I Introduction to the Interstellar Medium

"The Interstellar Medium is anything not in stars"

Donald Osterbrock
[1984 Jan 13 – author's UCSC course notes].

Astronomy 871 is a survey of the physics of the Interstellar Medium (henceforth ISM) of the Milky Way Galaxy. The emphasis will be on physical processes that prevail in each of the distinct phases of interstellar gas. In the end, you should have a thorough grounding in the basic physical principles necessary to pursue research in the field. As such, we will spend little time on current research results or questions, except insofar as they will illustrate the physics of interest to us here. This is not to say that the current research is unimportant, but rather that we are concerned with bringing you up to speed on the basic conceptual tools of that research.

The course is divided into the following parts:

1. Introduction to the ISM, a basic physical description of the ISM, and an overview of global ISM models (the multi-phase ISM).
2. Physics of neutral Hydrogen (HI) regions, concentrating on UV and visible absorption lines and the HI 21-cm hyperfine structure line.
3. Physics of ionized Hydrogen (HII) regions, with an emphasis on the physics of ionization and thermal equilibrium, recombination lines, nebular continuum, and spectral diagnostics of density, temperature, and abundances
4. Interstellar Dust, with an emphasis on interstellar extinction, spectral properties of dusty regions, the thermal properties of grains, grain formation and destruction, and depletion of gas-phase elements onto grains.
5. Physics of Molecular Clouds (H2), including mm and radio rotational line formation (esp. CO), diagnostics of cloud kinematics and structure, molecular hydrogen physics, especially UV and IR lines, H2 formation, destruction, and Photodissociation regions (PDRs).
6. The Hot ISM, primarily low-density, very hot X-ray gas seen in soft-X-ray continuum emission and OVI absorption lines.

Because of the nature of this course, I will be taking a somewhat broad approach to each of these topics, delving into detail only where it offers important physical insights into the processes involved. The syllabus is very ambitious, and I often add or delete material as time and the interests of the class permit. A number of recurring themes will become clear as the course progresses, for example, in each section we will see a different application of radiative transfer and encounter the processes of excitation, emission, and absorption operating in very different physical regimes. The central theme is that the ISM is a system far from local thermodynamic equilibrium and extremely rich in interesting astrophysics.

I-1 Historical Overview

Clouds in Space (pre 20th Century)

Before the 20th century, the Milky Way (which was equated by many astronomers with the entire Universe) was thought to consist of stars in a vacuum. The Herschels and others realized that there were constituents other than stars:

Bright Nebulae, bright "clouds" of gas that do not resolve into individual stars when viewed at high magnifications. These are roughly divided into diffuse nebulae (like Orion or the reflection nebulae around the Pleiades stars), planetary nebulae surrounding faint blue stars, and filamentary nebulae. The latter were later recognized to be supernova remnants. Diffuse nebulae have spectra dominated by either bright emission lines (work of William Huggins in the 1860s & James Keeler in the 1890s), or a reflected stellar absorption-line spectrum ("reflection nebulae"). Vesto Slipher of Lowell Observatory obtained a spectrum of the Pleiades reflection nebulosity in the early 20th century and found it to be reflected starlight, leading to his speculation that it was reflection from "small particles"

Dark Nebulae that were originally thought to be holes in the star clouds (in Hershel's description "*ein loch in Himmel*"), but that were later recognized to be dark clouds of obscuring material seen in silhouette against rich star fields. These are especially prominent in the brightest regions of the Milky Way (e.g., the Great Rift in Cygnus or the Coal Sack and associated clouds in the Southern Milky Way). Many were cataloged by E.E. Barnard who made the first systematic photographic survey of dark nebulae.

Figure I-1: Dark clouds towards the Milky Way in Sagittarius. [Credit: John P. Gleason, Celestial Images]

In general, however, these objects were viewed as isolated clouds in otherwise mostly empty space, and not as a manifestation of a general interstellar medium.

The Diffuse ISM (early 20th century)

The first observational evidence that there was a general ISM that pervaded the space between the stars came from photographic spectroscopy of spectroscopic binary stars early in the 20th century. It was noticed that in addition to the relatively broad absorption lines associated with the atmospheres of the stars, there were very narrow lines that appeared stationary with respect to the cyclical orbital Doppler shifts of the stellar lines. Hartmann first identified the "stationary lines" of CaII H and K (λ3933 and 3968Å rest) in the spectrum of the spectroscopic binary star δ Ori in 1904 (the stellar atmosphere lines showed the familiar orbital Doppler shifts, but the narrow stationary lines, as their name implies, did not move). Another strong stationary lines is the NaI "D" doublet (λ5890 and 5896Å). The stationary lines were particularly noticeable in the spectra of bright O and B stars where they stood out as being very narrow in width compared to the strong, pressure-broadened lines from the stellar atmospheres). The stationary lines had the following properties:

1. The lines are very narrow, with equivalent widths of a few mÅ.
2. Their strengths are correlated with distance: lines from more distant stars were stronger as shown by Plaskett & Pierce in the 1930s.
3. Their radial velocities (relative to the local standard of rest) showed a sine-wave pattern with galactic longitude consistent with galactic rotation.

In the 1930s, an important advance was the discovery of the Trumpler Effect by Lick Observatory astronomer Robert J. Trumpler. He found that the apparent diameters of star clusters got smaller slower than their overall brightnesses were diminishing (put another way, their apparent angular-diameter distances were systematically smaller than their photometric "luminosity" distances). This implied that there was a general absorbing medium that extinguished starlight; hence the clusters grew dimmer faster with distance than they would in a vacuum. This demonstration of the existence of a general diffuse and *absorbing* ISM had far-reaching implications for contemporaneous attempts to measure the distances to stars from brightnesses, and for galactic structure (e.g., Shapley greatly overestimated the size of the Milky Way because he assumed interstellar extinction was unimportant).

