,

$$\Psi(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \int A(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}. \quad (4.14)$$

The power spectrum at a wavenumber  $k$  then just tells us what fraction of the total power is in motions at that wavenumber, i.e. on that characteristic length scale. The power spectrum is another way of looking at the spatial scaling of turbulence. It tells us how much power there is in turbulent motions as a function of wavenumber  $k = 2\pi/\lambda$ . A power spectrum that peaks at low  $k$  means that most of the turbulent power is in large-scale motions, since small  $k$  corresponds to large  $\lambda$ . Conversely, a power spectrum that peaks at high  $k$  means that most of the power is in small-scale motions.

The power spectrum also tells us about the how the velocity dispersion will vary when it is measured over a region of some

characteristic size. Suppose we consider a volume of size  $\ell$ , and measure the velocity dispersion  $\sigma_v(\ell)$  within it. Further suppose that the power spectrum is described by a power law  $P(k) \propto k^{-n}$ . The total kinetic energy per unit within the region is, up to factors of order unity,

$$\text{KE} \sim \sigma_v(\ell)^2, \quad (4.15)$$

but we can also write the kinetic energy per unit mass in terms of the power spectrum, integrating over those modes that are small enough to fit within the volume under consideration:

$$\text{KE} \sim \int_{2\pi/\ell}^{\infty} P(k) dk \propto \ell^{n-1} \quad (4.16)$$

It therefore follows immediately that

$$\sigma_v = c_s \left( \frac{\ell}{\ell_s} \right)^{(n-1)/2}, \quad (4.17)$$

where we have normalized the relationship by defining the sonic scale  $\ell_s$  as the size of a region within which the velocity dispersion is equal to the thermal sound speed of the gas.

#### 4.2.2 The Kolmogorov Model and Turbulence Cascades

The closest thing we have to a model of turbulence is in the case of subsonic, hydrodynamic turbulence; the basic theory for that goes back to [Kolmogorov \(1941\)](#). Real interstellar clouds are neither subsonic nor hydrodynamic (as opposed to magnetohydrodynamic), but this theory is still useful for understanding how turbulence works.

Kolmogorov's theory of turbulence begins with the realization that turbulence is a phenomenon that occurs when  $\text{Re}$  is large, so that there is a large range of scales where dissipation is unimportant. It is possible to show by Fourier transforming the Navier-Stokes equation that for incompressible motion transfer of energy can only occur between adjacent wavenumbers. Energy at a length scale  $k$  cannot be transferred directly to some scale  $k' \ll k$ . Instead, it must cascade through intermediate scales between  $k$  and  $k'$ .

This gives a simple picture of how energy dissipates in fluids. Energy is injected into a system on some large scale that is dissipationless, and it cascades down to smaller scales until it reaches a small enough scale for  $\text{Re} \sim 1$ , at which point dissipation becomes significant. In this picture, if the turbulence is in statistical equilibrium, such that is neither getting stronger or weaker, the energy at some scale  $k$  should depend only on  $k$  and on the rate of injection or dissipation (which are equal)  $\psi$ .

A translation of [Kolmogorov \(1941\)](#) (which is in Russian) can be found in [Kolmogorov \(1991\)](#).

This allows us to make the following clever dimensional argument.  $k$  has units of  $1/L$ , i.e. one over length. The power spectrum has units of energy per unit mass per unit radius in  $k$ -space. The energy per unit mass is like a velocity squared, so it has units  $L^2/T^2$ , and this is divided by  $k$ , so  $P(k)$  has units of  $L^3/T^2$ . The injection / dissipation rate  $\psi$  has units of energy per unit mass per unit time, which is a velocity squared divided by a time, or  $L^2/T^3$ .

Since  $P(k)$  is a function only of  $k$  and  $\psi$ , we can write  $P(k) = Ck^\alpha\psi^\beta$  for some dimensionless constant  $C$ , then by dimensional analysis we have

$$\frac{L^3}{T^2} \sim L^{-\alpha} \left( \frac{L^2}{T^3} \right)^\beta \quad (4.18)$$

$$L^3 \sim L^{-\alpha+2\beta} \quad (4.19)$$

$$T^{-2} \sim T^{-3\beta} \quad (4.20)$$

$$\beta = \frac{2}{3} \quad (4.21)$$

$$\alpha = 2\beta - 3 = -\frac{5}{3} \quad (4.22)$$

This immediately tell us three critical things. First, the power in the flow varies with energy injection rate to the  $2/3$ . Second, this power is distributed such that the power at a given wavenumber  $k$  varies as  $k^{-5/3}$ . This means that most of the power is in the largest scale motions, since power diminishes as  $k$  increases. Third, if we now take this spectral slope and use it to derive the scale-dependent velocity dispersion from equation (4.17), we find that  $\sigma_v \propto \ell^{1/3}$ , i.e., velocity dispersion increase with size scale as size to the  $1/3$  power. This is an example of what is known in observational astronomy as a linewidth-size relation – linewidth because the observational diagnostic we use to characterize velocity dispersion is the width of a line. This relationship tells us that larger regions should have larger linewidths, with the linewidth scaling as the  $1/3$  power of size in the subsonic regime.

The subsonic regime can be tested experimentally on Earth, and Kolmogorov's model provides an excellent fit to observations. Figure 4.2 shows one example.

### 4.3 Supersonic Turbulence

#### 4.3.1 Velocity Statistics

We have seen that real interstellar clouds not only have  $\text{Re} \gg 1$ , they also have  $\mathcal{M} \gg 1$ , and so the flows within them are supersonic. This means that pressure is unimportant on size scales  $L \gg \ell_s$ . Since

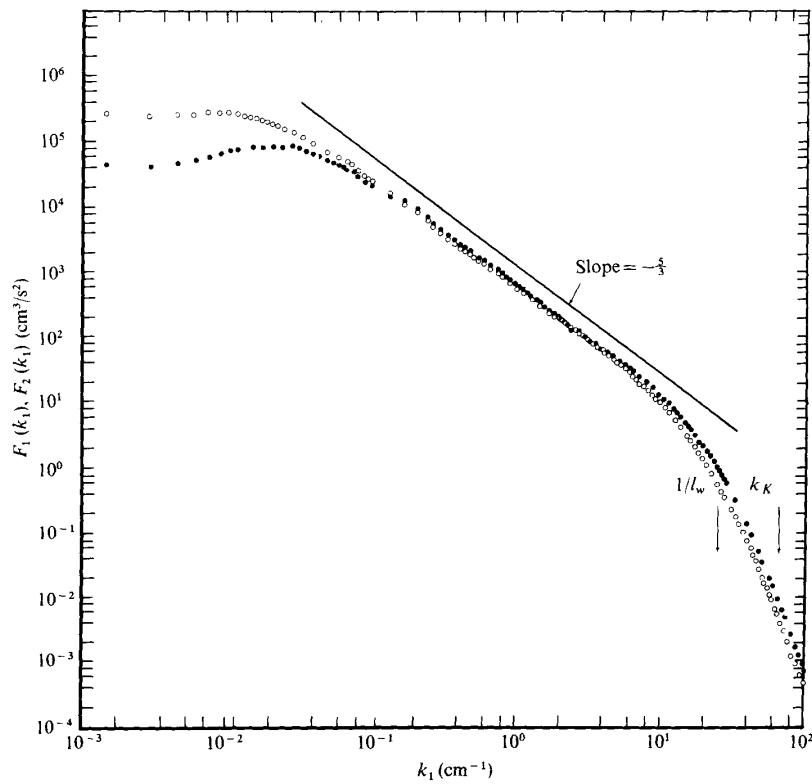


Figure 4.2: An experimentally-measured power spectrum for turbulence generated by an air jet (Champagne, 1978). The  $x$  axis is the wavenumber, and the open and filled points show the velocity power spectrum for the velocity components parallel and transverse to the stream, respectively.

viscosity is also unimportant on large scales (since  $\text{Re} \gg 1$ ), this means that gas tends to move ballistically on large scales.

On small scales this will produce very sharp gradients in velocity, since fast-moving volumes of fluid will simply overtake slower ones. Since the viscosity term gets more important on smaller scales, the viscosity term will eventually stop the fluid from moving ballistically. In practice this means the formation of shocks – regions where the flow velocity changes very rapidly, on a size scale determined by the viscosity.

We expect that the velocity field that results in this case will look like a series of step functions. The power spectrum of a step function is a power law  $P(k) \propto k^{-2}$ . One can establish this easily from direct calculation. Let's zoom in on the region around a shock, so that the change in velocity on either side of the shock is small. The Fourier transform of  $v$  in 1D is

$$\tilde{v}(k) = \frac{1}{\sqrt{2\pi}} \int v(x) e^{-ikx} dx \quad (4.23)$$

The periodic function vanishes for all periods in the regions where  $v$  is constant. It is non-zero only in the period that includes the shock. The amplitude of  $\int v(x) e^{-ikx} dx$  during that period is simply proportional to the length of the period, i.e. to  $1/k$ . Thus,  $\tilde{v}(k) \propto 1/k$ . It then follows that  $P(k) \propto k^{-2}$  for a single shock. An isotropic system of overlapping shocks should therefore also look approximately like a power law of similar slope. This gives a velocity dispersion versus size scale  $\sigma_v \propto \ell^{1/2}$ , and this is exactly what is observed. Figure 4.3 shows an example.

Note that, although the power spectrum is only slightly different than that of subsonic turbulence ( $-5/3$  versus about  $-2$ ), there is really an important fundamental difference between the two regimes. Most basically, in Kolmogorov turbulence decay of energy happens via a cascade from large to small scales, until a dissipative scale is reached. In the supersonic case, on the other hand, the decay of energy is via the formation of shocks, and as we have just seen a single shock generates a power spectrum  $\propto k^{-2}$ , i.e. it non-locally couples many scales. Thus, in supersonic turbulence there is no locality in  $k$ -space. All scales are coupled at shocks.

### 4.3.2 Density Statistics

In subsonic flows the pressure force is dominant, and so if the gas is isothermal, then the density stays nearly constant – any density inhomogeneities are ironed out immediately by the strong pressure forces. In supersonic turbulence, on the other hand, the flow is highly compressible. It is therefore of great interest to ask about the statistics

In real interstellar clouds the relevant viscosity is the magnetic one, as we shall see in Chapter 5.

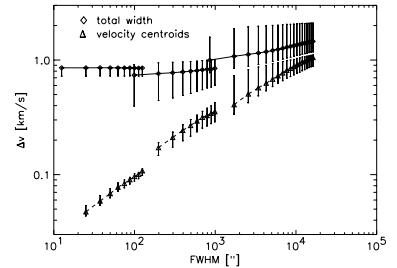


Figure 4.3: Linewidth versus size in the Polaris Flare Cloud obtained from CO observations (Ossenkopf & Mac Low, 2002). Diamonds show the total measured velocity width within apertures of the size indicated on the  $x$  axis, while triangles show the dispersion obtained by taking the centroid velocity in each pixel and measuring the dispersion of centroids. The three sets of points joined by lines represent measurements from three separate telescopes.

of the density field.

Numerical experiments and empirical arguments (but not fully rigorous proofs) indicate that the density field for a supersonically-turbulent, isothermal medium is well-described by a lognormal distribution,

$$p(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{(s-s_0)^2}{2\sigma_s^2}\right], \quad (4.24)$$

where  $s = \ln(\rho/\bar{\rho})$  is the log of the density normalized to the mean density  $\bar{\rho}$ . This distribution describes the probability that the density at a randomly chosen point will be such that  $\ln(\rho/\bar{\rho})$  is in the range from  $s$  to  $s + ds$ . The median of the distribution  $s_0$  and the dispersion  $\sigma_s$  must be related to one another, because we require that

$$\bar{\rho} = \int p(s)\rho ds. \quad (4.25)$$

With a bit of algebra, one can show that this equation is satisfied if and only if

$$s_0 = -\sigma_s^2/2. \quad (4.26)$$

Instead of computing the probability that a randomly chosen point in space will have a particular density, we can also compute the probability that a randomly chosen mass element will have a particular density. This more or less amounts to a simple change of variables. Consider some volume of interest  $V$ , and examine all the material with density such that  $\ln(\rho/\bar{\rho})$  is in the range from  $s$  to  $s + ds$ . This material occupies a volume  $dV = p(s)V$ , and thus must have a mass

$$dM = \rho p(s) dV \quad (4.27)$$

$$= \bar{\rho} e^s \cdot \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{(s-s_0)^2}{2\sigma_s^2}\right] dV \quad (4.28)$$

$$= \bar{\rho} \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{(s+s_0)^2}{2\sigma_s^2}\right] dV \quad (4.29)$$

It immediately follows that the mass PDF is simply

$$p_M(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{(s+s_0)^2}{2\sigma_s^2}\right], \quad (4.30)$$

i.e., exactly the same as the volume PDF but with the peak moved from  $-s_0$  to  $+s_0$ . Physically, the meaning of these shifts is that the typical volume element in a supersonic turbulent field is at a density lower than the mean, because much of the mass is collected into shocks. The typical mass element lives in one of these shocked regions, and thus is at higher-than-average density. Figure 4.4 shows an example of the density distribution produced in a numerical simulation of supersonic turbulence.

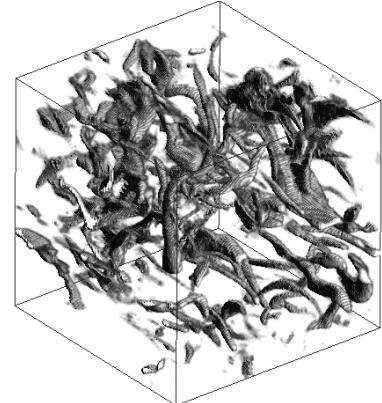


Figure 4.4: Volume rendering of the density field in a simulation of supersonic turbulence (Padoan & Nordlund, 1999). The surfaces shown are isosurfaces of density.

The lognormal functional form is not too surprising, given the central limit theorem. Supersonic turbulence consists of an alternative series of shocks, which cause the density to be multiplied by some factor, and supersonic rarefactions, which cause it to drop by some factor. The result of multiplying a lot of positive density increases by a lot of negative density drops at random tends to produce a normal distribution in the multiplicative factor, and thus a lognormal distribution in the density.

This argument does not, however, tell us about the dispersion of densities, which must be determined empirically from numerical simulations. The general result of these simulations (e.g., Federrath, 2013) is that

$$\sigma_s^2 \approx \ln \left( 1 + b^2 \mathcal{M}^2 \frac{\beta_0}{\beta_0 + 1} \right), \quad (4.31)$$

where the factor  $b$  is a number in the range  $1/3 - 1$  that depends on how compressive versus solenoidal the velocity field is, and  $\beta_0$  is the ratio of thermal to magnetic pressure at the mean density and magnetic field strength – we'll get to the magnetic case next.

In addition to the density PDF, there are higher order statistics describing correlations of the density field from point to point. We will defer a discussion of these until we get to models of the IMF in Chapter ??, where they play a major role.



# Problem Set 1

## 1. Molecular Tracers.

Here we will derive a definition of the critical density, and use it to compute some critical densities for important molecular transitions. For the purposes of this problem, you will need to know some basic parameters (such as energy levels and Einstein coefficients) of common interstellar molecules. You can obtain these from the Leiden Atomic and Molecular Database (LAMDA, <http://www.strw.leidenuniv.nl/~moldata>). It's also worth taking a quick look through the associated paper (Schöier et al., 2005)<sup>1</sup> so you get a feel for where these numbers come from.

- (a) Consider an excited state  $i$  of some molecule, and let  $A_{ij}$  and  $k_{ij}$  be the Einstein  $A$  coefficient and the collision rate, respectively, for transitions from state  $i$  to state  $j$ . Write down expressions for the rates of radiative and collisional de-excitations out of state  $i$  in a gas where the number density of collision partners is  $n$ .
- (b) We define the critical density  $n_{\text{crit}}$  of a state as the density for which the radiative and collisional de-excitation rates are equal.<sup>2</sup> Using your answer to the previous part, derive an expression for  $n_{\text{crit}}$  in terms of the Einstein coefficient and collision rates for the state.
- (c) When a state has a single downward transition that is far more common than any other one, as is the case for example for the rotational excitation levels of CO, it is common to refer to the critical density of the upper state of the transition as the critical density of the line. Compute critical densities for the following lines: CO  $J = 1 \rightarrow 0$ , CO  $J = 3 \rightarrow 2$ , CO  $J = 5 \rightarrow 4$ , and HCN  $J = 1 \rightarrow 0$ , using H<sub>2</sub> as a collision partner. Perform your calculation for the most common isotopes: <sup>12</sup>C, <sup>16</sup>O, and <sup>14</sup>N. Assume the gas temperature is 10 K, the H<sub>2</sub> molecules are all para-H<sub>2</sub>, and neglect hyperfine splitting.
- (d) Consider a molecular cloud in which the volume-averaged density is  $n = 100 \text{ cm}^{-3}$ . Assuming the cloud has a lognor-

<sup>1</sup> Schöier et. al, 2005, A&A, 432, 369

<sup>2</sup> There is some ambiguity in this definition. Some people define the critical density as the density for which the rate of radiative de-excitation equals the rate of *all* collisional transitions out of a state, not just the rate of collisional de-excitations out of it. In practice this usually makes little difference.

mal density distribution as given by equation (4.24), with a dispersion  $\sigma_s^2 = 5.0$ , compute the fraction of the cloud mass that is denser than the critical density for each of these transitions. Which transitions are good tracers of the bulk of the mass in a cloud? Which are good tracers of the denser, and thus presumably more actively star-forming, parts of the cloud?