In the early years of the 20th century the prevailing view was that the ISM was a homogeneous and diffuse, pervading space with near-constant density. This view was challenged in the late 1930s by the first high-resolution spectroscopy of bright stars that revealed that the so-called "stationary" lines were resolving into many narrower line components with different radial velocities. The complexity of the observed gas kinematics led to the recognition that the ISM was not homogeneous, but instead it was clumpy and organized into "clouds". The ISM suddenly became dynamic, with a kinetic energy density comparable to the thermal energy density.

The HI/HII Region Dichotomy (1940's)

In 1939, Bengt Strömgren developed the idea that the bright diffuse nebulae with strong emission-line spectra were regions of photoionized gas surrounding hot stars. Such idealized *Strömgren Spheres* are at the heart of our modern theory of ionized nebulae and will be discussed extensively in Chapter 3 on the physics of ionized Hydrogen regions.

The traditional nomenclature can be somewhat confusing. Spectroscopically, HII refers generically to Hydrogen in its first ionic state (H^+). Ionized Hydrogen is found throughout the Universe in a wide variety of forms. The term "HII Region", however, specifically refers to a bright diffuse nebulae photoionized by young O and B stars in regions of recent massive star formation. Osterbrock has

argued that when speaking generically of ionized Hydrogen regions we should call them "H^+ Regions" or "Ionized Hydrogen Regions" and reserve the term "HII Regions" for the ionized nebulae around hot young stars. Long-established practice and tradition, however, mean that we are generally stuck with the confusion. Beware.

At least three generic types of ionized nebulae are recognized in the ISM. Note that these are generally isolated objects, and not to be confused with the ionized phases of the general ISM that we will discuss later.

HII Regions are the classical "diffuse nebulae" described by Herschel and others that are characterized by strong emission-line spectra at all wavelengths and a weak continuum. These are regions of interstellar gas heated and ionized by UV ($h\nu \geq 13.6$eV) photons from the atmospheres of embedded O and B stars.

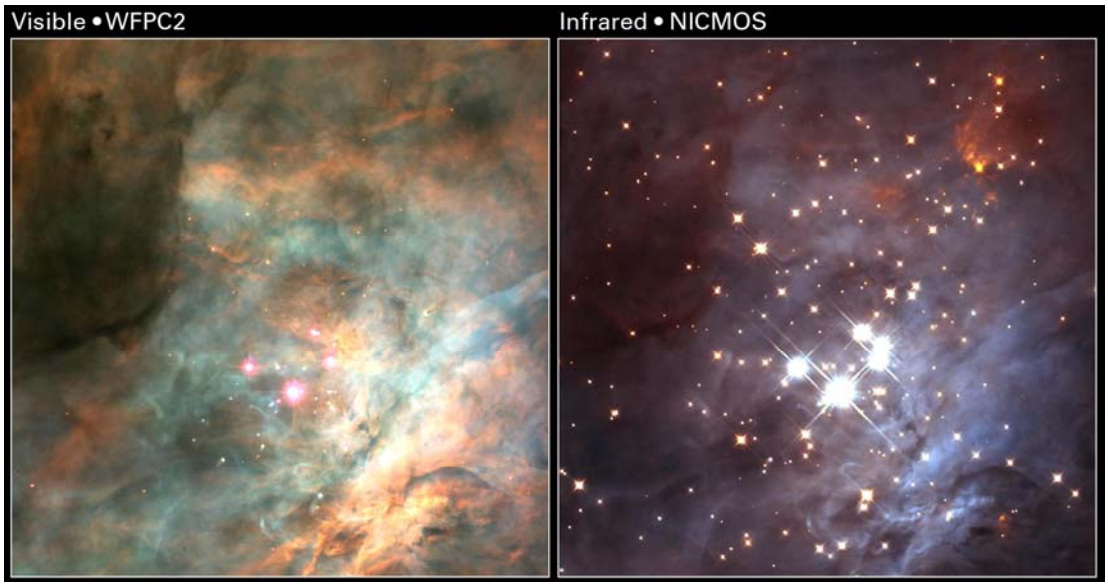

Figure I-2: HST Visible light (left) and Near-IR 1–2μm (right) images of the core of the Orion Nebula, a classical bright HII region [Credit: NASA and K. Luhman – STScI-PRC00-19]

Planetary Nebulae are the UV photoionized ejected stellar envelopes surrounding hot remnant stellar cores (on their way to becoming white dwarfs). While the excitation physics of PNe are similar to those of HII regions, they are the result of different physical processes (stellar death vs. stellar birth).

Figure I-3: Planetary Nebulae [Credit: Hubble Heritage Project, NASA/AURA/STScI, and NOAO (center)].

Supernova Remnants (SNRs). These are regions ionized by the passage of the shockwave expanding outwards from a supernova explosion into the ISM (either Type I or Type II supernovae). They differ from HII regions and PNe in the source of ionizing photons and the additional mechanical heating due to the hydrodynamical shocks. Two basic types of SNRs are recognized:

Young SNRs (e.g., Crab Nebula) are photoionized by UV synchrotron radiation emitted by relativistic electrons accelerated by the wind coming from the central pulsar. These are also often called "pulsar wind nebulae" or "Plerions" (from the greek "πλερες" meaning "full"). The Crab Nebula remnant of the supernova observed in 1054 AD is the prototype

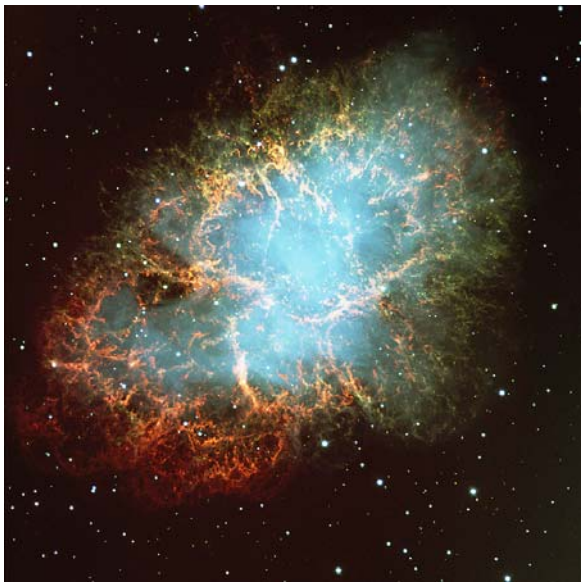

Figure I-4: Crab Nebula, young SNR (AD1054). [Credit: VLT Kueyen+FORS2]

Old SNRs (e.g., Cygnus Loop) are photoionized by X-rays emitted from dense cooling regions collisionally heated to 10^{5-6} K by the passage of the supernova blast wave through the ambient ISM. Most of the gas in the remnants has been plowed up by the shock, and we see the cooling zone where the gas has reached temperatures of ~10^4 K where line emission is most efficient .

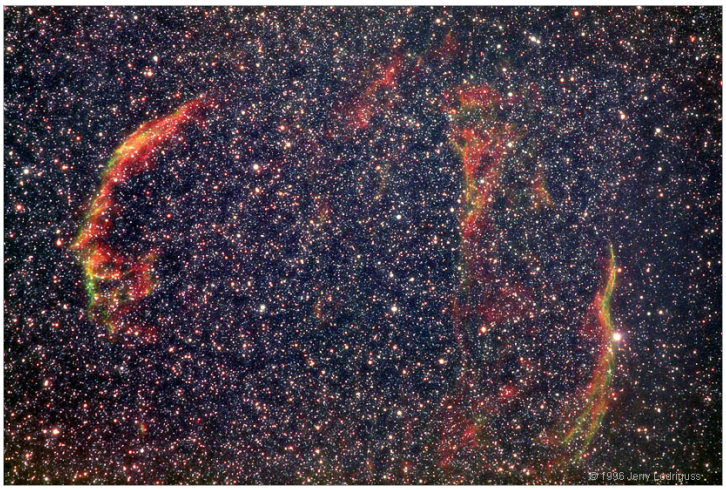

Figure I-5: The Cygnus Loop, an old SNR. This image shows emission from the shockwaves impinging on the ambient ISM (sharp filaments). [Credit/Copyright: Jerry Lodriguss, www.astropix.com]

The recognition that the spectra of photoionized regions contained a number of important diagnostics of the physical state of the gas (density, temperature, and elemental abundances to name the most important) was a major research field born out of Strömgren's seminal work on HII regions in the 1930s. The exploration of the physical processes involved is the subject of Chapter 3 of these notes.

Ionized hydrogen, however, is not confined to discrete regions, but at low surface-brightness is seen throughout the ISM. In fact, about 90% of the H^+ in the Galaxy lies *outside* classical HII regions. This gas makes up the "Warm Ionized Medium" ("warm" means kinetic temperatures of ~10^4 K). This map combined 3 all-sky imaging surveys (WHAM, VTSS, and SHASSA), and shows that low-surface brightness Hα emission essentially fills the sky.

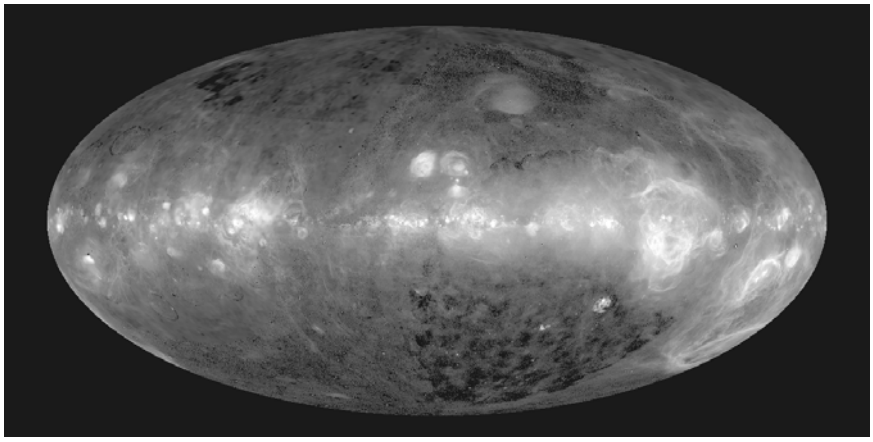

Figure I-6: Hα emission-line map of the Galaxy. [Credit: Douglas Finkbeiner (Harvard CfA)]

HI 21-cm Hyperfine Structure Emission (1950's)

That cold (10K) neutral hydrogen atoms can radiate line emission at 21-cm wavelengths via hyperfine atomic transitions in the ground state was predicted in 1944 by Hendrik van de Hulst, and first detected in1951 by Ewen and Purcell at Harvard (followed 6 weeks later by the Dutch astronomers Muller and Oort). Faint radio emission from cold HI was found to pervade the disk of the Milky Way.

This discovery initiated the era of radio-wavelength studies of the ISM and began the use of the ISM to trace Galactic structure. It has since become our principal tool for mapping the ISM in other galaxies and in the inter-galactic medium. Cold HI clouds make up much of the total mass of the ISM.