## 2. Infrared Luminosity as a Star Formation Rate Tracer.

We use a variety of indirect indicators to measure the star formation rate in galaxies, and one of the most common is to measure the galaxy's infrared luminosity. The underlying assumptions behind this method are that (1) most of the total radiant output in the galaxy comes from young, recently formed stars, and (2) that in a sufficiently dusty galaxy most of the starlight will be absorbed by dust grains within the galaxy and then re-radiated in the infrared. We will explore how well this conversion works using the popular stellar population synthesis package Starburst99 (Leitherer et al., 1999; Vázquez & Leitherer, 2005), <http://www.stsci.edu/~science/starburst99>.

- (a) Once you've read enough of the papers to figure out what Starburst99 does, use it with the default parameters to compute the total luminosity of a stellar population in which star formation occurs continuously at a fixed rate  $\dot{M}_*$ . What is the ratio of  $L_{\text{tot}}/\dot{M}_*$  after 10 Myr? After 100 Myr? After 1 Gyr? Compare these ratios to the conversion factor between  $L_{\text{TIR}}$  and  $\dot{M}_*$  given in Table 1 of Kennicutt & Evans (2012)<sup>3</sup>.
- (b) Plot  $L_{\text{tot}}/\dot{M}_*$  as a function of time for this population. Based on this plot, how old does a stellar population have to be before  $L_{\text{TIR}}$  becomes a good tracer of the total star formation rate?
- (c) Try making the IMF slightly top-heavy, by removing all stars below  $0.5 M_\odot$ . How much does the luminosity change for a fixed star formation rate? What do you infer from this about how sensitive this technique is to assumptions about the form of the IMF?

<sup>3</sup> Kennicutt & Evans, 2012, ARA&A, 50, 531

# 5

## *Magnetic Fields and Magnetized Turbulence*

In our treatment of fluid flow and turbulence in Chapter 4, we concentrated on the hydrodynamic case. However, real star-forming clouds are highly magnetized. We therefore devote this chapter to the question of how magnetic fields change the nature of molecular cloud fluid flow.

### *5.1 Observing Magnetic Fields*

#### *5.1.1 Zeeman Measurements*

How do we even know that magnetic field are present? There are several methods that can be used to measure magnetic fields, but the most direct is the Zeeman effect. The Zeeman effect is a slight shift in energy levels of an atom or molecule in the presence of a magnetic field. Ordinarily the energies of a level depend only the direction of the electron spin relative to its orbital angular momentum vector, not on the direction of the net angular momentum vector. However, in the presence of an external magnetic field, states with different orientations of the net angular momentum vector of the atom have slightly different energies due to the interaction of the electron magnetic moment with the external field. This causes a normally single spectral line produced transitions from that level to split into several separate lines at slightly different frequencies.

For the molecules with which we are concerned, the level is normally split into three sublevels – one at slightly higher frequency than the unperturbed line, one at slightly lower frequency, and one at the same frequency. The strength of this splitting varies depending on the electronic configuration of the atom or molecule in question. For OH, for example, the splitting is  $Z = 0.98 \text{ Hz}/\mu\text{G}$ , where the parameter  $Z$  is called the Zeeman sensitivity, and the shift is  $\Delta\nu = BZ$ , where  $B$  is the magnetic field strength. One generally wants to look for molecules where  $Z$  is as large as possible, and these are generally

#### **Suggested background reading:**

- Crutcher, R. M. 2012, ARA&A, 50, 29

#### **Suggested literature:**

- Li, P. S., McKee, C. F., Klein, R. I., & Fisher, R. T. 2008, ApJ, 684, 380

molecules or atoms that have an unpaired electron in their outer shell. Examples include atomic hydrogen, OH, CN, CH, CCS, SO, and O<sub>2</sub>.

The Doppler width of the line is  $\sigma_v = \nu_0(\sigma_v/c)$ , where for the relevant OH transition  $\nu_0 = 1.667$  GHz. If the OH molecule has a velocity dispersion of order 0.1 km s<sup>-1</sup>, as expected for the lowest observed velocity dispersion even on small scales in molecular clouds, then  $(\sigma_v/c) \sim 10^{-6}$ , so  $\sigma_v \sim 1$  kHz. This means that, unless the field is considerably larger than 1000 μG (1 mG), which it essentially never is, the Zeeman splitting is smaller than the Doppler line width, and we won't see the line split.

However, there is a trick to avoid this problem: radiation from the different Zeeman sublevels has different polarization. If the magnetic field is along the direction of propagation of the radiation, emission from the higher frequency Zeeman sublevel is right circularly polarized, while radiation from the lower frequency level is left circularly polarized. The unperturbed level is unpolarized. Thus although one cannot see the line split if one looks at total intensity (as measured by the Stokes  $I$  parameter), one can see that the different polarization components peak at slightly different frequencies, so that the circularly polarized spectrum (as measured by the Stokes  $V$  parameter) looks different than the total intensity spectrum. One can deduce the magnetic field strength along the line of sight from the difference between the total and polarized signals. Figure 5.1 shows a sample detection.

Applying this technique to molecular line emission from molecular clouds indicates that they are threaded by magnetic fields whose strengths range from tens to thousands of μG, with higher density gas generally showing stronger fields. We can attempt to determine if this is dynamically important by a simple energy argument.

For a low-density envelope of a GMC with  $n \sim 100$  cm<sup>-3</sup> ( $\rho \sim 10^{-22}$  g cm<sup>-3</sup>), we might have  $v$  of a few km s<sup>-1</sup>, giving a kinetic energy density

$$E_K \sim \rho v^2 \sim 10^{-22} \text{ g cm}^{-3} \times (3 \times 10^5 \text{ cm s}^{-1})^2 \sim 10^{-11} \text{ erg cm}^{-3}. \quad (5.1)$$

The energy density in a magnetic field is

$$E_B = \frac{B^2}{8\pi} \sim \frac{(10 \mu\text{G})^2}{8\pi} \sim 10^{-11} \text{ erg cm}^{-3}. \quad (5.2)$$

Thus the magnetic energy density is comparable to the kinetic energy density, and is dynamically significant in the flow.

### 5.1.2 The Chandrasekhar-Fermi Method

While the Zeeman effect provides by far the most direct method of measuring magnetic field strengths, it is not the only method

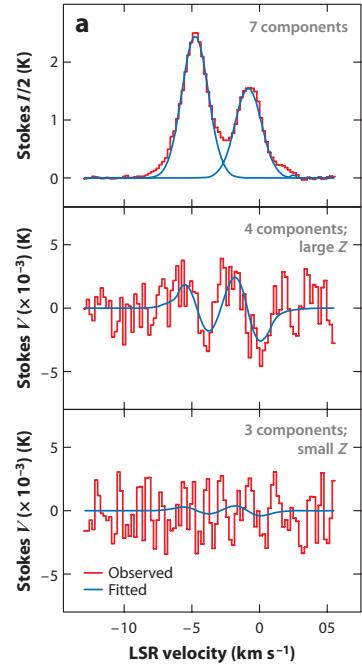


Figure 5.1: Sample Zeeman detection of an interstellar magnetic field using the CN line in the region DR21(OH) (Crutcher, 2012). The top panel shows the observed total intensity (Stokes  $I$ , red lines), which is well-fit by two different velocity components (blue lines). The CN molecule has 7 hyperfine components, of which 4 have a large Zeeman splitting and 3 have a small splitting. The middle panel shows the measured Stokes  $V$  (circularly polarized emission) for the sum of the 4 strong splitting components, while the bottom panel shows the corresponding measurement for the 3 weak components. The blue lines show the best fit, with the line of sight magnetic field as the fitting parameter.

for making this measurement. Another commonly-used technique, which we will not discuss in any detail, is the Chandrasekhar-Fermi method (Chandrasekhar & Fermi, 1953). This method relies on the fact that interstellar dust grains are non-spherical, which has two important implications. First, a non-spherical grain acts like an antenna, in that it interacts differently with electromagnetic waves that are oriented parallel and perpendicular to its long axis. As a result, grains both absorb and emit light preferentially along their long axis. This would not matter if the orientations of grains in the interstellar medium were random. However, there is a second effect. Most grains are charged, and as a result they tend to become preferentially aligned with the local magnetic field. The combination of these two effects means that the dust in a particular region of the ISM characterized by a particular large scale field will produce a net linear polarization in both the light it emits and any light passing through it. The direction of the polarization then reveals the orientation of the magnetic field on the plane of the sky.

By itself this effect tells us nothing about the strength of the field – in principle there should be some relationship between field strength and degree of dust polarization, but there are enough other compounding factors and uncertainties that we cannot with any confidence translate the observed degree of polarization into a field strength. However, if we have measurements of the field orientation as a function of position, then we can estimate the field strength from the morphology of the field. As we shall see below, the degree to which field lines are straight or bent is strongly correlated with the ratio of magnetic energy density to turbulent energy density, and so the degree of alignment becomes a diagnostic of this ratio. In fact, one can even attempt to make quantitative field strength estimates from this method, albeit with very large uncertainties.

## 5.2 Equations and Characteristic Numbers for Magnetized Turbulence

Now that we know that magnetic fields are present, let us discuss some basic theory for magnetized flow. To understand how magnetic fields affect the flows in molecular clouds, it is helpful to write down the fundamental evolution equation for the magnetic field in a plasma (this is derived in many places – my notation and discussion follow Shu 1992):

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}) = -\nabla \times (\eta \nabla \times \mathbf{B}) \quad (5.3)$$

Here  $\mathbf{B}$  is the magnetic field,  $\mathbf{v}$  is the fluid velocity (understood to be the velocity of the ions, which carry all the mass), and  $\eta$  is the

electrical resistivity. If the resistivity is constant in space, we can use the fact that  $\nabla \cdot \mathbf{B} = 0$  to simplify this slightly to get

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}) = \eta \nabla^2 \mathbf{B}. \quad (5.4)$$

The last term here looks very much like the  $\nu \nabla^2 \mathbf{v}$  term we had in the momentum equation to describe viscosity. That term described diffusion of momentum, while the one in this equation describes diffusion of the magnetic field.

(Note that we're simplifying a bit here – the real dissipation mechanism in molecular clouds is not simple resistivity, it is something more complex called ambipolar drift, which we'll discuss in more detail later. However, the qualitative point we can make is the same, and the algebra is simpler if we use a simple scalar resistivity.)

We can understand the implications of this equation using dimensional analysis much as we did for the momentum equation. Again, we let  $L$  be the characteristic size of the system and  $V$  be the characteristic velocity, so  $L/V$  is the characteristic timescale. We let  $B$  be the characteristic magnetic field strength. Inserting the same scalings as before, the terms vary as

$$\frac{BV}{L} + \frac{BV}{L} \sim \eta \frac{B}{L^2} \quad (5.5)$$

$$1 \sim \frac{\eta}{VL} \quad (5.6)$$

In analogy to the ordinary hydrodynamic Reynolds number, we define the magnetic Reynolds number by

$$Rm = \frac{LV}{\eta}. \quad (5.7)$$

Magnetic diffusion is significant only if  $Rm \sim 1$  or smaller.

What is  $Rm$  for a typical molecular cloud? As in the hydrodynamic case, we can take  $L$  to be a few tens of pc and  $V$  to be a few  $\text{km s}^{-1}$ . The magnetic field  $B$  is a few tens of  $\mu\text{G}$ . The electrical resistivity is a microphysical property of the plasma, which, for a weakly ionized plasma, depends on the ionization fraction in the gas and the ion-neutral collision rate. Its typical value for the molecular cloud example we've been using, which we will calculate in a bit, is  $\eta \sim 10^{22} - 10^{23} \text{ cm}^2 \text{ s}^{-1}$ . Since, as we discussed earlier,  $LV \sim 10^{25} \text{ cm}^2 \text{ s}^{-1}$ , this implies that the  $Rm$  for molecular clouds is hundreds to thousands.

Again in analogy to hydrodynamics, this means that on large scales magnetic diffusion is unimportant for molecular clouds – although it is important on smaller scales. The significance of a large value of  $Rm$  becomes clear if we write down the induction equation

with  $\eta = 0$  exactly:

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}) = 0. \quad (5.8)$$

To understand what this equation implies, it is useful consider the magnetic flux  $\Phi$  threading some fluid element. We define this as

$$\Phi = \int_A \mathbf{B} \cdot \hat{\mathbf{n}} dA, \quad (5.9)$$

where we integrate over some area  $A$  that defines the fluid element. Using Stokes's theorem, we can alternately write this as

$$\Phi = \oint_C \mathbf{B} \cdot d\mathbf{l}, \quad (5.10)$$

where  $C$  is the curve that bounds  $A$ . The time derivative of this is then

$$\frac{d\Phi}{dt} = \int_A \frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}} dA + \oint_C \mathbf{B} \cdot \mathbf{v} \times d\mathbf{l} \quad (5.11)$$

$$= \int_A \frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}} dA + \oint_C \mathbf{B} \times \mathbf{v} \cdot d\mathbf{l} \quad (5.12)$$

Here the second term on the right comes from the fact that, if the fluid is moving at velocity  $\mathbf{v}$ , the area swept out by a unit  $d\mathbf{l}$  per unit time is  $\mathbf{v} \times d\mathbf{l}$ , so the flux crossing this area is  $\mathbf{B} \cdot \mathbf{v} \times d\mathbf{l}$ . Then in the second step we used the fact that  $\nabla \cdot \mathbf{B} = 0$  to exchange the dot and cross products.

If we now apply Stokes theorem again to the second term, we get

$$\frac{d\Phi}{dt} = \int_A \frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}} dA + \int_A \nabla \times (\mathbf{B} \times \mathbf{v}) \cdot \hat{\mathbf{n}} dA \quad (5.13)$$

$$= \int_A \left[ \frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}) \right] \cdot \hat{\mathbf{n}} dA \quad (5.14)$$

$$= 0 \quad (5.15)$$

The meaning of this is that, when  $Rm$  is large, the magnetic flux through each fluid element is conserved. This is called flux-freezing, since we can envision it geometrically as saying that magnetic field lines are frozen into the fluid, and move along with it.

Thus on large scales the magnetic field moves with the fluid. However, on smaller scales the magnetic Reynolds number is  $\sim 1$ , and the field lines are not tied to the gas. We will calculate this scale in a bit. Before that, however, we want to calculate another important dimensionless number describing the MHD flows in molecular clouds.

The momentum equation including magnetic forces is

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) = -\nabla \cdot (\rho \mathbf{v} \mathbf{v}) - \nabla P + \rho v \nabla^2 \mathbf{v} + \frac{1}{4\pi}(\nabla \times \mathbf{B}) \times \mathbf{B}, \quad (5.16)$$

and if we make order of magnitude estimates of the various terms in this, we have

$$\frac{\rho V^2}{L} \sim -\frac{\rho V^2}{L} + \frac{\rho c_s^2}{L} + \frac{\rho \nu V}{L^2} + \frac{B^2}{L} \quad (5.17)$$

$$1 \sim 1 + \frac{c_s^2}{V^2} + \frac{\nu}{VL} + \frac{B^2}{\rho V^2} \quad (5.18)$$

The second and third terms on the right hand side we have already defined in terms of  $\mathcal{M} = V/c_s$  and  $\text{Re} = LV/\nu$ . We now define our fourth and final characteristic number,

$$\mathcal{M}_A \equiv \frac{V}{v_A}, \quad (5.19)$$

where

$$v_A = \frac{B}{\sqrt{4\pi\rho}} \quad (5.20)$$

is the Alfvén speed – the speed of the wave that, in magnetohydrodynamics, plays a role somewhat analogous to the sound wave in hydrodynamics. In flows with  $\mathcal{M}_A \gg 1$ , the magnetic force term is unimportant, while in those with  $\mathcal{M}_A \ll 1$  it is dominant.

Using our characteristic numbers  $n \sim 100 \text{ cm}^{-3}$ ,  $B$  of a few tens of  $\mu\text{G}$ , and  $V$  of a few  $\text{km s}^{-1}$ , we see that  $v_A$  is of order a few  $\text{km s}^{-1}$ , about the same as the velocity. Thus the flows in molecular clouds are highly supersonic ( $\mathcal{M} \gg 1$ ), but only trans-Alfvénic ( $\mathcal{M}_A \sim 1$ ), and magnetic forces have a significant influence. Simulations of turbulence with magnetic fields confirm this, as shown in Figure 5.2.

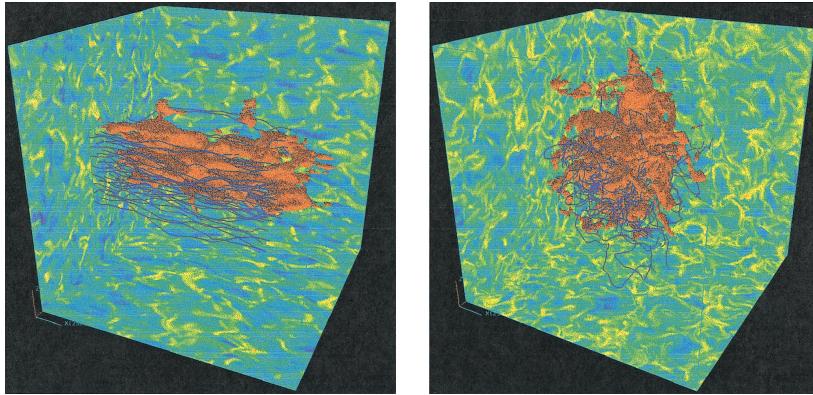


Figure 5.2: Simulations of Alfvénic (left) and sub-Alfvénic (right) turbulence. Colors on the cube surface are slices of the logarithm of density, blue lines are magnetic field lines, and red surfaces are isodensity surfaces for a passive contaminant added to the flow.

### 5.3 Non-Ideal Magnetohydrodynamics

We have just shown that the magnetic Reynolds number is a critical parameter for magnetized turbulence, and that this depends on the

resistivity  $\eta$ . In the final part of this Chapter we will discuss in a bit more detail the physical origins of resistivity and related effects.

### 5.3.1 Ion-Neutral Drift

Molecular clouds are not very good plasmas. Most of the gas in a molecular cloud is neutral, not ionized. The ion fraction may be  $10^{-6}$  or lower. Since only ions and electrons can feel the Lorentz force directly, this means that fields only exert forces on most of the particles in a molecular cloud indirectly. The indirect mechanism is that the magnetic field exerts forces on the ions and electrons (and mostly ions matter for this purpose), and these then collide with the neutrals, transmitting the magnetic force.

If the collisional coupling is sufficiently strong, then the gas acts like a perfect plasma. However, when the ion fraction is very low, the coupling is imperfect, and ions and neutrals don't move at exactly the same speed. The field follows the ions, since they are much less resistive, and flux freezing for them is a very good approximation, but the neutrals are able to drift across field lines. This slippage between ions and neutrals is called ion-neutral drift, or ambipolar diffusion.

To estimate how this process works, we need to think about the forces acting on both ions and neutrals. The ions feel the Lorentz force we wrote down in our derivation of the virial theorem:

$$\mathbf{f}_L = \frac{1}{4\pi} (\nabla \times \mathbf{B}) \times \mathbf{B} \quad (5.21)$$

The other force in play is the drag force due to ion-neutral collisions, which is

$$\mathbf{f}_d = \gamma \rho_n \rho_i (\mathbf{v}_i - \mathbf{v}_n), \quad (5.22)$$

where the subscript  $i$  and  $n$  refer to ions and neutrals, respectively, and  $\gamma$  is the drag coefficient, which can be computed from the micro-physics of the plasma.

In a very weakly ionized fluid, the neutral and ions very quickly reach terminal velocity with respect to one another, so the drag force and the Lorentz force must balance. Equating our expressions and solving for  $\mathbf{v}_d = \mathbf{v}_i - \mathbf{v}_n$ , the drift velocity, we get

$$\mathbf{v}_d = \frac{1}{4\pi\gamma\rho_n\rho_i} (\nabla \times \mathbf{B}) \times \mathbf{B} \quad (5.23)$$

To figure out how this affects the fluid, we write down the equation of magnetic field evolution under the assumption that the field is perfectly frozen into the ions:

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}_i) = 0. \quad (5.24)$$

To figure out how the field behaves with respect to the neutrals, which constitute most of the mass, we simply use our expression for the drift speed  $\mathbf{v}_d$  to eliminate  $\mathbf{v}_i$ .

With a little algebra, the result is

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}_n) = \nabla \times \left\{ \frac{\mathbf{B}}{4\pi\gamma\rho_n\rho_i} \times [\mathbf{B} \times (\nabla \times \mathbf{B})] \right\}. \quad (5.25)$$

Referring back to the MHD evolution equation,

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}) = -\nabla \times (\eta \nabla \times \mathbf{B}), \quad (5.26)$$

we can see that the resistivity produced by ambipolar drift isn't a scalar, and that it is non-linear, in the sense that it depends on  $\mathbf{B}$  itself.

However, our scaling analysis still applies. The magnitude of the resistivity produced by ambipolar drift is

$$\eta_{AD} = \frac{B^2}{4\pi\rho_i\rho_n\gamma}. \quad (5.27)$$

Thus, the magnetic Reynolds number is

$$Rm = \frac{LV}{\eta_{AD}} = \frac{4\pi LV\rho_i\rho_n\gamma}{B^2} \approx \frac{4\pi LV\rho^2 x\gamma}{B^2}, \quad (5.28)$$

where  $x = n_i/n_n$  is the ion fraction, which we've assumed is  $\ll 1$  in the last step. Ion-neutral drift will allow the magnetic field lines to drift through the fluid on length scales  $L$  such that  $Rm \lesssim 1$ . Thus, we can define a characteristic length scale for ambipolar diffusion by

$$L_{AD} = \frac{B^2}{4\pi\rho^2 x\gamma V} \quad (5.29)$$

In order to evaluate this numerically, we must calculate two things from microphysics: the ion-neutral drag coefficient  $\gamma$  and the ionization fraction  $x$ . For  $\gamma$ , the dominant effect at low speeds is that ions induce a dipole moment in nearby neutrals, which allows them to undergo a Coulomb interaction. This greatly enhances the cross-section relative to the geometric value. We won't go into details of that calculation, and will simply adopt the results:  $\gamma \approx 9.2 \times 10^{13} \text{ cm}^3 \text{ s}^{-1} \text{ g}^{-1}$  (Smith & Mac Low 1997; note that Shu 1992 gives a value that is lower by a factor of  $\sim 3$ , based on an earlier calculation).

The remaining thing we need to know to compute the drag force is the ion density. In a molecular cloud the gas is almost all neutral, and the high opacity excludes most stellar ionizing radiation. The main source of ions is cosmic rays, which can penetrate the cloud, although nearby strong x-ray sources can also contribute if present. We've already discussed them in the context of cloud heating.

Calculating the ionization fraction requires balancing this against the recombination rate, which is a nasty problem. That is because recombination is dominated by different processes at different densities, and recombinations are usually catalyzed by dust grains rather than occurring in the gas phase. Rather than trying to model all this, which is the subject of many research papers, we will simply adopt the results of a calculation by Tielens (2005):

$$n_i \approx 2 \times 10^{-3} \text{ cm}^{-3} \left( \frac{n_H}{10^4 \text{ cm}^{-3}} \right)^{1/2} \left( \frac{\zeta}{10^{-16} \text{ s}^{-1}} \right)^{1/2} \quad (5.30)$$

Thus at a density  $n_H \sim 100 \text{ cm}^{-3}$ , we expect  $x \approx 10^{-6}$ .

Plugging this into our formulae, along with our characteristic numbers  $L$  of a few tens of pc,  $V \sim$  a few  $\text{km s}^{-1}$ , and  $B \sim 10 \mu\text{G}$ , we find

$$\text{Rm} \approx 50 \quad (5.31)$$

$$L_{\text{AD}} \sim 0.5 \text{ pc.} \quad (5.32)$$

If we put in numbers for  $L$  and  $V$  more appropriate for cores than entire GMCs, we get  $L_{\text{AD}} \sim 0.05 \text{ pc}$ . Thus we expect the gas to act like a fully ionized gas on scales larger than this, but to switch over to behaving hydrodynamically on small scales.

### 5.3.2 Turbulent Reconnection

A final non-ideal MHD effect that has gotten a lot of attention lately is turbulent reconnection. The general idea of reconnection is that, in regions of non-zero resistivity where oppositely directed field lines are brought into close contact, the field lines can break and the field geometry can relax to a lower energy configuration. This may allow the field to slip out of the matter, and it always involves a reduction in magnetic pressure and energy density. The released energy is transformed into heat.

The simplest model of reconnection, the Sweet-Parker model, considers two regions of oppositely-directed field that meet at a plane. On that plane, a large current must flow in order to maintain the oppositely-directed fields on either side of it. Within this sheet reconnection can occur. As with ion-neutral drift, we can define a characteristic Reynolds-like number for this process, in this case called the Lundqvist number:

$$\mathcal{R}_L = \frac{LV}{\eta}, \quad (5.33)$$

where here  $\eta$  is the true microphysical resistivity, as opposed to the term describing ambipolar diffusion.

The rate at which reconnection can occur in the Sweet-Parker model is dictated by geometry. Matter is brought into the thin reconnection region, it reconnects, and then it must exit so that new reconnecting matter can be brought in. Matter can only exit the layer at the Alfvén speed, and since the cross-section of the reconnection layer is small, this sets severe limits on the rate at which reconnection can occur. It turns out that one can show that the maximum speed at which matter can be brought into the reconnection region is of order  $\mathcal{R}_L^{1/2} v_A$ .

To figure out this speed, we need to know the resistivity, which is related to the electrical conductivity  $\sigma$  by

$$\eta = \frac{c^2}{4\pi\sigma}. \quad (5.34)$$

Deriving the conductivity of a plasma is beyond the scope of this class, but it can be done in a fairly straightforward manner. Just to sketch out the steps: the conductivity is simply the proportionality constant between the applied electric field and the resulting current:

$$J = \sigma E. \quad (5.35)$$

In a plasma the current is carried by motions of the electrons, which move much faster than the protons, and the current is simply the electron charge times the electron number density times the mean electron speed:  $J = e n_e v_e$ . To compute the mean electron speed, one balances the electric force against the drag force exerted by collisions with neutrals (which dominate in a weakly ionized plasma), in precisely the same way we derived the mean ion-neutral drift speed by balancing the drag force against the Lorentz force. Not surprisingly  $v_e$  ends up being proportional to  $E$ , and inversely proportional to the number density of H<sub>2</sub> and the cross section for H<sub>2</sub>-electron collisions. The final result of this procedure is

$$\sigma = \frac{n_e e^2}{m_e n_{H_2} \langle \sigma v \rangle_{e-H_2}} \approx 10^{17} x \text{ s}^{-1}, \quad (5.36)$$

where  $\langle \sigma v \rangle_{e-H_2} \approx 10^{-9} \text{ cm}^3 \text{ s}^{-1}$  is the mean cross-section times velocity for electron-ion collisions. Plugging this into the resistivity gives

$$\eta \approx \frac{10^3 \text{ cm}^2 \text{ s}^{-1}}{x} \quad (5.37)$$

Plugging in our typical value  $x \sim 10^{-6}$  gives  $\eta \sim 10^9 \text{ cm}^2 \text{ s}^{-1}$ , and using  $L \sim 10 \text{ pc}$  and  $V$  of a few km s<sup>-1</sup>, typical molecular cloud numbers, this implies

$$\mathcal{R}_L \sim 10^{16}. \quad (5.38)$$

Of course this makes the reconnection speed truly tiny, of order  $10^{-8}$  of  $v_A$ . So why is reconnection at all interesting? Why is it worth considering? The answer turns on the word turbulent. It turns out that the Sweet-Parker model underpredicts the observed reconnection rate in laboratory experiments or observed in Solar flares and the Earth's magnetosphere. Indeed, if Sweet-Parker were right, there would be no such things as Solar flares.

We currently lack a good understanding of reconnection, but a rough idea is that, in a turbulent medium, reconnection sheets are can be much wider due to turbulent broadening, and that this in turn removes the geometric constraint that makes the reconnection velocity much smaller than the Alfvén speed. Exactly how and when this is important in molecular clouds is a subject of very active debate in the literature right now.



# 6

## *Gravitational Instability and Collapse*

The previous two chapters provided a whirlwind tour of fluid dynamics and turbulence. However, in that discussion we completely omitted gravity, which is obviously critical to the process of star formation. We will now remedy that omission by bringing gravity back into the discussion.

### 6.1 *The Virial Theorem*

To open this topic, we will start by proving a powerful and general theorem about the behavior of fluids, known as the virial theorem. To derive the virial theorem, we begin with the MHD equations of motion, without either viscosity or resistivity (since neither of these are important for GMCs on large scales) but with gravity. We leave in the pressure forces, even though they are small, because they're also trivial to include. We will omit the viscosity term, since we convinced ourselves last time that it is truly negligible. Thus we have

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}) \quad (6.1)$$

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) = -\nabla \cdot (\rho \mathbf{v} \mathbf{v}) - \nabla P + \frac{1}{4\pi}(\nabla \times \mathbf{B}) \times \mathbf{B} - \rho \nabla \phi. \quad (6.2)$$

Here  $\phi$  is the gravitational potential, so  $-\rho \nabla \phi$  is the gravitational force per unit volume. These equations are the Eulerian equations written in conservative form.

Before we begin, life will be a bit easier if we re-write the entire second equation in a manifestly tensorial form – this simplifies the analysis tremendously. To do so, we define two tensors: the fluid pressure tensor  $\Pi$  and the Maxwell stress tensor  $\mathbf{T}_M$ , as follows:

$$\Pi \equiv \rho \mathbf{v} \mathbf{v} + P \mathbf{I} \quad (6.3)$$

$$\mathbf{T}_M \equiv \frac{1}{4\pi} \left( \mathbf{B} \mathbf{B} - \frac{B^2}{2} \mathbf{I} \right) \quad (6.4)$$

#### Suggested background reading:

- Krumholz, M. R. 2014, Phys. Rep., 539, 49, section 3.4

Like the equations of motion, there is both an Eulerian form and a Lagrangian form of the virial theorem, depending on which version of the equations of motion we start with. We'll derive the Eulerian form in class, following the original proof by McKee & Zweibel (1992), but the derivation of the Lagrangian form proceeds in a similar manner, and can be found in many standard textbooks, for example Shu (1992).

Here  $\mathbf{I}$  is the identity tensor. In tensor notation, these are

$$(\mathbf{\Pi})_{ij} \equiv \rho v_i v_j + P \delta_{ij} \quad (6.5)$$

$$(\mathbf{T}_M)_{ij} \equiv \frac{1}{4\pi} \left( B_i B_j - \frac{1}{2} B_k B_k \delta_{ij} \right) \quad (6.6)$$

With these definitions, the momentum equation just becomes

$$\frac{\partial}{\partial t} (\rho \mathbf{v}) = -\nabla \cdot (\mathbf{\Pi} - \mathbf{T}) - \rho \nabla \phi. \quad (6.7)$$

The substitution for  $\mathbf{\Pi}$  is obvious. The equivalence of  $\nabla \cdot \mathbf{T}_M$  to  $1/(4\pi)(\nabla \times \mathbf{B}) \times \mathbf{B}$  is easy to establish with a little vector manipulation, which is most easily done in tensor notation:

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = \epsilon_{ijk} \epsilon_{jmn} \left( \frac{\partial}{\partial x_m} B_n \right) B_k \quad (6.8)$$

$$= -\epsilon_{jik} \epsilon_{jmn} \left( \frac{\partial}{\partial x_m} B_n \right) B_k \quad (6.9)$$

$$= (\delta_{in} \delta_{km} - \delta_{im} \delta_{kn}) \left( \frac{\partial}{\partial x_m} B_n \right) B_k \quad (6.10)$$

$$= B_k \frac{\partial}{\partial x_k} B_i - B_k \frac{\partial}{\partial x_i} B_k \quad (6.11)$$

$$= \left( B_k \frac{\partial}{\partial x_k} B_i + B_i \frac{\partial}{\partial x_k} B_k \right) - B_k \frac{\partial}{\partial x_i} B_k \quad (6.12)$$

$$= \frac{\partial}{\partial x_k} (B_i B_k) - \frac{1}{2} \frac{\partial}{\partial x_i} (B_k^2) \quad (6.13)$$

$$= \nabla \cdot \left( \mathbf{B} \mathbf{B} - \frac{B^2}{2} \mathbf{I} \right) \quad (6.14)$$

To derive the virial theorem, we begin by imagining a cloud of gas enclosed by some fixed volume  $V$ . The surface of this volume is  $S$ . We want to know how the overall distribution of mass changes within this volume, so we begin by writing down a quantity the represents the mass distribution. This is the moment of inertia:

$$I = \int_V \rho r^2 dV. \quad (6.15)$$

We want to know how this changes in time, so we take its time derivative:

$$\dot{I} = \int_V \frac{\partial \rho}{\partial t} r^2 dV \quad (6.16)$$

$$= - \int_V \nabla \cdot (\rho \mathbf{v}) r^2 dV \quad (6.17)$$

$$= - \int_V \nabla \cdot (\rho \mathbf{v} r^2) dV + 2 \int_V \rho \mathbf{v} \cdot \mathbf{r} dV \quad (6.18)$$

$$= - \int_S (\rho \mathbf{v} r^2) \cdot d\mathbf{S} + 2 \int_V \rho \mathbf{v} \cdot \mathbf{r} dV \quad (6.19)$$

In the first step we used the fact that the volume  $V$  does not vary in time to move the time derivative inside the integral. Then in the second step we used the equation of mass conservation to substitute. In the third step we brought the  $r^2$  term inside the divergence. Finally in the fourth step we used the divergence theorem to replace the volume integral with a surface integral.

Now we take the time derivative again, and multiply by 1/2 for future convenience:

$$\frac{1}{2} \ddot{I} = -\frac{1}{2} \int_S r^2 \frac{\partial}{\partial t} (\rho \mathbf{v}) \cdot d\mathbf{S} + \int_V \frac{\partial}{\partial t} (\rho \mathbf{v}) \cdot \mathbf{r} dV \quad (6.20)$$

$$= -\frac{1}{2} \frac{d}{dt} \int_S r^2 (\rho \mathbf{v}) \cdot d\mathbf{S} - \int_V \mathbf{r} \cdot [\nabla \cdot (\mathbf{\Pi} - \mathbf{T}_M) + \rho \nabla \phi] dV \quad (6.21)$$

The term involving the tensors is easy to simplify using a handy identity, which applies to an arbitrary tensor. This is a bit easier to follow in tensor notation:

$$\int_V \mathbf{r} \cdot \nabla \cdot \mathbf{T} dV = \int_V x_i \frac{\partial}{\partial x_j} T_{ij} dV \quad (6.22)$$

$$= \int_V \frac{\partial}{\partial x_j} (x_i T_{ij}) dV - \int_V T_{ij} \frac{\partial}{\partial x_j} x_i dV \quad (6.23)$$

$$= \int_S x_i T_{ij} dS_j - \int_V \delta_{ij} T_{ij} dV \quad (6.24)$$

$$= \int_S \mathbf{r} \cdot \mathbf{T} \cdot d\mathbf{S} - \int_V \text{Tr } \mathbf{T} dV, \quad (6.25)$$

where  $\text{Tr } \mathbf{T} = T_{ii}$  is the trace of the tensor  $\mathbf{T}$ .