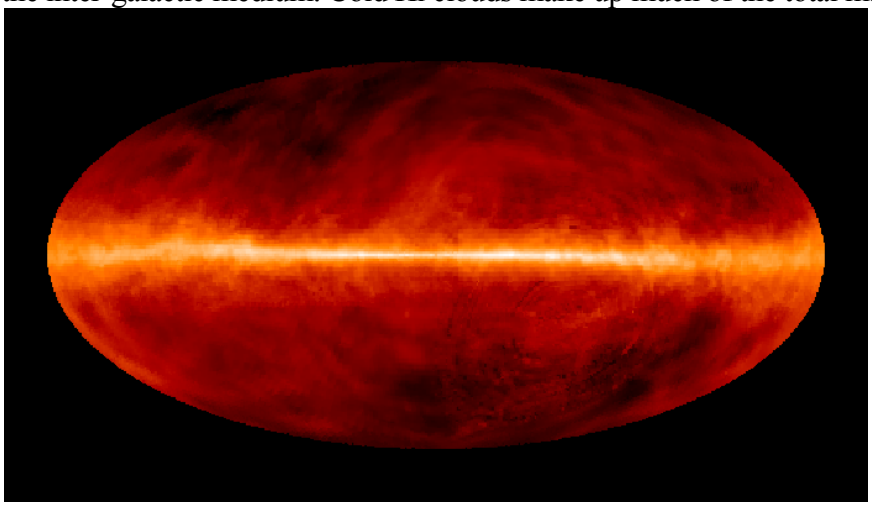

Figure I-7: HI 21cm emission-line map of the Galaxy. [Dickey & Lockman 1990]

Interstellar Molecules (1960's)

While Arthur Eddington predicted the existence of interstellar H_2 as early as 1926, and visible-wavelength absorption lines from diatomic molecules like CH, CN, and CH^+ were discovered in the late 1930s through the work of Swings, Rosenfeld, McKellar, and Adams, direct detection of emission lines at radio wavelengths was not achieved until 1963 with the discovery of the OH molecule at λ=18 cm by Townes et al., and emission from the polyatomic molecules of NH_3 (ammonia) at 1.2 cm and H_2CO (formaldehyde) detected in discrete interstellar clouds in 1968. The CO J=1-0 emission line at 2.6mm, one of the most important of the molecular tracers at mm wavelengths, was first detected in 1970 by Wilson, Jefferts, and Penzias (the are the same Wilson & Penzias of cosmic microwave background fame), with later detections of interstellar CN and HCN through the 1970s. As receiver and antenna technology improved through the 1980s and 90s, many 100s of increasingly complex interstellar molecules have been detected, primarily at mm wavelengths (the largest molecules are chains containing 13 atoms). This has ushered in the study of cold, dense molecular clouds in our own galaxy (and later in other galaxies, especially CO studies), and introduced *chemistry* as a feature of ISM physics.

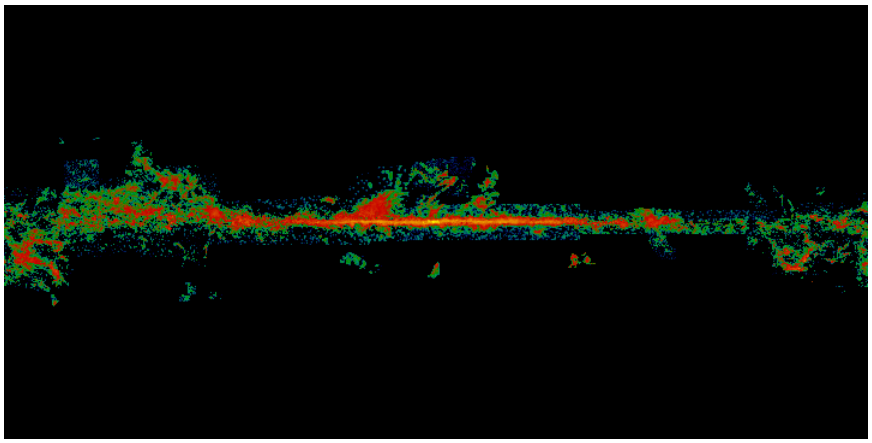

Figure I-8: CO J=1–0 (λ=1.6mm) emission-line map of the Galaxy [Dame, Hartmann, & Thaddeus, 2001, ApJ]

The most abundant interstellar molecule, H_2, is also the most elusive. It has no permanent dipole moment because it is a homonuclear diatomic molecule, and so it does not emit radio-wavelength lines, unlike CO, OH, and others. Interstellar H_2 was first detected in absorption towards hot stars in the satellite UV in 1973 by Carruthers et al. It was also in the 1970s that near-IR (1–2μm) emission-lines of collisionally or fluorescently excited H_2 were first detected in star formation regions and dense reflection nebulae. Most of the cold (3–10K) H_2, however, remains invisible to direct study.

Space Astrophysics (1970's to present)

A number of space-based observatories have been flown above the atmosphere, opening new wavelength windows, and with them new diagnostics of the ISM. Some highlights at different wavelength ranges:

Infrared: IRAS (1983) mapped most of the sky at 12, 25, 60, and 100μm wavelengths, revealing the pervasive Infrared Cirrus component of the ISM, and detecting thermal emission from warm dust in other galaxies. The **ISO** satellite mission (1996–1998) provided pointed imaging and spectrophotometry at near to far-IR wavelengths. While not "space", the **Kuiper Airborne**

Observatory (1974-1996) had a 36-inch telescope mounted in a converted C141 cargo plane. The KAO made extensive studies of the ISM and nebulae in the 1−100μm (and beyond) wavelength region. **COBE** (1989-1993), while primarily designed to map the Cosmic Microwave Background, also produced superb IR maps of the Galaxy (you have to get the Galaxy out of the way to see the CMB).