Applying this to our result our tensors, we note that

$$\text{Tr } \mathbf{\Pi} = 3P + \rho v^2 \quad (6.26)$$

$$\text{Tr } \mathbf{T}_M = -\frac{B^2}{8\pi} \quad (6.27)$$

Inserting this result into our expression for  $\ddot{I}$  give the virial theorem, which I will write in a more suggestive form to make its physical interpretation clearer:

$$\frac{1}{2} \ddot{I} = 2(\mathcal{T} - \mathcal{T}_S) + \mathcal{B} + \mathcal{W} - \frac{1}{2} \frac{d}{dt} \int_S (\rho \mathbf{v} r^2) \cdot d\mathbf{S}, \quad (6.28)$$

where

$$\mathcal{T} = \int_V \left( \frac{1}{2} \rho v^2 + \frac{3}{2} P \right) dV \quad (6.29)$$

$$\mathcal{T}_S = \int_S \mathbf{r} \cdot \mathbf{\Pi} \cdot d\mathbf{S} \quad (6.30)$$

$$\mathcal{B} = \frac{1}{8\pi} \int_V B^2 dV + \int_S \mathbf{r} \cdot \mathbf{T}_M \cdot d\mathbf{S} \quad (6.31)$$

$$\mathcal{W} = - \int_V \rho \mathbf{r} \cdot \nabla \phi dV \quad (6.32)$$

Written this way, we can give a clear interpretation to what these terms mean.  $\mathcal{T}$  is just the total kinetic plus thermal energy of the cloud.  $\mathcal{T}_S$  is the confining pressure on the cloud surface, including both the thermal pressure and the ram pressure of any gas flowing across the surface.  $\mathcal{B}$  is the difference between the magnetic pressure in the cloud interior, which tries to hold it up, and the magnetic pressure plus magnetic tension at the cloud surface, which try to crush it.  $\mathcal{W}$  is the gravitational energy of the cloud. If there is no external gravitational field, and  $\phi$  comes solely from self-gravity, then  $\mathcal{W}$  is just the gravitational binding energy. The final integral represents the rate of change of the momentum flux across the cloud surface.

$\ddot{\mathcal{I}}$  is the integrated form of the acceleration. For a cloud of fixed shape, it tells us the rate of change of the cloud's expansion of contraction. If it is negative, the terms that are trying to collapse the cloud (the surface pressure, magnetic pressure and tension at the surface, and gravity) are larger, and the cloud accelerates inward. If it is positive, the terms that favor expansion (thermal pressure, ram pressure, and magnetic pressure) are larger, and the cloud accelerates outward. If it is zero, the cloud neither accelerates nor decelerates.

We get a particularly simple form of the virial theorem if there is no gas crossing the cloud surface (so  $\mathbf{v} = 0$  at  $S$ ) and if the magnetic field at the surface to be a uniform value  $B_0$ . In this case the virial theorem reduces to

$$\frac{1}{2}\ddot{\mathcal{I}} = 2(\mathcal{T} - \mathcal{T}_S) + \mathcal{B} + \mathcal{W} \quad (6.33)$$

with

$$\mathcal{T}_S = \int_S r P dS \quad (6.34)$$

$$\mathcal{B} = \frac{1}{8\pi} \int_V (B^2 - B_0^2) dV. \quad (6.35)$$

In this case  $\mathcal{T}_S$  just represents the mean radius times pressure at the virial surface, and  $\mathcal{B}$  just represents the total magnetic energy of the cloud minus the magnetic energy of the background field over the same volume.

Notice that, if a cloud is in equilibrium ( $\ddot{\mathcal{I}} = 0$ ) and magnetic and surface forces are negligible, then we have  $2\mathcal{T} = -\mathcal{W}$ . Based on this result, we define the virial ratio

$$\alpha_{\text{vir}} = \frac{2\mathcal{T}}{|\mathcal{W}|}. \quad (6.36)$$

For an object for which magnetic and surface forces are negligible, and with no flow across the virial surface, a value of  $\alpha_{\text{vir}} > 1$  implies  $\ddot{\mathcal{I}} > 0$ , and a value  $\alpha_{\text{vir}} < 1$  implies  $\ddot{\mathcal{I}} < 0$ . Thus  $\alpha_{\text{vir}} = 1$  roughly

divides clouds that have enough internal pressure or turbulence to avoid collapse from those that do not.

## 6.2 Stability Conditions

Armed with the virial theorem, we are now in a position to understand, at least qualitatively, under what conditions a cloud of gas will be stable against gravitational contraction, and under what conditions it will not be. If we examine the terms on the right hand side of the virial theorem, we can group them into those that are generally or always positive, and thus oppose collapse, and those that are generally or always negative, and thus encourage it. The main terms opposing collapse are  $\mathcal{T}$ , which contains parts describing both thermal pressure and turbulent motion, and  $\mathcal{B}$ , which describes magnetic pressure and tension. The main terms favoring collapse are  $\mathcal{W}$ , representing self-gravity, and  $\mathcal{T}_S$ , representing surface pressure. The final term, the surface one, could be positive or negative depending on whether mass is flowing into our out of the virial volume. We will begin by examining the balance among these terms, and the forces they represent.

### 6.2.1 Thermal Support and the Jeans Instability

Gas pressure is perhaps the most basic force in opposing collapse. Unlike turbulent motions, which can compress in some places even as they provide overall support, gas pressure always tries to smooth out the gas. Similarly, self-gravity is the most reliable promoter of collapse. A full, formal analysis of the interaction between pressure and self-gravity was provided by James Jeans in 1902 [Jeans \(1902\)](#), and we'll go through that in a moment. However, we can already see what the basic result will have to look like just from the virial theorem. We expect the dividing line between stability and instability to lie at  $\alpha_{\text{vir}} \approx 1$ . For an isolated, isothermal cloud of mass  $M$  and radius  $R$  with only thermal pressure, we have

$$\mathcal{T} = \frac{3}{2}Mc_s^2 \quad (6.37)$$

$$\mathcal{W} = -a\frac{GM^2}{R}, \quad (6.38)$$

where  $a$  is a factor of order unity that depends on the internal density structure. Thus the condition  $\alpha_{\text{vir}} \gtrsim 1$  corresponds to

$$Mc_s^2 \gtrsim \frac{GM^2}{R} \implies R \gtrsim \frac{GM}{c_s^2}, \quad (6.39)$$

or, rewriting in terms of the mean density  $\rho \sim M/R^3$ ,

$$R \gtrsim \frac{c_s}{\sqrt{G\rho}}. \quad (6.40)$$

The formal analysis proceeds as follows. Consider a uniform, infinite, isothermal medium at rest. The density is  $\rho_0$ , the pressure is  $P_0 = \rho_0 c_s^2$ , and the velocity is  $\mathbf{v}_0 = 0$ . We will write down the equations of hydrodynamics and self-gravity for this gas:

$$\frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (6.41)$$

$$\frac{\partial}{\partial t} (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \mathbf{v}) = -\nabla P - \rho \nabla \phi \quad (6.42)$$

$$\nabla^2 \phi = 4\pi G \rho. \quad (6.43)$$

Here the first equation represents conservation of mass, the second represents conservation of momentum, and the third is the Poisson equation for the gravitational potential  $\phi$ . We take the background density  $\rho_0$ , velocity  $\mathbf{v}_0 = 0$ , pressure  $P_0$ , and potential  $\phi_0$  to be an exact solution of these equations, so that all time derivatives are zero as long as the gas is not disturbed.

Note that this involves the "Jeans swindle": this assumption is actually not really consistent, because the Poisson equation cannot be solved for an infinite uniform medium unless  $\rho_0 = 0$ . In other words, there is no function  $\phi_0$  such that  $\nabla^2 \phi_0$  is equal to a non-zero constant value on all space. That said, we will ignore this complication, since the approximation of a uniform infinite medium is a reasonable one for a very large but finite uniform medium. It is possible to construct the argument without the Jeans swindle, but doing so adds mathematical encumbrance without physical insight, so we will not do so.

That digression aside, now let us consider what happens if we perturb this system. We will write the density as  $\rho = \rho_0 + \epsilon \rho_1$ , where  $\epsilon \ll 1$ . Similarly, we write  $\mathbf{v} = \epsilon \mathbf{v}_1$  and  $\phi = \phi_0 + \epsilon \phi_1$ . Since we can always use Fourier analysis to write an arbitrary perturbation as a sum of Fourier components, without loss of generality we will take the perturbation to be a single, simple Fourier mode. The reason to do this is that, as we will say, differential equations are trivial to solve when the functions in question are simple plane waves.

Thus we write  $\rho_1 = \rho_a \exp[i(kx - \omega t)]$ . Note that we implicitly understand that we use only the real part of this exponential. It is just easier to write things in terms of an  $e^{i(kx - \omega t)}$  than it is to keep track of a bunch of sines and cosines. In writing this equation, we have chosen to orient our coordinate system so that the wave vector  $\mathbf{k}$  of the perturbation is in the  $x$  direction. Again, there is no loss of generality in doing so.

Given this density perturbation, what is the corresponding perturbation to the potential? From the Poisson equation, we have

$$\nabla^2(\phi_0 + \epsilon\phi_1) = 4\pi G(\rho_0 + \epsilon\rho_1). \quad (6.44)$$

Since by assumption  $\rho_0$  and  $\phi_0$  are exact solutions to the Poisson equation, we can cancel the  $\phi_0$  and  $\rho_0$  terms out of the equation, leaving

$$\nabla^2\phi_1 = 4\pi G\rho_1 = 4\pi G\rho_a e^{i(kx-\omega t)}. \quad (6.45)$$

This equation is trivial to solve, since it is just of the form  $y'' = ae^{bx}$ .

The solution is

$$\phi_1 = -\frac{4\pi G\rho_a}{k^2} e^{i(kx-\omega t)} \quad (6.46)$$

By analogy to what we did for  $\rho_1$ , we write this solution as  $\phi_1 = \phi_a e^{i(kx-\omega t)}$ , with

$$\phi_a = -\frac{4\pi G\rho_a}{k^2}. \quad (6.47)$$

Now that we have found the perturbed potential, let us determine what motion this will induce in the fluid. To do so, we first take the equations of mass and momentum conservation and we linearize them. This means that we substitute in  $\rho = \rho_0 + \epsilon\rho_1$ ,  $\mathbf{v} = \epsilon\mathbf{v}_1$ ,  $P = P_0 + \epsilon P_1 = c_s^2(\rho_0 + \epsilon\rho_1)$ , and  $\phi = \phi_0 + \epsilon\phi_1$ . Note that  $\mathbf{v}_0 = 0$ . We then expand the equations in powers of  $\epsilon$ , and we drop all the terms that are of order  $\epsilon^2$  or higher on the grounds that they become negligible in the limit of small  $\epsilon$ .

Linearizing the equation of motion we get

$$\frac{\partial}{\partial t}(\rho_0 + \epsilon\rho_1) + \nabla \cdot [(\rho_0 + \epsilon\rho_1)(\epsilon\mathbf{v}_1)] = 0 \quad (6.48)$$

$$\frac{\partial}{\partial t}\rho_0 + \epsilon \frac{\partial}{\partial t}\rho_1 + \epsilon \nabla \cdot (\rho_0\mathbf{v}_1) = 0 \quad (6.49)$$

$$\frac{\partial}{\partial t}\rho_1 + \nabla \cdot (\rho_0\mathbf{v}_1) = 0 \quad (6.50)$$

In the second step, we dropped a term of order  $\epsilon^2$ . In the third step we used the fact that  $\rho_0$  is constant, i.e. that the background density has zero time derivative, to drop that term.

Applying the same procedure to the momentum / Euler equation, we get

$$\begin{aligned} \frac{\partial}{\partial t}[(\rho_0 + \epsilon\rho_1)(\epsilon\mathbf{v}_1)] + \nabla \cdot [(\rho_0 + \epsilon\rho_1)(\epsilon\mathbf{v}_1)(\epsilon\mathbf{v}_1)] \\ = -c_s^2 \nabla(\rho_0 + \epsilon\rho_1) \\ - (\rho_0 + \epsilon\rho_1) \nabla(\phi_0 + \epsilon\phi_1) \end{aligned} \quad (6.51)$$

$$\begin{aligned} \epsilon \frac{\partial}{\partial t}(\rho_0\mathbf{v}_1) = -c_s^2 \nabla\rho_0 - \rho_0 \nabla\phi_0 \\ - \epsilon \left( c_s^2 \nabla\rho_1 + \rho_1 \nabla\phi_0 + \rho_0 \nabla\phi_1 \right) \end{aligned} \quad (6.52)$$

$$\frac{\partial}{\partial t}(\rho_0\mathbf{v}_1) = -c_s^2 \nabla\rho_1 - \rho_0 \nabla\phi_1 \quad (6.53)$$

In the second step we dropped terms of order  $\epsilon^2$ , and in the third step we used the fact that the background state is uniform to drop terms involving gradients of  $\rho_0$  and  $\phi_0$ .

Now that we have our linearized equations, we're ready to find out what  $\mathbf{v}_1$  must be. By analogy to what we did for  $\rho_1$  and  $\phi_1$ , we take  $\mathbf{v}_1$  to be a single Fourier mode, of the form

$$\mathbf{v}_1 = \mathbf{v}_a e^{i(kx - \omega t)} \quad (6.54)$$

Substituting for  $\rho_1$ ,  $\phi_1$ , and  $\mathbf{v}_1$  into the linearized mass conservation equation, we get

$$\frac{\partial}{\partial t} (\rho_a e^{i(kx - \omega t)}) + \nabla \cdot (\rho_0 \mathbf{v}_a e^{i(kx - \omega t)}) = 0 \quad (6.55)$$

$$-i\omega \rho_a e^{i(kx - \omega t)} + ik\rho_0 v_{a,x} e^{i(kx - \omega t)} = 0 \quad (6.56)$$

$$-\omega \rho_a + k\rho_0 v_{a,x} = 0 \quad (6.57)$$

$$\frac{\omega \rho_a}{k\rho_0} = v_{a,x} \quad (6.58)$$

where  $v_{a,x}$  is the  $x$  component of  $\mathbf{v}_a$ .

We have now found the velocity perturbation in terms of  $\rho_a$ ,  $\omega$ , and  $k$ . Similarly substituting into the momentum equation

$$\frac{\partial}{\partial t} (\rho_0 \mathbf{v}_a e^{i(kx - \omega t)}) = -c_s^2 \nabla (\rho_a e^{i(kx - \omega t)}) - \rho_0 \nabla (\phi_a e^{i(kx - \omega t)}) \quad (6.59)$$

$$-i\omega \rho_0 \mathbf{v}_a e^{i(kx - \omega t)} = -ikc_s^2 \rho_a \hat{\mathbf{x}} e^{i(kx - \omega t)} - ik\rho_0 \phi_a e^{i(kx - \omega t)} \hat{\mathbf{x}} \quad (6.60)$$

$$\omega \rho_0 v_{a,x} = k (c_s^2 \rho_a + \rho_0 \phi_a) \quad (6.61)$$

Now let us take this equation and substitute in the values for  $\phi_a$  and  $v_{a,x}$  that we previously determined:

$$\omega \rho_0 \left( \frac{\omega \rho_a}{k\rho_0} \right) = kc_s^2 \rho_a - k\rho_0 \left( \frac{4\pi G \rho_a}{k^2} \right) \quad (6.62)$$

$$\omega^2 = c_s^2 k^2 - 4\pi G \rho_0 \quad (6.63)$$

This expression is known as a dispersion relation, because it describes the dispersion of the plane wave solution we have found, i.e. how that wave's spatial frequency  $k$  relates to its temporal frequency  $\omega$ .

To see what this implies, let's consider what happens when we put in a perturbation with a short wavelength or a large spatial frequency. In this case  $k$  is large, and  $c_s^2 k^2 - 4\pi G \rho_0 > 0$ , so  $\omega$  is a positive or negative real number. The density is  $\rho = \rho_0 + \rho_a e^{i(kx - \omega t)}$ , which represents a uniform background density with a small oscillation

in space and time on top of it. Since  $|e^{i(kx-\omega t)}| < 1$  at all times and places, the oscillation does not grow.

On the other hand, suppose that we impose a perturbation with a large spatial range, or a small spatial frequency. In this case  $c_s^2 k^2 - 4\pi G \rho_0 < 0$ , so  $\omega$  is a positive or negative imaginary number. For an imaginary  $\omega$ ,  $|e^{-i\omega t}|$  either decays to zero or grows infinitely large in time, depending on whether we take the positive or negative imaginary root. Thus at least one solution for the perturbation will not remain small. It will grow in amplitude without limit.

This represents an instability: if we impose an arbitrarily small amplitude perturbation on the density at a sufficiently large wavelength, that perturbation will eventually grow to be large. Of course once  $\rho_1$  becomes large enough, our linearization procedure of dropping terms proportional to  $\epsilon^2$  becomes invalid, since these terms are no longer small. In this case we must follow the full non-linear behavior of the equations, usually with simulations.

We have, however, shown that there is a critical size scale beyond which perturbations that are stabilized only by pressure must grow to non-linear amplitude. The critical length scale is set by the value of  $k$  for which  $\omega = 0$ :

$$k_J = \sqrt{\frac{4\pi G \rho_0}{c_s^2}}. \quad (6.64)$$

The corresponding wavelength is

$$\lambda_J = \frac{2\pi}{k_J} = \sqrt{\frac{\pi c_s^2}{G \rho_0}}. \quad (6.65)$$

This is known as the Jeans length. One can also define a mass scale associated with this: the Jeans mass,  $M_J = \rho \lambda_J^3$ .

If we plug in some typical numbers for a GMC,  $c_s = 0.2 \text{ km s}^{-1}$  and  $\rho_0 = 100 m_p$ , we get  $\lambda_J = 3.4 \text{ pc}$ . Since every GMC we have seen is larger than this size, and there are clearly always perturbations present, this means that molecular clouds cannot be stabilized by gas pressure against collapse. Of course you could have guessed this just by evaluating terms in the virial theorem: the gas pressure term is very small compared to the gravitational one. Ultimately, the virial theorem and the Jeans instability analysis are just two different ways of extracting the same information from the equations of motion.

One nice thing about the Jeans analysis, however, is that it makes it obvious how fast we should expect the perturbation to grow. Suppose we have a very unstable system, where  $c_s^2 k^2 \ll 4\pi G \rho_0$ . This is the case for GMC, for example. There are perturbations on the size of the entire cloud, which might be 50 pc in size. This is a spatial frequency  $k = 2\pi/(50 \text{ pc}) = 0.12 \text{ pc}^{-1}$ . Plugging this in with  $c_s = 0.2$

$\text{km s}^{-1}$  and  $\rho_0 = 100m_p$  gives  $c_s^2 k^2 / (4\pi G \rho) = 0.005$ . In this case we have

$$\omega \approx \pm i\sqrt{4\pi G \rho_0}. \quad (6.66)$$

Taking the negative  $i$  root, which corresponds to the growing mode, we find that

$$\rho_1 \propto \exp([4\pi G \rho_0]^{1/2} t). \quad (6.67)$$

Thus the  $e$ -folding time for the disturbance to grow is  $\sim 1/\sqrt{G\rho_0}$ . We define the free-fall time as

$$t_{\text{ff}} = \sqrt{\frac{3\pi}{32G\rho_0}}, \quad (6.68)$$

where the numerical coefficient of  $\sqrt{3\pi/32}$  comes from doing the closely related problem of the collapse of a pressureless sphere – we'll do that later this week. The free-fall time is the characteristic time scale required for a medium with negligible pressure support to collapse.

The Jeans analysis is of course only appropriate for a uniform medium, and it requires the Jeans swindle. Problem Set 2 contains a calculation of the maximum mass of a spherical cloud that can support itself against collapse by thermal pressure, called the Bonnor-Ebert mass (Ebert, 1955; Bonnor, 1956). Not surprisingly, the Bonnor-Ebert mass is simply  $M_J$  times factors of order unity.

### 6.2.2 Magnetic Support and the Magnetic Critical Mass

Let us now consider another term that generally opposes collapse: the magnetic one. Let us consider a uniform spherical cloud of radius  $R$  threaded by a magnetic field  $\mathbf{B}$ . We imagine that  $\mathbf{B}$  is uniform inside the cloud, but that outside the cloud the field lines quickly spread out, so that the magnetic field drops down to some background strength  $\mathbf{B}_0$ , which is also uniform but has a magnitude much smaller than  $\mathbf{B}$ .

Here it is easiest to work directly with the virial theorem. The magnetic term in the virial theorem is

$$\mathcal{B} = \frac{1}{8\pi} \int_V B^2 dV + \int_S \mathbf{x} \cdot \mathbf{T}_M \cdot d\mathbf{S} \quad (6.69)$$

where

$$\mathbf{T}_M = \frac{1}{4\pi} \left( \mathbf{B}\mathbf{B} - \frac{B^2}{2} \mathbf{I} \right). \quad (6.70)$$

If the field inside the cloud is much larger than the field outside it, then the first term, representing the integral of the magnetic pressure within the cloud, is

$$\frac{1}{8\pi} \int_V B^2 dV \approx \frac{B^2 R^3}{6}. \quad (6.71)$$

Here we have dropped any contribution from the field outside the cloud. The second term, representing the surface magnetic pressure and tension, is

$$\int_S \mathbf{x} \cdot \mathbf{T}_M \cdot d\mathbf{S} = \int_S \frac{B_0^2}{8\pi} \mathbf{x} \cdot d\mathbf{S} \approx \frac{B_0^2 R_0^3}{6} \quad (6.72)$$

Since the field lines that pass through the cloud must also pass through the virial surface, it is convenient to rewrite everything in terms of the magnetic flux. The flux passing through the cloud is  $\Phi_B = \pi B R^2$ , and since these field lines must also pass through the virial surface, we must have  $\Phi_B = \pi B_0 R_0^2$  as well. Thus, we can rewrite the magnetic term in the virial theorem as

$$\mathcal{B} \approx \frac{B^2 R^3}{6} - \frac{B_0^2 R_0^2}{6} = \frac{1}{6\pi^2} \left( \frac{\Phi_B^2}{R} - \frac{\Phi_B^2}{R_0} \right) \approx \frac{\Phi_B^2}{6\pi^2 R}. \quad (6.73)$$

In the last step we used the fact that  $R \ll R_0$  to drop the  $1/R_0$  term. Now let us compare this to the gravitational term, which is

$$\mathcal{W} = -\frac{3}{5} \frac{GM^2}{R} \quad (6.74)$$

for a uniform cloud of mass  $M$ . Comparing these two terms, we find that

$$\mathcal{B} + \mathcal{W} = \frac{\Phi_B^2}{6\pi^2 R} - \frac{3}{5} \frac{GM^2}{R} \equiv \frac{3}{5} \frac{G}{R} \left( M_\Phi^2 - M^2 \right) \quad (6.75)$$

where

$$M_\Phi \equiv \sqrt{\frac{5}{2}} \left( \frac{\Phi_B}{3\pi G^{1/2}} \right) \quad (6.76)$$

We call  $M_\Phi$  the magnetic critical mass. Since  $\Phi_B$  does not change as a cloud expands or contracts (due to flux-freezing), this magnetic critical mass does not change either.

The implication of this is that clouds that have  $M > M_\Phi$  always have  $\mathcal{B} + \mathcal{W} < 0$ . The magnetic force is unable to halt collapse no matter what. Clouds that satisfy this condition are called magnetically supercritical, because they are above the magnetic critical mass  $M_\Phi$ . Conversely, if  $M < M_\Phi$ , then  $\mathcal{B} + \mathcal{W} > 0$ , and gravity is weaker than magnetism. Clouds satisfying this condition are called subcritical. For a subcritical cloud, since  $\mathcal{B} + \mathcal{W} \propto 1/R$ , this term will get larger and larger as the cloud shrinks.

In other words, not only is the magnetic force resisting collapse is stronger than gravity, it becomes larger and larger without limit as the cloud is compressed to a smaller radius. Unless the external pressure is also able to increase without limit, which is unphysical, then there is no way to make a magnetically subcritical cloud collapse. It will always stabilize at some finite radius. The only way to get

around this is to change the magnetic critical mass, which requires changing the magnetic flux through the cloud. This is possible only via ambipolar diffusion or some other non-ideal MHD effect that violates flux-freezing.

Of course our calculation is for a somewhat artificial configuration of a spherical cloud with a uniform magnetic field. In reality a magnetically-supported cloud will not be spherical, since the field only supports it in some directions, and the field will not be uniform, since gravity will always bend it some amount. Figuring out the magnetic critical mass in that case requires solving for the cloud structure numerically. A calculation of this effect by Tomisaka (1998) gives

$$M_\Phi = 0.12 \frac{\Phi_B}{G^{1/2}} \quad (6.77)$$

for clouds for which pressure support is negligible. The numerical coefficient we got for the uniform cloud case is 0.17, so this is obviously a small correction. It is also possible to derive a combined critical mass that incorporates both the flux and the sound speed, and which limits to the Bonnor-Ebert mass for negligible field and the magnetic critical mass for negligible pressure.

It is not so easy to determine observationally whether the magnetic fields are strong enough to hold up molecular clouds. The observations are somewhat complicated by the fact that, using the most common technique of Zeeman splitting, one can only measure the line of sight component of the field. This therefore gives only a lower limit on the magnetic critical mass. Nonetheless, for a large enough sample, one can estimate true magnetic field strengths statistically under the assumption of random orientations. When this analysis is performed, the current observational consensus is that magnetic fields in molecular clouds are not, by themselves, strong enough to prevent gravitational collapse. Figure 6.1 shows a summary of the current observations. Clearly atomic gas is magnetically subcritical, but molecular gas is supercritical.

### 6.2.3 *Turbulent Support*

There is one more positive term in the virial theorem, which is the turbulent component of  $\mathcal{T}$ . This one is not at all well understood, largely because we don't understand turbulence itself. This term almost certainly provides some support against collapse, but the amount is not well understood, and we will defer any further discussion of this effect until we get to our discussions of the star formation rate in Chapter ??.

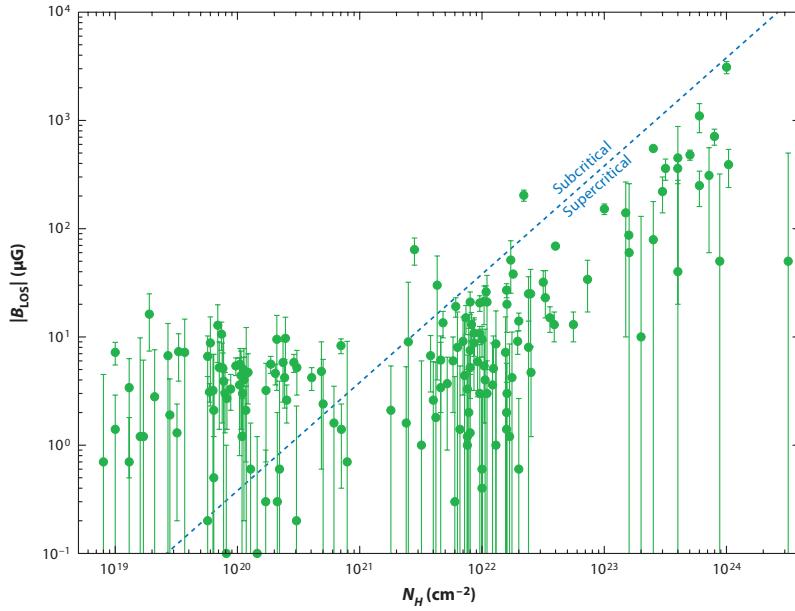


Figure 6.1: Measurements of the line of sight magnetic field strength from the Zeeman effect, versus total gas column density in H atoms  $\text{cm}^{-2}$  (Crutcher, 2012). The three clumps of points represent, from left to right, measurements from the Zeeman splitting of H I, OH, and CN. The dashed blue line indicates the separation between field strengths that are large enough to render the gas subcritical, and those weak enough for it to be supercritical.

### 6.3 Pressureless Collapse

As a final topic for this chapter, let us consider what we should expect to happen if gas does begin to collapse. Let us consider the simplest case of an initially-spherical cloud with an initial density distribution  $\rho(r)$ . We would like to know how the gas moves under the influence of gravity and thermal pressure, under the assumption of spherical symmetry. For convenience we define the enclosed mass

$$M_r = \int_0^r 4\pi r'^2 \rho(r') dr' \quad (6.78)$$

or equivalently

$$\frac{\partial M_r}{\partial r} = 4\pi r^2 \rho. \quad (6.79)$$

The equation of mass conservation for the gas in spherical coordinates is

$$\frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (6.80)$$

$$\frac{\partial}{\partial t} \rho + \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \rho v) = 0, \quad (6.81)$$

where  $v$  is the radial velocity of the gas.

It is useful to write the equations in terms of  $M_r$  rather than  $\rho$ , so we take the time derivative of  $M_r$  to get

$$\frac{\partial}{\partial t} M_r = 4\pi \int_0^{r'} r'^2 \frac{\partial}{\partial t} \rho dr' \quad (6.82)$$

$$= -4\pi \int_0^{r'} \frac{\partial}{\partial r'} (r'^2 \rho v) dr' \quad (6.83)$$

$$= -4\pi r^2 \rho v \quad (6.84)$$

$$= -v \frac{\partial}{\partial r} M_r. \quad (6.85)$$

In the second step we used the mass conservation equation to substitute for  $\partial\rho/\partial t$ , and in the final step we used the definition of  $M_r$  to substitute for  $\rho$ .

To figure out how the gas moves, we write down the Lagrangean version of the momentum equation:

$$\rho \frac{Dv}{Dt} = -\frac{\partial}{\partial r} P - \mathbf{f}_g, \quad (6.86)$$

where  $\mathbf{f}_g$  is the gravitational force. For the momentum equation, we take advantage of the fact that the gas is isothermal to write  $P = \rho c_s^2$ . The gravitational force is  $\mathbf{f}_g = -GM_r/r^2$ . Thus we have

$$\frac{Dv}{Dt} = \frac{\partial}{\partial t} v + v \frac{\partial}{\partial r} v = -\frac{c_s^2}{\rho} \frac{\partial}{\partial r} \rho - \frac{GM_r}{r^2}. \quad (6.87)$$

To go further, let us make one more simplifying assumption: that the sound speed  $c_s$  is zero. This is not as bad an approximation as you might think. Consider the virial theorem: the thermal pressure term is just proportional to the mass, since the gas sound speed stays about constant. On the other hand, the gravitational term varies as  $1/R$ . Thus, even if pressure starts out competitive with gravity, as the core collapses the dominance of gravity will increase, and before too long the collapse will resemble a pressureless one.

In this case the momentum equation is trivial:

$$\frac{Dv}{Dt} = -\frac{GM_r}{r^2}. \quad (6.88)$$

This just says that a shell's inward acceleration is equal to the gravitational force per unit mass exerted by all the mass interior to it, which is constant. We can then solve for the velocity as a function of position:

$$v = \dot{r} = -\sqrt{2GM_r} \left( \frac{1}{r_0} - \frac{1}{r} \right)^{1/2}, \quad (6.89)$$

where  $r_0$  is the position at which a particular fluid element starts.

The integral can be evaluated by the trigonometric substitution  $r = r_0 \cos^2 \xi$ . The solution, first obtained by Hunter (1962), is

$$-2r_0(\cos \xi \sin \xi) \dot{\xi} = -\sqrt{\frac{2GM_r}{r_0}} \left( \frac{1}{\cos^2 \xi} - 1 \right)^{1/2} \quad (6.90)$$

$$2(\cos \xi \sin \xi) \dot{\xi} = \sqrt{\frac{2GM_r}{r_0^3}} \tan \xi \quad (6.91)$$

$$2 \cos^2 \xi d\xi = \sqrt{\frac{2GM_r}{r_0^3}} dt \quad (6.92)$$

$$\xi + \frac{1}{2} \sin 2\xi = t \sqrt{\frac{2GM_r}{r_0^3}}. \quad (6.93)$$

We are interested in the time at which a given fluid element reaches the origin,  $r = 0$ . This corresponds to  $\xi = \pi/2$ , so this time is

$$t = \frac{\pi}{2} \sqrt{\frac{r_0^3}{2GM_r}}. \quad (6.94)$$

Suppose that the gas we started with was of uniform density  $\rho$ , so that  $M_r = (4/3)\pi r_0^3 \rho$ . In this case we have

$$t = t_{\text{ff}} = \sqrt{\frac{3\pi}{32G\rho}}, \quad (6.95)$$

where we have defined the free-fall time  $t_{\text{ff}}$ : it is the time required for a uniform sphere of pressureless gas to collapse to infinite density. This is of course just the characteristic growth time for the Jeans instability in the regime of negligible pressure, up to a factor of order unity.

For a uniform fluid this means that the collapse is synchronized – all the mass reaches the origin at the exact same time. A more realistic case is for the initial state to have some level of central concentration, so that the initial density rises inward. Let us take the initial density profile to be  $\rho = \rho_c(r/r_c)^{-\alpha}$ , where  $\alpha > 0$  so the density rises inward. The corresponding enclosed mass is

$$M_r = \frac{4}{3-\alpha} \pi \rho_c r_c^3 \left(\frac{r}{r_c}\right)^{3-\alpha} \quad (6.96)$$

Plugging this in, the collapse time is

$$t = \sqrt{\frac{(3-\alpha)\pi}{32G\rho_c}} \left(\frac{r_0}{r_c}\right)^{\alpha/2}. \quad (6.97)$$

Since  $\alpha > 0$ , this means that the collapse time increases with initial radius  $r_0$ . This illustrates one of the most basic features of a collapse, which will continue to hold even in the case where the pressure is non-zero. Collapse of centrally concentrated objects occurs inside-out, meaning that the inner parts collapse before the outer parts.

Within the collapsing region near the star, the density profile also approaches a characteristic shape. If the radius of a given fluid element  $r$  is much smaller than its initial radius  $r_0$ , then its velocity is roughly

$$v \approx v_{\text{ff}} \equiv -\sqrt{\frac{2GM_r}{r}}, \quad (6.98)$$

where we have defined the free-fall velocity  $v_{\text{ff}}$  as the characteristic speed achieved by an object collapsing freely onto a mass  $M_r$ . The mass conservation equation is

$$\frac{\partial M_r}{\partial t} = -v \frac{\partial M_r}{\partial r} = -4\pi r^2 v \rho \quad (6.99)$$

If we are near the star so that  $v \approx v_{\text{ff}}$ , then this implies that

$$\rho = \frac{(\partial M_r / \partial t) r^{-3/2}}{4\pi \sqrt{2GM_r}}. \quad (6.100)$$

To the extent that we look at a short interval of time, over which the accretion rate does not change much (so that  $\partial M_r / \partial t$  is roughly constant), this implies that the density near the star varies as  $\rho \propto r^{-3/2}$ .