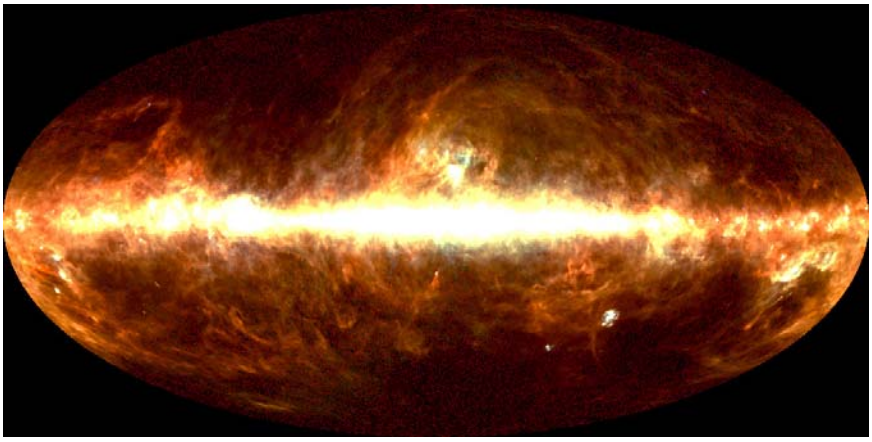

Figure I-9: COBE Far-Infrared (60-240μm) map of the Galaxy [Credit: NASA/GSFC]

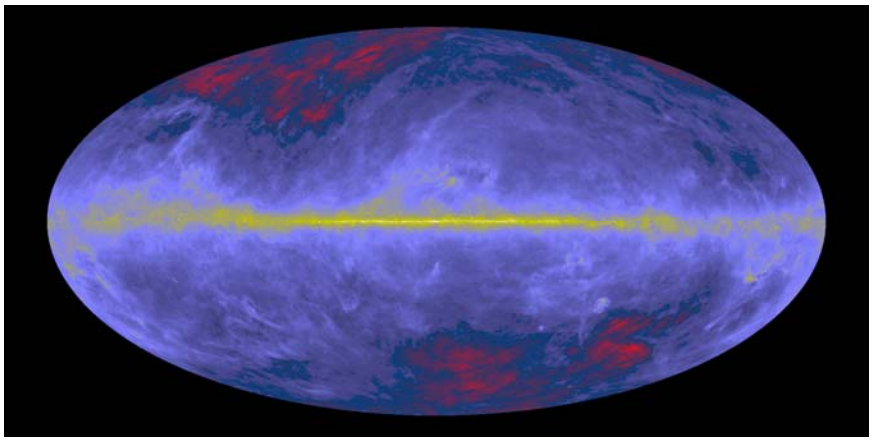

Figure I-10: Dust column density map of the Galaxy from IRAS & COBE data [Schlegel, Finkbeiner & Davis 1998]

The WMAP (Wilkinson Microwave Anisotropy Probe) satellite launched in 2001 as a high-resolution successor to COBE has already (2003) produced exquisite all-sky maps for CMB research, and will eventually produce as a by-product the highest resolution FIR and dust maps of the Galaxy (which needs to be subtracted to see the CMB in detail). The **Spitzer Space Telescope** (the observatory formerly known as SIRTF) was launched in 2003, and has made remarkable contributions in a wide variety of areas, but is a pointed observatory not an all-sky survey, and works primarily at near- and mid-Infrared wavelengths. Its primary mission ran through 2009 (limited by cryogen and its drift-away orbit taking it out of telemetry range of Earth). **Herschel** is an ESA mission operating a 3.5-meter Far-IR and Sub-mm telescope orbiting at the Earth-Sun L2 point for a 3-year primary mission (extendable by 1 year). It was launched in 2009 and has started to produce remarkable data. **SOFIA**, the successor to the KAO, is a 2.5-m telescope in a modified 747 that began flight testing in 2008 and should operate for about 20 years. It can observe at 0.3μm to 1.6mm.

X-Rays: Observations of a hot (10^{5-6} K) component of the ISM and individual SNRs have been made with satellites like **UHURU**, **Ariel**, **Einstein**, **ROSAT**, etc. X-ray spectroscopy of these regions as come from many, but especially **ASCA** and the **Chandra**, and **XMM/Newton** missions are bringing X-ray spectrophotometry of the ISM to the same level of sophistication enjoyed at UV, visible, and IR wavelengths. These recent observatories, however, are not making all-sky surveys, being dedicated instead to making "pointed" observations.

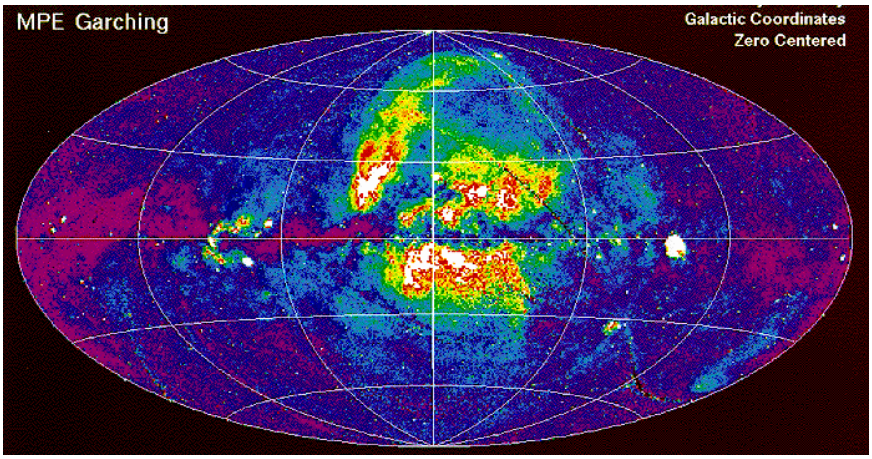

Figure I-11: ROSAT Soft X-ray (0.5-0.9keV) map of the Galaxy [Credit: RASS, MPE Garching]

γ-Rays: The only major γ-ray study of the Milky Way from space to date was carried out by the Compton/Gamma-Ray Observatory (**CGRO**) **EGRET** instrument. γ-rays are produced by high-energy nuclear processes, like e^+-e^- annihilation, particles accelerated by pulsar magnetic fields or supernova shocks, and relativistic outflows from exotic objects like X-ray binaries with black holes. **GLAST** (Gamma-Ray Large-Area Space Telescope) was launched in June 2008 and began science operations in the fall of 2008. With greater sensitivity and angular resolution (arcminutes), it has greatly extended the work done by EGRET.

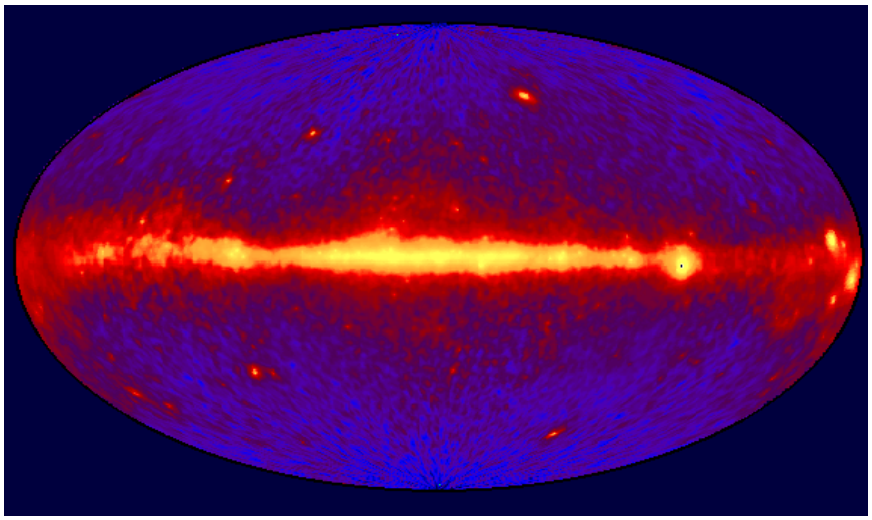

Figure I-12: CGRO EGRET γ-ray (>100MeV) map of the Galaxy [Credit: NASA/GSFC]

Ultraviolet: Copernicus (OSO-2) 1974 discovered UV absorption from diffuse H_2 regions (Lyman and Werner electronic bands), and mapped strong atomic HI Lyman-series absorption as well as other elements in the ISM. **IUE** (International Ultraviolet Explorer, 1980–1996) provided low-resolution

spectroscopy that contributed significantly to ISM line studies in our Galaxy and other galaxies. **EUVE** (Extreme Ultraviolet Explorer 1994–1999) probed the Universe below the Lyman continuum limit (λ<912Å), mapping the structure of the local ISM and Galactic EUV background, as well as contributing to extragalactic studies**. ORFEUS-SPAS** flew on the Shuttle in 1993 and 1996 and carried 3 instruments for Far- and Extreme-UV spectroscopy.

The **HST** (Hubble Space Telescope, 1990-present) has carried UV spectrometers, including:

> **GHRS** (Goddard High-Resolution Spectrometer) that has greatly contributed to UV absorption line studies of the ISM,
>
> **STIS** (Space Telescope Imaging Spectrometer) currently installed on HST, but it failed in 2005. It may be repaired in the 2008 servicing mission
>
> **COS** (Cosmic Origins Spectrograph) was installed during the 2009 final servicing mission. At UV wavelengths it is more sensitive at short wavelengths than STIS.

Lyman/FUSE (Far-Ultraviolet Spectroscopic Explorer) was launched in June 1999 and operated until early 2008 when its last gyro failed. FUSE was the first long-duration UV satellite to be devoted to the Far-UV region (912–1100Å) since the 1972 *Copernicus* mission. This spectral region includes the principal HI and H_2 absorption lines in the ISM, and resonance lines of many other important metal-lines species. Because FUSE was about 10^5 times more sensitive than Copernicus, it revolutionized ISM studies in the Far-UV, and was the first Far-UV mission to make significant contributions to extragalactic studies.

The ISM in the Computer

On the theoretical side, advances in computers have contributed to the study of the ISM in two ways:

Atomic & Molecular Physics: Advances in computational power has made possible *ab initio* calculations of fundamental atomic and molecular parameters (ionization cross sections, collision strengths, transition probabilities, etc,) that are essential to interpreting the spectra of interstellar plasmas (or any other kind of astrophysical or laboratory plasmas, for that matter). Examples include the atomic opacities projects at Los Alamos (LAOP) and the international Opacities and Iron Projects (including the work of Anil Pradhan and collaborators at OSU).

Nebular Modeling Codes: A number of groups have developed detailed photoionization equilibrium codes for modeling the spectra of nebulae. These include packages like CLOUDY (Gary Ferland), XSTAR (Tim Kallman), MAPPINGs (Dopita, Binette, et al.), and others. These programs all depend on inputs from the atomic and molecular parameter calculations described above. Combined with modern fast computers, they allow exploration of the observational parameter space to understand how to interpret spectroscopic data, and wring from them interesting quantitative results.

I-2 Physical Description of the ISM

The ISM is described physically in terms of thermodynamic properties: density, temperature, pressure, etc. In particular, the idea of *thermal phases* is an important element of this description. It is important to emphasize from the outset, however, that the ISM is a system that is very far from thermal equilibrium. This is not to say that other types of equilibria do not prevail. Four principal forms of equilibria are encountered: *Kinetic Equilibrium, Excitation Equilibrium, Ionization Equilibrium*, and *Pressure Equilibrium*.

Kinetic Equilibrium

Because most collisions in the ISM are elastic, the particle velocities quickly setup a Maxwellian distribution:

$$f(\mathbf{v})d^3v = \left(\frac{\mu}{2\pi kT}\right)^{3/2} \exp\left(-\frac{\mu v^2}{2kT}\right) dv_x dv_y dv_z$$

$$v^2 = (v_x^2 + v_y^2 + v_z^2)$$

$\mu \equiv$ reduced mass of the particles

A system with a Maxwellian distribution of velocities is said to be in **Kinetic Equilibrium**, and the temperature, T, characterizing that velocity distribution is the **Kinetic Temperature** of the system.