What sort of accretion rate do we expect from a collapse like this? For a core of mass  $M_c = [4/(3-\alpha)]\pi\rho_c r_c^3$ , the last mass element arrives at the center at a time

$$t_c = \sqrt{\frac{(3-\alpha)\pi}{32G\rho_c}} = \sqrt{\frac{3-\alpha}{3}} t_{\text{ff}}(\rho_c), \quad (6.101)$$

so the time-averaged accretion rate is

$$\langle \dot{M} \rangle = \sqrt{\frac{3}{3-\alpha}} \frac{M_c}{t_{\text{ff}}(\rho_c)}. \quad (6.102)$$

In order to get a sense of the numerical value of this, let us suppose that our collapsing object is a marginally unstable Bonnor-Ebert sphere. Such an object does not have negligible pressure, but the pressure will only change the collapse rate at order unity. Problem Set 2 includes a calculation of the structure of a maximum-mass Bonnor-Ebert sphere, so we will just quote the value. The maximum mass is

$$M_{\text{BE}} = 1.18 \frac{c_s^4}{\sqrt{G^3 P_s}}, \quad (6.103)$$

where  $P_s$  is the pressure at the surface of the sphere and  $c_s$  is the thermal sound speed in the core.

Let's suppose that the surface of the core, at radius  $r_c$ , is in thermal pressure balance with its surroundings. Thus  $P_s = \rho_c c_s^2$ , so we may rewrite the Bonnor-Ebert mass as

$$M_{\text{BE}} = 1.18 \frac{c_s^3}{\sqrt{G^3 \rho_c}}. \quad (6.104)$$

A Bonnor-Ebert sphere doesn't have a powerlaw structure, but if we substitute into our equation for the accretion rate and say that the

factor of  $\sqrt{3/(3-\alpha)}$  is a number of order unity, we find that the accretion rate is

$$\langle \dot{M} \rangle \approx \frac{c_s^3 / \sqrt{G^3 \rho_c}}{1 / \sqrt{G \rho_c}} = \frac{c_s^3}{G}. \quad (6.105)$$

This is an extremely useful expression, because we know the sound speed  $c_s$  from microphysics. Thus, we have calculated the rough accretion rate we expect to be associated with the collapse of any object that is marginally stable based on thermal pressure support. Plugging in  $c_s = 0.19 \text{ km s}^{-1}$ , we get  $\dot{M} \approx 2 \times 10^{-6} M_\odot \text{ yr}^{-1}$  as the characteristic accretion rate for these objects. Since the typical stellar mass is a few tenths of  $M_\odot$ , based on the peak of the IMF, this means that the characteristic star formation time is of order  $10^5 - 10^6$  yr. Of course this conclusion about the accretion rate only applies to collapsing objects that are supported mostly by thermal pressure. Other sources of support produce higher accretion rates, as we will see when we get to massive stars.



# 7

## *Stellar Feedback*

The final piece of physics we will discuss before moving on to the star formation process itself is the interaction of stellar radiation, winds, and other forms of feedback with the interstellar medium. Our goal is to understand the myriad ways that stars influence their environments. This is particularly relevant to star formation because for the most part these influences are exerted most strongly by young stellar populations, and therefore constitute an important part in the regulation of star formation.

### *7.1 General Formalism*

#### *7.1.1 IMF-Averaged Yields*

In most cases when considering feedback, we will be averaging over many, many stars. Consequently, it makes sense to focus not on individual stars, but on the collective properties of stellar populations. For this reason, a very useful first step is to consider budgets of mass, momentum, and energy.

We've already encountered a formalism of this sort in our discussion of galactic star formation rate indicators, and the idea is similar here. To begin, let us fix the IMF

$$\xi(m) \equiv \frac{dn}{d \ln m}, \quad (7.1)$$

with the normalization chosen so that  $\int \xi(m) dm = 1$ . We will assume that this is invariant, for lack of really convincing evidence otherwise (although this is hotly debated). The mean stellar mass is

$$\bar{m} = \frac{\int_{-\infty}^{\infty} m \xi(m) d \ln m}{\int_{-\infty}^{\infty} \xi(m) d \ln m} = \frac{1}{\int_{-\infty}^{\infty} \xi(m) d \ln m}, \quad (7.2)$$

where the second step follows from our choice of normalization.

The numerator here represents the total mass of the stars, and the

#### **Suggested background reading:**

- Krumholz, M. R., et al. 2014, in "Protostars and Planets VI", ed. H. Beuther et al., pp. 243-266

#### **Suggested literature:**

- Murray, N., Quataert, E., & Thompson, T. A. 2010, ApJ, 709, 191
- Dale, J. E., Ngoumou, J., Ercolano, B., & Bonnell, I. A. 2014, MNRAS, 442, 694

denominator is the number of stars. Note that  $\xi(m)$  is presumably zero outside some finite interval in mass – we are writing the limits of integration as  $-\infty$  to  $\infty$  only for convenience.

We will further assume that, from stellar evolution, we know the rate  $q$  at which stars produce some quantity  $Q$  as a function of their starting mass and age, where  $\dot{Q} = q$ . For example if the quantity  $Q$  we are concerned with is total radiant energy  $E$ , then  $q$  is the bolometric luminosity  $L(m, t)$  of a star of mass  $m$  and age  $t$ . Now consider a population of stars that forms in a single burst at time  $0$ . The instantaneous production rate for these stars is

$$q(t) = M \int_{-\infty}^{\infty} d \ln m \xi(m) q(m, t). \quad (7.3)$$

We use this equation to define the IMF-averaged production rate,

$$\left\langle \frac{q}{M} \right\rangle = \int_{-\infty}^{\infty} d \ln m \xi(m) q(m, t). \quad (7.4)$$

Note that this rate is a function of the age of the stellar population  $t$ . We can also define a lifetime-averaged yield. Integrating over all time, the total amount of the quantity produced is

$$Q = M \int_{-\infty}^{\infty} d \ln m \xi(m) \int_0^{\infty} dt q(M, t). \quad (7.5)$$

We therefore define the IMF-averaged yield

$$\left\langle \frac{Q}{M} \right\rangle = \int_{-\infty}^{\infty} d \ln m \xi(m) \int_0^{\infty} dt q(M, t). \quad (7.6)$$

The meaning of these quantities is that  $\langle q/M \rangle$  is the instantaneous rate at which the stars are producing  $Q$  per unit stellar mass, and  $\langle Q/M \rangle$  is the total amount produced per unit mass of stars formed over the stars' entire lifetimes.

In practice we can't really integrate to infinity for most quantities, since the lifetimes of some stars may be very, very long compared to what we're interested in. For example the luminous output of a stellar population will have a large contribution for  $\sim 5$  Myr coming from massive stars, which is mostly what is of interest. However, if we integrate for 1000 Gyr, we will find that the luminous output is dominated by the vast numbers of  $\sim 0.2 M_{\odot}$  stars near the peak of the IMF that are fully convective and thus are able to burn all of their hydrogen to He. In reality, though, this is longer than the age of the universe. In practice, therefore, we must define our lifetime averages as cutting off after some finite time.

It can also be useful to define a different IMF average. The quantities we have discussed thus far are yields per unit mass that goes into stars. Sometimes we are instead interested in the yield per unit mass

that stays locked in stellar remnants for a long time, rather than the mass that goes into stars for  $\sim 3 - 4$  Myr and then comes back out in supernovae. Let us define the mass of the remnant that a star of mass  $m$  leaves as  $w(m)$ . If the star survives for a long time,  $w(m) = m$ . In this case, the mass that is ejected back into the ISM is

$$M_{\text{remnant}} = M \int_{-\infty}^{\infty} d \ln m \xi(m) [m - w(m)] \equiv RM, \quad (7.7)$$

where we define  $R$  as the return fraction. The mass fraction that stays locked in remnants is  $1 - R$ .

Of course "long time" here is a vague term. By convention (defined by Tinsley 1980), we choose to take  $w(m) = m$  for  $m = 1 M_{\odot}$ . We take  $w(m) = 0.7 M_{\odot}$  for  $m = 1 - 8 M_{\odot}$  and  $w(m) = 1.4 M_{\odot}$  for  $m > 8 M_{\odot}$ , i.e. to assume that stars from  $1 - 8 M_{\odot}$  leave behind  $0.7 M_{\odot}$  white dwarfs, and stars larger than that mass form  $1.4 M_{\odot}$  neutron stars. If one puts this in for a Chabrier (2005) IMF, the result is  $R = 0.46$ , meaning that these averages are larger by a factor of  $1/0.56$ .

Given this formalism, it is straightforward to use a set of stellar evolutionary tracks plus an IMF to compute  $\langle q/M \rangle$  or  $\langle Q/M \rangle$  for any quantity of interest. Indeed, this is effectively what starburst99 and programs like it do. The quantities of greatest concern for massive star feedback are the bolometric output, ionizing photon output, wind momentum and energy output, and supernova output.

### 7.1.2 Energy- versus Momentum-Driven Feedback

Before discussing individual feedback mechanisms in detail, it is also helpful to lay out two general categories that can be used to understand them. Let us consider a population of stars surrounded by initially-uniform interstellar gas. Those stars eject both photons and baryons (in the form of stellar winds) into the surrounding gas, and these photons and baryons carry both momentum and energy. We want to characterize how the ISM will respond.

One important consideration is that, as we have already shown, it is very hard to raise the temperature of molecular gas (or even dense atomic gas) because it is able to radiate so efficiently. A factor of 10 increase in the radiative heating rate might yield only a tens of percent increase in temperature. This is true as long as the gas is cold and dense, but at sufficiently high temperatures or if the gas is continuously illuminated then the cooling rate begins to drop off, and it is possible for gas to remain hot.

A critical distinction is therefore between mechanisms that are able to keep the gas hot for a time that is long enough to be significant (generally of order the crossing time of the cloud or longer), and

those where the cooling time is much shorter. For the latter case, the energy delivered by the photons and baryons will not matter, only the momentum delivered will. The momentum cannot be radiated away. We refer to feedback mechanism where the energy is lost rapidly as momentum-driven feedback, and to the opposite case where the energy is retained for at least some time as energy-driven, or explosive, feedback.

To understand why the distinction between the two is important, let us consider two extreme limiting cases. We place a cluster of stars at the origin and surround it by a uniform region of gas with density  $\rho$ . At time  $t = 0$ , the stars "turn on" and begin emitting energy and momentum, which is then absorbed by the surrounding gas. Let the momentum and energy injection rates be  $\dot{p}_w$  and  $\dot{E}_w$ ; it does not matter if the energy and momentum are carried by photons or baryons, so long as the mass swept up is significantly greater than the mass carried by the wind.

The wind runs into the surrounding gas and causes it to begin moving radially outward, which in turn piles up material that is further away, leading to an expanding shell of gas. Now let us compute the properties of that shell in the two extreme limits of all the energy being radiated away, and all the energy being kept. If all the energy is radiated away, then at any time the radial momentum of the shell must match the radial momentum injected up to that time, i.e.,

$$p_{\text{sh}} = M_{\text{sh}} v_{\text{sh}} = \dot{p}_w t. \quad (7.8)$$

The kinetic energy of the shell is

$$E = p_{\text{sh}}^2 / 2M_{\text{sh}} = \frac{1}{2} v_{\text{sh}} \dot{p}_w t. \quad (7.9)$$

For comparison, if none of the energy is radiated away, the energy is simply

$$E = \dot{E}_w t. \quad (7.10)$$

Thus the energy in the energy-conserving case is larger by a factor of

$$\frac{1}{v_{\text{sh}}} \cdot \frac{2\dot{E}_w}{\dot{p}_w}. \quad (7.11)$$

If the energy injected by the stars is carried by a wind of baryons, then  $2\dot{E}_w / \dot{p}_w$  is simply the speed of that wind, while if it is carried by photons, then  $2\dot{E}_w / \dot{p}_w = 2c$ . Thus the energy in the energy-conserving case is larger by a factor of  $2c/v_{\text{sh}}$  for a photon wind, and  $v_w/v_{\text{sh}}$  for a baryon wind. These are not small factors: observed expanding shells typically have velocities of at most a few tens of  $\text{km s}^{-1}$ , while wind speeds from massive stars, for example, can be thousands of  $\text{km s}^{-1}$ . Thus it matters a great deal where a particular

feedback mechanism lies between the energy- and momentum-conserving limits.

## 7.2 Momentum-Driven Feedback Mechanisms

We are now ready to consider individual mechanisms by which stars can deliver energy and momentum to the gas around them. Our goal is to understand what forms of feedback are significant and to estimate their relative budgets of momentum and energy.

### 7.2.1 Radiation Pressure and Radiatively-Driven Winds

The simplest form of feedback to consider is radiation pressure. Since the majority of the radiant energy deposited in the ISM will be re-radiated immediately, radiation pressure is (probably) a momentum-driven feedback. To evaluate the momentum it deposits, one need merely evaluate the integrals over the IMF we have written down using the bolometric luminosities of stars. Murray & Rahman (2010) find

$$\left\langle \frac{L}{M} \right\rangle = 1140 L_{\odot} M_{\odot}^{-1} = 2200 \text{ erg g}^{-1}, \quad (7.12)$$

and the corresponding momentum injection rate is

$$\left\langle \frac{p_{\text{rad}}}{M} \right\rangle = \frac{1}{c} \left\langle \frac{L}{M} \right\rangle = 7.3 \times 10^{-8} \text{ cm s}^{-2} = 23 \text{ km s}^{-1} \text{ Myr}^{-1} \quad (7.13)$$

The physical meaning of this expression is that for every gram of matter that goes into stars, those stars produce enough light over 1 Myr to accelerate another gram of matter to a speed of 23 km s<sup>-1</sup>. For very massive stars, radiation pressure also accelerates winds off the star's surfaces; for such stars, the wind carries a bit under half the momentum of the radiation field. Including this factor raises the estimate by a few tens of percent. However, these winds may also be energy conserving, a topic we will approach momentarily.

Integrated over the lifetimes of the stars, out 100 Myr the total energy production is

$$\left\langle \frac{E_{\text{rad}}}{M} \right\rangle = 1.1 \times 10^{51} \text{ erg M}_{\odot}^{-1} \quad (7.14)$$

The majority of this energy is produced in the first  $\sim 5$  Myr of a stellar population's life, when the massive stars live and die.

It is common to quote the energy budget in units of  $c^2$ , which gives a dimensionless efficiency with which stars convert mass into radiation. Doing so gives

$$\epsilon = \frac{1}{c^2} \left\langle \frac{E_{\text{rad}}}{M} \right\rangle = 6.2 \times 10^{-4}. \quad (7.15)$$

The radiation momentum budget is simply this over  $c$ ,

$$\left\langle \frac{p_{\text{rad,tot}}}{M} \right\rangle = 190 \text{ km s}^{-1}. \quad (7.16)$$

This is an interesting number, since it is not all that different than the circular velocity a typical galaxy. It is a suggestion that the radiant momentum output by stars may be interesting in pushing matter around in galaxies – probably not by itself, but perhaps in conjunction with other effects.

### 7.2.2 Protostellar Winds

A second momentum-driven mechanism, that we will discuss in more detail when we get to disks, is protostellar jets. All accretion disks appear to produce some sort of wind that carries away some of the mass and angular momentum, and protostars are no exception. The winds from these stars carry a mass flux of order a few tens of percent of the mass coming into the stars, and eject it with a velocity of order the Keplerian speed at the stellar surface. Note that these winds are distinct from the radiatively-driven ones that come from main sequence O stars. They are very different in both their driving mechanism and physical characteristics.

Why do we expect protostellar winds to be a momentum-driven feedback mechanism instead of an energy-driven one? The key lies in their characteristic speeds. Consider a star of mass  $M_*$  and radius  $R_*$ . Its wind will move at a speed of order

$$v_w \sim \sqrt{\frac{GM_*}{R_*}} = 250 \text{ km s}^{-1} \left( \frac{M_*}{M_\odot} \right)^{1/2} \left( \frac{R_*}{3R_\odot} \right)^{-1/2}, \quad (7.17)$$

where the scalings are for typical protostellar masses and radii. The kinetic energy per unit mass carried by the wind is  $v_w^2/2$ , and when the wind hits the surrounding ISM it will shock and this kinetic energy will be converted to thermal energy. We can therefore find the post-shock temperature from energy conservation. The thermal energy per unit mass is  $(3/2)k_B T / \mu m_H$ , where  $\mu$  is the mean particle mass in H masses. Thus the post-shock temperature will be

$$T = \frac{\mu m_H v_w^2}{3k_B} \sim 5 \times 10^6 \text{ K} \quad (7.18)$$

for the fiducial speed above. This is low enough that gas at this temperature will be able to cool fairly rapidly, leaving us in the momentum-conserving limit.

So how much momentum can we extract? To answer that, we will use our formalism for IMF averaging. Let us consider stars forming over some timescale  $t_{\text{form}}$ . This can be a function of mass if

we want; it doesn't matter. Similarly, let us assume for simplicity that the accretion rate during the formation stage is constant; again, this assumption actually makes no difference to the result, it just makes the math easier. Thus a star of mass  $m$  accretes at a rate  $\dot{m} = m/t_{\text{form}}$  over a time  $t_{\text{form}}$ , and during this time it produces a wind with a mass flux  $f\dot{m}$  that is launched with a speed  $v_K$ . Thus IMF-averaged yield of wind momentum is

$$\left\langle \frac{p_w}{M} \right\rangle = \int_{-\infty}^{\infty} d \ln m \xi(m) \int_0^{t_{\text{form}}} dt \frac{f m v_K}{t_{\text{form}}}. \quad (7.19)$$

In reality  $v_K$ ,  $f$ , and the accretion rate probably vary over the formation time of a star, but to get a rough answer we can assume that they are constant, in which case the integral is trivial and evaluates to

$$\left\langle \frac{p_w}{M} \right\rangle = f v_K \int_{-\infty}^{\infty} d \ln m \xi(m) m = f v_K \quad (7.20)$$

where the second step follows from the normalization of the IMF. Thus we learn that winds supply momentum to the ISM at a rate of order  $f v_K$ . Depending on the exact choices of  $f$  and  $v_K$ , this amounts to a momentum supply of a few tens of  $\text{km s}^{-1}$  per unit mass of stars formed.