The timescale for elastic electron-electron collisions is very short:

$$t_{ee} \approx 10^4 \left(\frac{E_e}{1\text{ eV}}\right)^{3/2} n_e^{-1} \text{ sec}$$

By contrast, the timescale for electron-Hydrogen collisions is

$$t_{eH} \approx 2\times 10^7 \left(\frac{E_e}{1\text{ eV}}\right) n_e^{-1} \text{ sec}$$

Which of these dominates depends on the ionized gas fraction. In ionized regions, the electron-electron collision timescale dominates, and is much shorter than all other timescales in the system (e.g., recombination or photoionization), whereas in cold neutral and molecular regions the electron-Hydrogen timescale is relevant. Because both timescales are so much shorter than those of all other processes, the systems very quickly **thermalize** and setup a Maxwellian velocity distribution characterized by a kinetic temperature, T. In ionized gas regions, this temperature is more specifically the kinetic temperature of the *electrons*, whereas in cold neutral and molecular clouds it will be the kinetic temperature of neutral or molecular hydrogen.

Excitation Equilibrium

For a gas in Local Thermodynamic Equilibrium (LTE) we denote the relative level populations within an atom or molecule n_i^* in terms of the Boltzmann Equation:

$$\frac{n_u^*}{n_l^*} = \frac{g_u}{g_l} e^{-\Delta E_{ul}/kT}$$

Where $g_i = (2J_i + 1)$ is the statistical weight of the i[th] energy level, and T is the *Kinetic Temperature* of the system. Another way of parameterizing level populations is relative to the population in the ground state, n_0:

$$\frac{n_u^*}{n_0^*} = \frac{g_u}{g_0} e^{-\Delta E_u / kT}$$

Alternatively we might wish to know the fractional population. We first need to know the sum over all states n^*, which is generally expressed in terms of a *Partition Function* $f(T)$ for the system:

$$n^* \equiv \sum_{i=0}^{\infty} n_i^* = \frac{n_0^*}{g_0} \sum_{i=0}^{\infty} g_i e^{-E_i/kT} = \frac{n_0^*}{g_0} f(T)$$

The fractional level population in some level j is then:

$$\frac{n_j^*}{n^*} = \frac{g_j}{f(T)} e^{-E_j/kT}$$

All of these level populations are for LTE. Because the ISM is a system far from thermal equilibrium, we need to compute the true level populations n_j. There are two formalisms for expressing true level populations n_j in terms of LTE populations n_j^*: departure coefficients and excitation temperature:

Departure Coefficients: $b_j \equiv n_j / n_j^*$

For example:

$$\frac{n_u}{n_l} = \frac{b_u}{b_l} \frac{n_u^*}{n_l^*}$$

These measure the "departure" of the true level populations from the LTE populations predicted by the Boltzmann equation.

Excitation Temperature: T_{exc}

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} e^{-\Delta E_{ul}/kT_{exc}}$$

This *looks* like the Boltzmann equation, but we have substituted the *Excitation Temperature* (T_{exc}) for the usual Kinetic Temperature ($T_{kinetic}$). In general, $T_{exc} \neq T_{kinetic}$, except at very high densities when the excitation temperature begins to approach the kinetic temperature of the gas.

The first question that is usually asked is "what is T_{exc} the temperature of?" The short answer is that it is *not the physical temperature of anything*; it is a convenient way to parameterize the relative non-LTE level populations in terms of a single number (T_{exc}) instead of a pair of decidedly inscrutable numbers (departure coefficients).

These two different formalisms are completely interchangeable; although in general we will more often work in terms of excitation temperatures instead of departure coefficients (the latter is more common for detailed radiative transfer calculations). The excitation temperature formalism is the most convenient for many of the problems we will encounter here, but you must be careful not to confuse the excitation temperature of the system (a parameterization of the relative non-LTE level populations) with the physical (kinetic) temperature of the gas.

Ionization Equilibrium

In the ISM, the radiation field is primarily responsible for ionization, unlike the case of stellar atmospheres or very hot plasmas ($T_e > 10^6$ K) where collisional ionization is dominant. These photons can come from either the general **Interstellar Radiation Field** (**ISRF**) or from a nearby hot star. Ionization is balanced by recombination of a free electron with an ion:

$$X^r + h\nu \leftrightarrow X^{(r+1)} + e^-$$

Ionization Equilibrium occurs when the photoionization rate exactly balances the recombination rates:

$$N(X^r)N(photons)\sigma_{ionize}c = N(X^{(r+1)})N_e\sigma_{recomb}\left\|\mathbf{v}_X - \mathbf{v}_\mathbf{e}\right\|$$

The quantity in $\| \ \|'s$ is the relative velocity between the ion and electron.

We can rewrite this in a form resembling the Saha Equation:

$$\frac{n(X^{(r+1)})n_e}{n(X^r)} = \frac{n_{photons}\left\langle\sigma_{ionize}\right\rangle c}{\left\langle\sigma_{recomb}v\right\rangle}$$

Here the averages inside the ⟨⟩'s are computed over photon energies in the numerator, and over the Maxwellian distribution of electron velocities in the denominator.

Since the recombination cross-section, σ_{recomb}, is proportional to v^{-2} (it is easier to recombine at slower relative velocities), we can recast the equation above in the form:

$$\frac{n(X^{(r+1)})}{n(X^r)} = \frac{1}{n_e}\left(\frac{\Gamma}{\alpha(T)}\right)$$

where

$$\alpha(T) = \text{"Recombination Rate"}$$
$$\Gamma = \text{"Radiation Field Term"}$$

If the electron density is low, then recombinations are rare and the species stays in the ionized state most of the time.

Where do the electrons come from?

> **HII Regions:** the dominant source of electrons is Hydrogen ionized by photons with energies above 13.6eV, with a proportionally smaller contribution from Helium photoionization. Because H and He are the most abundant elements (all others are 10^{-4} or smaller relative to H), ionization of metals (O, N, etc.) contributes very little to the electron population.