Thus in terms of momentum budget, protostellar winds carry over the full lifetimes of the stars that produce them about as much momentum as is carried by the radiation each Myr. Thus if one integrates over the full lifetime of even a very massive, short-lived star, it puts out much more momentum in the form of radiation than it does in the form of outflows. So why worry about outflows at all, in this case?

There are two reasons. First, because the radiative luminosities of stars increase steeply with stellar mass, the luminosity of a stellar population is dominated by its few most massive members. In small star-forming regions with few or no massive stars, the radiation pressure will be much less than our estimate, which is based on assuming full sampling of the IMF, suggests. On the other hand, protostellar winds produce about the same amount of momentum per unit mass accreted no matter what stars is doing the accreting – this is just because  $v_K$  is not a very strong function of stellar mass. (This is a bit of an oversimplification, but it's true enough for this purpose.) This means that winds will be significant even in regions that lack massive stars, because they can be produced by low-mass stars too.

Second, while winds carry less momentum integrated over stars' lifetimes, when they are on they are much more powerful. Typical formation times, we shall see, are of order a few times  $10^5$  yr, so the instantaneous production rate of wind momentum is typically

$\sim 100 \text{ km s}^{-1} \text{ Myr}^{-1}$ , a factor of several higher than radiation pressure. Thus winds can dominate over radiation pressure significantly during the short phase when they are on.

### 7.3 (Partly) Energy-Driven Feedback Mechanisms

#### 7.3.1 Ionizing Radiation

Massive stars produce significant amounts of ionizing radiation. From Murray & Rahman (2010), the yield of ionizing photons from a zero-age population is

$$\left\langle \frac{S}{M} \right\rangle = 6.3 \times 10^{46} \text{ photons s}^{-1} M_{\odot}^{-1}. \quad (7.21)$$

The corresponding lifetime-averaged production of ionizing photons is

$$\left\langle \frac{S_{\text{tot}}}{M} \right\rangle = 4.2 \times 10^{60} \text{ photons M}_{\odot}^{-1} \quad (7.22)$$

*H II Region Expansion* We will not go into tremendous detail on how these photons interact with the ISM, but to summarize: photons capable of ionizing hydrogen will be absorbed with a very short mean free path, producing a bubble of fully ionized gas within which all the photons are absorbed. The size of this bubble can be found by equating the hydrogen recombination rate with the ionizing photon production rate, giving

$$S = \frac{4}{3} \pi r_i^3 n_e n_p \alpha_B, \quad (7.23)$$

where  $r_i$  is the radius of the ionized region,  $n_e$  and  $n_p$  are the number densities of electrons and protons, and  $\alpha_B$  is the recombination rate coefficient for case B, and which has a value of roughly  $3 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1}$ . Cases A and B, what they mean, and how this quantity is computed are all topics discussed in the class on diffuse matter, and here we will simply take  $\alpha_B$  as a known constant. The radius of the ionized bubble is known as the Strömgren radius after the person who first calculated it.

If we let  $\mu_H$  be the mean mass per hydrogen nucleus in the gas, and  $\rho_0$  be the initial density before the photoionizing stars turn on, then  $n_p = \rho_0 / \mu_H$  and  $n_e = 1.1 \rho_0 / \mu_H$ , with the factor of 1.1 coming from assuming that He is singly ionized (since its ionization potential is not that different from hydrogen's) and from a ratio of 10 He nuclei per H nucleus. Inserting these factors and solving for  $r_i$ , we obtain the Strömgren radius, the equilibrium radius of a sphere of

gas ionized by a central source:

$$r_s = \left( \frac{3S\mu_H^2}{4(1.1)\pi\alpha_B\rho_0^2} \right)^{1/3} = 2.8S_{49}^{1/3}n_2^{-2/3} \text{ pc}, \quad (7.24)$$

where  $S_{49} = S/10^{49} \text{ s}^{-1}$ ,  $n_2 = (\rho_0/\mu_H)/100 \text{ cm}^{-3}$ , and we have used  $\alpha_B = 3.46 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1}$ .

The photoionized gas will be heated to  $\sim 10^4 \text{ K}$  by the energy deposited by the ionizing photons. The corresponding sound speed in the ionized gas will be

$$c_i = \sqrt{\frac{k_B T_i}{\mu_H/2.2}} = 11T_{i,4}^{1/2} \text{ km s}^{-1}, \quad (7.25)$$

where  $T_{i,4} = T_i/10^4 \text{ K}$ ,  $\mu_H = 2.3 \times 10^{-24} \text{ g}$  is the mean mass per H nucleus, and the factor of 2.2 arises because there are 2.2 particles per H nucleus, and the extra particles are electrons that are essentially massless. The pressure in the ionized region is  $\rho_0 c_i^2$ , which is generally much larger than the pressure  $\rho_0 c_0^2$  outside the ionized region, where  $c_0$  is the sound speed in the neutral gas. As a result, the ionized region is hugely over-pressured compared to the neutral gas around it. The gas in this region will therefore begin to expand dynamically.

The time to reach ionization balance is short compared to dynamical timescales (as is demonstrated in the diffuse matter class), so we can assume that ionization balance is always maintained as the expansion occurs. Consequently, when the ionized region has reached a radius  $r_i$ , the density inside the ionized region must obey

$$\rho_i = \left( \frac{3S\mu_H^2}{4(1.1)\pi\alpha_B r_i^3} \right)^{1/2}. \quad (7.26)$$

At the start of expansion  $\rho_i = \rho_0$ , but we see here that the density drops as  $r_i^{-3/2}$  as expansion proceeds. Since the expansion is highly supersonic with respect to the external gas (as we will see shortly), there is no time for sound waves to propagate away from the ionization front and pre-accelerate the neutral gas. Instead, this gas must be swept up by the expanding H II region. However, since  $\rho_i \ll \rho_0$ , the mass that is swept up as the gas expands must reside not in the ionized region interior, but in a dense neutral shell at its edges. At late times, when  $r_i \gg r_s$ , we can neglect the mass in the shell interior in comparison to that in the shell, and simply set the shell mass equal to the total mass swept up. We therefore have a shell mass

$$M_{\text{sh}} = \frac{4}{3}\pi\rho_0 r_i^3. \quad (7.27)$$

We can write down the equation of motion for this shell. If we neglect the small ambient pressure, then the only force acting on the shell is the pressure  $\rho_i c_i^2$  exerted by ionized gas in the H II region interior. Conservation of momentum therefore requires that

$$\frac{d}{dt} (M_{\text{sh}} \dot{r}_i) = 4\pi r_i^2 \rho_i c_i^2. \quad (7.28)$$

Rewriting everything in terms of  $r_i$ , we arrive at an ODE for  $r_i$ :

$$\frac{d}{dt} \left( \frac{1}{3} r_i^3 \dot{r}_i \right) = c_i^2 r_i^2 \left( \frac{r_i}{r_S} \right)^{-3/2}, \quad (7.29)$$

where we have used the scaling  $\rho_i = \rho_0 (r_i/r_S)^{-3/2}$ .

This ODE is straightforward to solve numerically, but if we focus on late times when  $r_i \gg r_S$ , we can solve it analytically. For  $r_i \gg r_S$ , we can take  $r_i \approx 0$  as  $t \rightarrow 0$ , and with this boundary condition the ODE cries out for a similarity solution. As a trial, consider  $r_i = f r_S (t/t_S)^\eta$ , where

$$t_S = \frac{r_S}{c_i} = 240 S_{49}^{1/3} n_2^{-2/3} T_{i,4}^{-1/2} \text{ kyr} \quad (7.30)$$

and  $f$  is a dimensionless constant. Substituting this trial solution in, there are numerous cancellations, and in the end we obtain

$$\frac{1}{4} \eta (4\eta - 1) f^4 \left( \frac{t}{t_S} \right)^{4\eta-2} = f^{1/2} \left( \frac{t}{t_S} \right)^{\eta/2}. \quad (7.31)$$

Clearly we can obtain a solution only if  $4\eta - 2 = \eta/2$ , which requires  $\eta = 4/7$ . Solving for  $f$  gives  $f = (49/12)^{2/7}$ . We therefore have a solution

$$r_i = r_S \left( \frac{7t}{2\sqrt{3}t_S} \right)^{4/7} = 9.4 S_{49}^{1/7} n_2^{-2/7} T_{i,4}^{2/7} t_6^{4/7} \text{ pc} \quad (7.32)$$

at late times, where  $t_6 = t/1$  Myr.

*Feedback Effects of H II Regions* Given this result, what can we say about the effects of an expanding H II region? There are several possible effects: ionization can eject mass, drive turbulent motions, and possibly even disrupt clouds entirely. First consider mass ejection. In our simple calculation, we have taken the ionized gas to be trapped inside a spherical H II region interior. In reality, though, once the H II region expands to the point where it encounters a low density region at a cloud edge, it will turn into a “blister” type region, and the ionized gas will freely escape into the low density medium. The mass flux carried in this ionized wind will be roughly

$$\dot{M} = 4\pi r_i^2 \rho_i c_i, \quad (7.33)$$

i.e. the area from which the wind flows times the characteristic density of the gas at the base of the wind times the characteristic speed of the wind. Substituting in our similarity solution, we have

$$\dot{M} = 4\pi r_S^2 \rho_0 c_i \left( \frac{7t}{2\sqrt{3}t_S} \right)^{2/7} = 7.2 \times 10^{-3} t_6^{2/7} S_{49}^{4/7} n_2^{-1/7} T_{i,4}^{1/7} M_\odot \text{ yr}^{-1}. \quad (7.34)$$

We therefore see that, over the roughly 3 – 4 Myr lifetime of an O star, it can eject  $\sim 10^3 - 10^4 M_\odot$  of mass from its parent cloud, provided that cloud is at a relatively low density (i.e.  $n_2$  is not too big). Thus massive stars can eject many times their own mass from a molecular cloud. In fact, we can use this effect to make an estimate of the star formation efficiency in GMCs – see Matzner (2002).

We can also estimate the energy contained in the expanding shell. This is

$$E_{\text{sh}} = \frac{1}{2} M_{\text{sh}} \dot{r}_i^2 = \frac{32}{147} \pi \rho_0 \frac{r_S^5}{t_S^2} \left( \frac{7t}{2\sqrt{3}t_S} \right)^{6/7} = 8.1 \times 10^{47} t_6^{6/7} S_{49}^{5/7} n_2^{-10/7} T_{i,4}^{10/7} \text{ erg}. \quad (7.35)$$

For comparison, the gravitational binding energy of a  $10^5 M_\odot$  GMC with a surface density of  $0.03 \text{ g cm}^{-2}$  is  $\sim 10^{50}$  erg. Thus a single O star's H II region provides considerably less energy than this. On the other hand, the collective effects of  $\sim 10^2$  O stars, with a combined ionizing luminosity of  $10^{51} \text{ s}^{-1}$  or so, can begin to produce H II regions whose energies rival the binding energies of individual GMCs. This means that H II region shells may sometimes be able to unbind GMCs entirely. Even if they cannot, they may be able to drive significant turbulent motions within GMCs.

We can also compute the momentum of the shell, for comparison to the other forms of feedback we discussed previously. This is

$$p_{\text{sh}} = M_{\text{sh}} \dot{r}_i = 1.1 \times 10^5 n_2^{-1/7} T_{i,4}^{-8/7} S_{49}^{4/7} t_6^{9/7} M_\odot \text{ km s}^{-1}. \quad (7.36)$$

Since this is non-linear in  $S_{49}$  and in time, the effects of HII regions will depend on how the stars are clustered together, and how long they live. To get a rough estimate, though, we can take the typical cluster to have an ionizing luminosity around  $10^{49}$ , since by number most clusters are small, and we can adopt an age of 4 Myr. This means that (also using  $n_2 = 1$  and  $T_{i,4} = 1$ ) the momentum injected per  $10^{49}$  photons  $\text{s}^{-1}$  of luminosity is  $p = 3 - 5 \times 10^5 M_\odot \text{ km s}^{-1}$ . Recalling that we get  $6.3 \times 10^{46}$  photons  $\text{s}^{-1} M_\odot^{-1}$  for a zero-age population, this means that the momentum injection rate for HII regions is roughly

$$\left\langle \frac{\dot{p}_{\text{HII}}}{M} \right\rangle \sim 3 \times 10^3 \text{ km s}^{-1}. \quad (7.37)$$

This is obviously a very rough calculation, and it can be done with much more sophistication, but this analysis suggests that HII regions are likely the dominant feedback mechanism compared to winds and HII regions.

There is one important caveat to make, though. Although in the similarity solution we formally have  $v_i \rightarrow \infty$  as  $r_i \rightarrow 0$ , in reality the ionized region cannot expand faster than roughly the ionized gas sound speed: one cannot drive a  $100 \text{ km s}^{-1}$  expansion using gas with a sound speed of  $10 \text{ km s}^{-1}$ . As a result, all of these effects will not work in any cluster for which the escape speed or the virial velocity exceeds  $\sim 10 \text{ km s}^{-1}$ . This is not a trivial limitation, since for very massive star clusters the escape speed can exceed this value. An example is the R136 cluster in the LMC, which has a present-day stellar mass of  $5.5 \times 10^4 M_\odot$  inside a radius of 1 pc. The escape speed from the stars alone is roughly  $20 \text{ km s}^{-1}$ . Assuming there was gas in the past when the cluster formed, the escape speed must have been even higher. For a region like this, H II regions cannot be important.

### 7.3.2 Stellar Winds

Next let us consider the effects of stellar winds. As we alluded to earlier, O stars launch winds with velocities of  $v_w \sim 1000 - 2500 \text{ km s}^{-1}$  and mass fluxes of  $\dot{M}_w \sim 10^{-7} M_\odot \text{ yr}^{-1}$ . We have already seen that the momentum carried by these winds is fairly unimportant in comparison to the momentum of the protostellar outflows or the radiation field, let alone the momentum provided by HII regions. However, because of the high wind velocities, repeating the analysis we performed for protostellar jets yields a characteristic post-shock temperature that is closer to  $10^8 \text{ K}$  than  $10^6 \text{ K}$ . Gas at such high temperatures has a very long cooling time, so we might end up with an energy-driven feedback. We therefore consider that case next.

Since the winds are radiatively driven, they tend to carry momenta comparable to that carried by the stellar radiation field. The observed correlation between stellar luminosity and wind momentum (e.g. Repolust et al. 2004) is that

$$\dot{M}_w v_w \approx 0.5 \frac{L_*}{c}, \quad (7.38)$$

where  $L_*$  is the stellar luminosity. This implies that the mechanical luminosity of the wind is

$$L_w = \frac{1}{2} \dot{M}_w v_w^2 = \frac{L_*^2}{8\dot{M}_w c^2} = 850 L_{*,5}^2 \dot{M}_{w,-7}^{-1} L_\odot. \quad (7.39)$$

This is not much compared to the star's radiant luminosity, but that radiation will mostly not do into pushing the ISM around. The wind,

on the other hand might. Also notice that over the integrated power output is

$$E_w = L_w t = 1.0 \times 10^{50} L_{*,5}^2 \dot{M}_{w,-7}^{-1} t_6 \text{ erg}, \quad (7.40)$$

so over the  $\sim 4$  Myr lifetime of a massive star, the total mechanical power in the wind is not much less than the amount of energy released when the star goes supernova.

If energy is conserved, and we assume that about half the available energy goes into the kinetic energy of the shell and half is in the hot gas left in the shell interior, conservation of energy then requires that

$$\frac{d}{dt} \left( \frac{2}{3} \pi \rho_0 r_b^3 \dot{r}_b^2 \right) \approx \frac{1}{2} L_w. \quad (7.41)$$

As with the H II region case, this cries out for similarity solution.

Letting  $r_b = At^\eta$ , we have

$$\frac{4}{3} \pi \eta^2 (5\eta - 2) \rho_0 A^5 t^{5\eta-3} \approx L_w. \quad (7.42)$$

Clearly we must have  $\eta = 3/5$  and  $A = [25L_w/(12\pi\rho_0)]^{1/5}$ . Putting in some numbers,

$$r_b = 16 L_{*,5}^{2/5} \dot{M}_{w,-7}^{-1/5} n_2^{-1/5} t_6^{3/5} \text{ pc}. \quad (7.43)$$

Note that this is greater than the radius of the comparable H II region, so the wind will initially move faster and drive the H II region into a thin ionized layer between the hot wind gas and the outer cool shell – *if the energy-driven limit is correct*. A corollary of this is that the wind would be even more effective than the ionized gas at ejecting mass from the cloud.

However, this may not be correct, because this solution assumes that the energy carried by the wind will stay confined within a closed shell. This may not be the case: the hot gas may instead break out and escape, imparting relatively little momentum. Whether this happens or not is difficult to determine theoretically, but can be addressed by observations. In particular, if the shocked wind gas is trapped inside the shell, it should produce observable x-ray emission. We can quantify how much x-ray emission we should see with a straightforward argument. It is easiest to phrase this argument in terms of the pressure of the x-ray emitting gas, which is essentially what an x-ray observation measures.

Consider an expanding shell of matter that began its expansion a time  $t$  ago. In the energy-driven case, the total energy within that shell is, up to factors of order unity,  $E_w = L_w t$ . The pressure is simply  $2/3$  of the energy density (since the gas is monatomic at these temperatures). Thus,

$$P_X = \frac{2E_w}{3[(4/3)\pi r^3]} = \frac{L_{*,5}^2 t}{16\pi \dot{M}_w c^2 r^3}. \quad (7.44)$$

It is useful to compute the ratio of this to the pressure exerted by the radiation, which is simply twice that exerted by the wind in the momentum-driven limit. This is

$$P_{\text{rad}} = \frac{L_*}{4\pi r^2 c}. \quad (7.45)$$

We define this ratio as the trapping factor:

$$f_{\text{trap}} = \frac{P_X}{P_{\text{rad}}} = \frac{L_* t}{4\dot{M}_w c r} \approx \frac{L_*}{4\dot{M}_w c v}, \quad (7.46)$$

where in the last step we used  $v \approx r/t$ , where  $v$  is the expansion velocity of the shell. If we now use the relation  $\dot{M}_w v_w \approx (1/2)L_*/c$ , we finally arrive at

$$f_{\text{trap}} \approx \frac{v_w}{2v}. \quad (7.47)$$

Thus if shells expand in the energy-driven limit due to winds, the pressure of the hot gas within them should exceed the direct radiation pressure by a factor of roughly  $v_w/v$ , where  $V$  is the shell expansion velocity and  $v_w$  is the wind launch velocity. In contrast, the momentum driven gas gives  $P_X/P_{\text{rad}} \sim 1/2$ , since the hot gas exerts a force that is determined by the wind momentum, which is roughly has the momentum carried by the stellar radiation field.

[Lopez et al. \(2011\)](#) observed the 30 Doradus H II region, which is observed to be expanding with  $v \approx 20 \text{ km s}^{-1}$ , giving a predicted  $f_{\text{trap}} = 20$  for a conservative  $v_w = 1000 \text{ km s}^{-1}$ . They then measured the hot gas pressure from the x-rays and the direct radiation pressure from the stars optical emission. The result is that  $f_{\text{trap}}$  is much closer to 0.5 than 20 for 30 Doradus, indicating that the momentum-driven solution is closer to reality there. [Harper-Clark & Murray \(2009\)](#) reached a similar conclusion about the Carina Nebula.

### 7.3.3 Supernovae

We can think of the energy and momentum budget from supernovae as simply representing a special case of the lifetime budgets we've computed. In this case, we can simply think of  $q(M, t)$  as being a  $\delta$  function: all the energy and momentum of the supernova is released in a single burst at a time  $t = t_l(m)$ , where  $t_l(m)$  is the lifetime of the star in question. We normally assume that the energy yield per star is  $10^{51} \text{ erg}$ , and have to make some estimate of the minimum mass at which a SN will occur, which is roughly  $8 M_\odot$ . We can also, if we want, imagine mass ranges where other things happen, for example direct collapse to black hole, pair instability supernova that produce more energy, or something more exotic. These choices usually don't make much difference, though, because they affect very massive stars, and since the supernova energy yield (unlike the luminosity) is not

a sharp function of mass, the relative rarity of massive stars means they make a small contribution. Thus it is usually safe to ignore these effects.

Given this preamble, we can write the approximate supernova energy yield per unit mass as

$$\left\langle \frac{E_{\text{SN}}}{M} \right\rangle = E_{\text{SN}} \int_{m_{\text{min}}}^{\infty} d \ln m \xi(m) \equiv E_{\text{SN}} \left\langle \frac{N_{\text{SN}}}{M} \right\rangle, \quad (7.48)$$

where  $E_{\text{SN}} = 10^{51}$  erg is the constant energy per SN, and  $m_{\text{min}} = 8 M_{\odot}$  is the minimum mass to have a supernova. Note that the integral, which we have named  $\langle N_{\text{SN}}/M \rangle$ , is simply the number of stars above  $m_{\text{min}}$  per unit mass in stars total, which is just the expected number of supernovae per unit mass of stars. For a Chabrier IMF from  $0.01 - 120 M_{\odot}$ , we have

$$\left\langle \frac{N_{\text{SN}}}{M} \right\rangle = 0.011 M_{\odot}^{-1} \quad \left\langle \frac{E_{\text{SN}}}{M} \right\rangle = 1.1 \times 10^{49} \text{ erg } M_{\odot}^{-1} = 6.1 \times 10^{-6} c^2. \quad (7.49)$$

A more detailed calculation from starburst99 agrees very well with this crude estimate. Note that this, plus the Milky Way's SFR of  $\sim 1 M_{\odot} \text{ yr}^{-1}$ , is the basis of the oft-quoted result that we expect  $\sim 1$  supernova per century in the Milky Way.

The momentum yield from SN can be computed in the same way. This is slightly more uncertain, because it is easier to measure the SN energy than its momentum – the latter requires the ability to measure the velocity or mass of the ejecta before they are mixed with significant amounts of ISM. However, roughly speaking the ejection velocity is  $v_{\text{ej}} \approx 10^9 \text{ cm s}^{-1}$ , which means that the momentum is  $p_{\text{SN}} = 2E_{\text{SN}}/v_{\text{ej}}$ . Adopting this value, we have

$$\left\langle \frac{p_{\text{SN}}}{M} \right\rangle = \frac{2}{v_{\text{ej}}} \left\langle \frac{E_{\text{SN}}}{M} \right\rangle = 55 v_{\text{ej},9}^{-1} \text{ km s}^{-1}. \quad (7.50)$$

Physically, this means that every  $M_{\odot}$  of matter that goes into stars provides enough momentum to raise another  $M_{\odot}$  of matter to a speed of  $55 \text{ km s}^{-1}$ . This is not very much compared to other feedbacks, but of course supernovae, like stellar winds, may have an energy-conserving phase where their momentum deposition grows. We will discuss the question of supernova momentum deposition more in the next few classes as we discuss models for regulation of the star formation rate.



## **Part III**

# **Star Formation Processes and Problems**



*A*

*Solutions to Problem Sets*



# Solutions to Problem Set 1

## 1. Molecular Tracers.

(a) The radiative de-excitation rate is

$$\left( \frac{dn_i}{dt} \right)_{\text{spon. emiss.}} = -n_i \sum_{j < i} A_{ij}.$$

The collisional de-excitation rate is

$$\left( \frac{dn_i}{dt} \right)_{\text{coll.}} = -nn_i \sum_{j < i} k_{ij}.$$

(b) Setting the results from the previous part equal and solving, we obtain

$$n_i \sum_{j < i} A_{ij} = n_{\text{crit}} n_i \sum_{j < i} k_{ij} \implies n_{\text{crit}} = \frac{\sum_{i < j} A_{ij}}{\sum_{i < j} k_{ij}}.$$

(c) Using numbers taken from the LAMBDA website for the  $A_{ij}$  and  $\gamma_{ij}$  values, we have

Line	$n_{\text{crit}}$ [cm $^{-3}$ ]
CO( $J = 1 \rightarrow 0$ )	$2.2 \times 10^3$
CO( $J = 3 \rightarrow 2$ )	$1.9 \times 10^4$
CO( $J = 5 \rightarrow 4$ )	$7.6 \times 10^4$
HCN( $J = 1 \rightarrow 0$ )	$1.0 \times 10^6$

(d) The fraction of the mass above some specified density  $\rho_c$  can be obtained by integrating the PDF for mass:

$$f_M(\rho > \rho_0) = \frac{\int_{s_c}^{\infty} p_M(s) ds}{\int_{-\infty}^{\infty} p_M(s) ds} \quad (\text{A.1})$$

where  $s_c = \ln(\rho_c / \bar{\rho})$  and the mass PDF is

$$p_M = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp \left[ -\frac{(s + s_0)^2}{2\sigma_s^2} \right] \quad (\text{A.2})$$

with  $s_0 = -\sigma_s^2 / 2$ . Using the critical densities obtained in the previous part to compute  $s_c$ , and then to evaluate the integral, we obtain

Line	$s_c$	$f_M(n > n_{\text{crit}})$
CO( $J = 1 \rightarrow 0$ )	3.1	0.39
CO( $J = 3 \rightarrow 2$ )	5.2	0.11
CO( $J = 5 \rightarrow 4$ )	6.6	0.032
HCN( $J = 1 \rightarrow 0$ )	9.2	0.0013

It appears that CO( $J = 1 \rightarrow 0$ ) and (to some extent) CO( $J = 3 \rightarrow 2$ ) are good tracers of the bulk of the mass, while CO( $J = 5 \rightarrow 4$ ) and HCN( $J = 1 \rightarrow 0$ ) are better tracers of the denser parts of the cloud.

## 2. Inferring Star Formation Rates in the Infrared.

- (a) This problem can be done by using the default parameters with `starburst99` and writing out the bolometric luminosity on a logarithmic grid from 0.1 Myr to 1 Gyr, for continuous star formation at a rate of  $1 M_\odot \text{ yr}^{-1}$ . Taking the output luminosities, the results are

$$\begin{aligned} \text{SFR}[M_\odot \text{ yr}^{-1}] &= 4.3 \times 10^{-44} L_{\text{tot}}[\text{erg s}^{-1}] & (10 \text{ Myr}) \\ \text{SFR}[M_\odot \text{ yr}^{-1}] &= 2.9 \times 10^{-44} L_{\text{tot}}[\text{erg s}^{-1}] & (100 \text{ Myr}) \\ \text{SFR}[M_\odot \text{ yr}^{-1}] &= 2.2 \times 10^{-44} L_{\text{tot}}[\text{erg s}^{-1}] & (1 \text{ Gyr}). \end{aligned}$$

In comparison, the corresponding coefficient given by Kennicutt (1998) is  $3.9 \times 10^{-44}$ , the same to within a factor of 2.

- (b) The plot of the `starburst99` output is shown in Figure A.1. The solid line is the output with a normal IMF, and the dashed line is the output with a top-heavy IMF, for part (c).

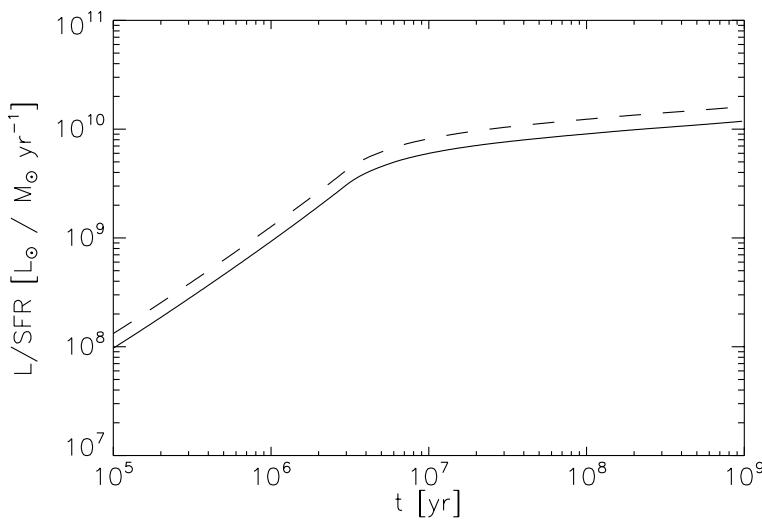


Figure A.1: Luminosity normalized by star formation rate for a normal IMF (solid line) and a top-heavy IMF (dashed line).

- (c) To generate this IMF, I told starbust99 to use a 1 section IMF with a slope of  $-2.3$  running from  $0.5$  to  $100 M_{\odot}$ . At equal ages, the numbers change to

$$\text{SFR}[M_{\odot} \text{ yr}^{-1}] = 3.2 \times 10^{-44} L_{\text{tot}}[\text{erg s}^{-1}] \quad (\text{10 Myr})$$

$$\text{SFR}[M_{\odot} \text{ yr}^{-1}] = 2.1 \times 10^{-44} L_{\text{tot}}[\text{erg s}^{-1}] \quad (\text{100 Myr})$$

$$\text{SFR}[M_{\odot} \text{ yr}^{-1}] = 1.6 \times 10^{-44} L_{\text{tot}}[\text{erg s}^{-1}] \quad (\text{1 Gyr}).$$

These are a few tens of percent lower, because the IMF contains fewer low mass stars that contribute little light. The effect is mild, but that is partly because the change in IMF is mild. These results do suggest that the IR to SFR conversion does depend on the IMF.



## *Bibliography*

- Arzoumanian, D., André, P., Didelon, P., et al. 2011, *Astron. & Astrophys.*, 529, L6
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, *Annu. Rev. Astron. Astrophys.*, 48, 339
- Bonnor, W. B. 1956, *Mon. Not. Roy. Astron. Soc.*, 116, 351
- Chabrier, G. 2003, *Proc. Astron. Soc. Pac.*, 115, 763
- Chabrier, G. 2005, in *Astrophysics and Space Science Library*, Vol. 327, *The Initial Mass Function 50 Years Later*, ed. E. Corbelli, F. Palla, & H. Zinnecker (Dordrecht: Springer), 41–+
- Champagne, F. H. 1978, *J. Fluid Mech.*, 86, 67
- Chandrasekhar, S., & Fermi, E. 1953, *Astrophys. J.*, 118, 116
- Crutcher, R. M. 2012, *Annu. Rev. Astron. Astrophys.*, 50, 29
- Dale, J. E., Ngoumou, J., Ercolano, B., & Bonnell, I. A. 2014, *Mon. Not. Roy. Astron. Soc.*, 442, 694
- Draine, B. T. 2003, *Annu. Rev. Astron. Astrophys.*, 41, 241
- Dunham, M. M., Stutz, A., Allen, L. E., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning (University of Arizona Press), arXiv:1401.1809
- Ebert, R. 1955, *Zeitschrift fur Astrophysics*, 37, 217
- Federrath, C. 2013, *Mon. Not. Roy. Astron. Soc.*, 436, 1245
- Glover, S. C. O., Federrath, C., Mac Low, M., & Klessen, R. S. 2010, *Mon. Not. Roy. Astron. Soc.*, 404, 2
- Harper-Clark, E., & Murray, N. 2009, *Astrophys. J.*, 693, 1696
- Imara, N., Bigiel, F., & Blitz, L. 2011, *Astrophys. J.*, 732, 79

- Jeans, J. H. 1902, Royal Society of London Philosophical Transactions Series A, 199, 1
- Kennicutt, R. C., & Evans, N. J. 2012, *Annu. Rev. Astron. Astrophys.*, 50, 531
- Kennicutt, Jr., R. C. 1992, *Astrophys. J.*, 388, 310
- Kolmogorov, A. 1941, Akademii Nauk SSSR Doklady, 30, 301
- Kolmogorov, A. N. 1991, Royal Society of London Proceedings Series A, 434, 9
- Kroupa, P. 2001, *Mon. Not. Roy. Astron. Soc.*, 322, 231
- . 2002, *Science*, 295, 82
- Krumholz, M. R. 2014, *Physics Reports*, 539, 49
- Krumholz, M. R., & Tan, J. C. 2007, *Astrophys. J.*, 654, 304
- Krumholz, M. R., Bate, M. R., Arce, H. G., et al. 2014, *Protostars and Planets VI*, 243
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *Astrophys. J. Supp.*, 123, 3
- Li, P. S., McKee, C. F., Klein, R. I., & Fisher, R. T. 2008, *Astrophys. J.*, 684, 380
- Lombardi, M., Alves, J., & Lada, C. J. 2006, *Astron. & Astrophys.*, 454, 781
- Lopez, L. A., Krumholz, M. R., Bolatto, A. D., Prochaska, J. X., & Ramirez-Ruiz, E. 2011, *Astrophys. J.*, 731, 91
- McKee, C. F., & Zweibel, E. G. 1992, *Astrophys. J.*, 399, 551
- Murray, N., Quataert, E., & Thompson, T. A. 2010, *Astrophys. J.*, 709, 191
- Offner, S. S. R., Hennebelle, P., Whitworth, A. P., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning (University of Arizona Press), arXiv:1312.5326
- Ossenkopf, V., & Mac Low, M.-M. 2002, *Astron. & Astrophys.*, 390, 307
- Padoan, P., & Nordlund, Å. 1999, *Astrophys. J.*, 526, 279
- Ridge, N. A., Di Francesco, J., Kirk, H., et al. 2006, *Astron. J.*, 131, 2921

- Schinnerer, E., Meidt, S. E., Pety, J., et al. 2013, *Astrophys. J.*, **779**, 42
- Schöier, F. L., van der Tak, F. F. S., van Dishoeck, E. F., & Black, J. H. 2005, *Astron. & Astrophys.*, **432**, 369
- Shu, F. H. 1992, Physics of Astrophysics, Vol. II (University Science Books)
- Smith, M. D., & Mac Low, M.-M. 1997, *Astron. & Astrophys.*, **326**, 801
- Sun, K., Kramer, C., Ossenkopf, V., et al. 2006, *Astron. & Astrophys.*, **451**, 539
- Tafalla, M., Santiago, J., Johnstone, D., & Bachiller, R. 2004, *Astron. & Astrophys.*, **423**, L21
- Taylor, G. I. 1964, Low-Reynolds-Number Flows, National Committe for Fluid Mechanics Films, <https://www.youtube.com/watch?v=51-6QCJTAjU&list=PLoEC6527BE871ABA3&index=7>
- Tomisaka, K. 1998, *Astrophys. J.*, **502**, L163+
- Vázquez, G. A., & Leitherer, C. 2005, *Astrophys. J.*, **621**, 695