HI Regions: The stellar radiation field is sharply truncated at energies above 13.6eV by the neutral hydrogen ionization threshold, so only those species with ionization potentials <13.6eV are photoionized and contribute to the electron population. The primary source of electrons in neutral gas regions in the diffuse ISM is the photoejection of electrons from the surfaces of dust grains, with neutral Carbon, C^0, followed by Si, Fe, and S, as the most important gas-phase electron donors in HI regions. Cosmic rays and diffuse X-rays from the hot (10^{5-6}K) components of the ISM can also ionize hydrogen, particularly deep inside dense clouds providing an additional source of free electrons in regions where UV photons cannot readily penetrate.

Despite the fact that the dust-to-gas ratio in the ISM is approximately 1:100, dust grains contribute about 2 orders of magnitude more electrons than the gas. This can be demonstrated by a simple order-of-magnitude calculation. In the absence of dust, a lower limit on the fractional ionization, *x*, in HI regions is given by the relative abundance of C and H:

$$x \equiv \frac{n_e}{n} \geq \frac{n_C}{n_H} \approx 4 \times 10^{-4}$$

Observationally, a direct estimate of the interstellar free electron density comes from the timing of pulsar signals at radio wavelengths. The free electrons in the ISM act like a transmitting medium with an index of refraction that decreases with increasing frequency, so that that high frequencies propagate faster than low frequencies, and causing pulses to arrive from a pulsar sooner at high frequencies. The dispersion ("spread") in pulse arrival times between two widely separated frequencies is just the mean electron density integrated along the line of sight to the pulsar.

The mean interstellar electron density derived from pulsar dispersion measures is $\langle n_e \rangle \approx 0.03$ cm^{-3}, which for typical interstellar Hydrogen densities of ~1 cm^{-3} implies an ionization fraction of $x \approx 0.01–0.04$. This is about 2 orders of magnitude larger than the limit derived from the C/H abundance. The source of electrons is photoejection from dust grains, illustrating a general idea that we will see repeated throughout this class: dust grains play an important role in nearly every phase of the interstellar medium.

Pressure Equilibrium

The simple cloud picture of the ISM, motivated by early absorption-line studies, envisions cold, dense interstellar clouds embedded in a warm, low density, ionized intercloud medium. Such co-existence may be understood in terms of pressure balance between the two components.

If only thermal pressure is important, then

$$n_H T_{cool} \approx (n_e + n_p) T_{warm}$$

In the 1950s Lyman Spitzer speculated that this kind of thermal pressure balance might permit the existence of a third, very hot ($T \approx 10^6$K) but very low density galactic "corona" (by analogy with the solar corona) in approximate pressure equilibrium with the cold and warm gas.

Note, however, that we have only assumed that *thermal* pressure is important. Other possible pressure sources that can contribute significantly include:

Magnetic Pressure $B^2 / 8\pi$ from interstellar magnetic fields

Cosmic Ray Pressure from protons with energies of a few MeV

Hydrodynamic Pressure (e.g., ram pressure) from moving gas streams, stellar winds, or supernova blast waves.

Radiation Pressure from starlight and the cosmic background radiation.

In the ISM, cold clouds typically have densities and temperatures of $n_H \geq 10$ cm^{-3} and T≤300 K, while the warm ionized intercloud medium has $n_H \leq 0.3$ cm^{-3} and T≥10^4 K. Each has a thermal pressure of

$$\left(P/k\right) = nT \approx 3000 \text{ cm}^{-3} \text{ K}$$

By comparison, typical interstellar magnetic fields have strengths in the range 2–8μGauss, which imply typical magnetic field pressures, in similar units, of

$$\left(P/k\right)_B \approx \frac{B_0^2}{8\pi k} \approx 2600 \text{ cm}^{-3} \text{ K} \quad (\text{for } B_0 = 3\mu G)$$

If relatively ordered magnetic fields are present on the scale of cold interstellar clouds (a few parsecs), magnetic pressures can be at least as important (if not more so) than thermal pressure in establishing the overall pressure balance between clouds and their confining intercloud medium. Disordered (random) magnetic fields would produce about 1/3 the pressure of ordered fields (as a rule of thumb, ordered fields have a pressure of $B^2/8\pi$, whereas random fields are produce a pressure of ~$B^2/24\pi$).

It is also useful to express the strengths of the various sources of pressure in the ISM in terms of the **energy density**, u, of each contribution expressed in units of eV cm^{-3}. Using energy density lets us think in terms of the energy available to be deposited into a region of the ISM, rather than in terms of pressure (though they are physically equivalent ways of framing the problem). There are six principal sources of energy in the ISM.

Thermal Energy, u_{th}

$$u_{th} = \frac{3}{2}P = 0.39\frac{P/k}{3000 \text{ cm}^{-3} \text{ K}} \text{ eV cm}^{-3}$$

Hydrodynamic Energy, u_{hydro}

$$u_{hydro} = \frac{1}{2}\rho\left\langle v^2 \right\rangle = 0.13\left(\frac{n_H}{cm^{-3}}\right)\left(\frac{\sigma_{1D}}{5 \text{ km s}^{-1}}\right)^2 \text{ eV cm}^{-3}$$

Magnetic Energy, u_{mag}

$$u_{mag} = \frac{B^2}{8\pi} = 0.22\left(\frac{B}{3\ \mu\text{G}}\right)^2 \text{ eV cm}^{-3}$$

Cosmic Rays, u_{CR}

$$u_{CR} = 0.8 \text{ eV cm}^{-3}$$

Starlight, u_{stars}

$$u_{stars} = 0.5 \text{ eV cm}^{-3}$$

Cosmic Background Radiation, u_{CBR}

$$u_{CBR} = 0.26 \text{ eV cm}^{-3} \text{ for T=2.725K}$$

Note that all of these sources of energy density are comparable to within an order of magnitude.The coincidence of these six very different sources of pressure/energy density emphasizes the fact that the ISM is a dynamic and complex system in which matter, momentum, and energy are constantly being exchanged between stars and interstellar gas in its various phases.

I-3 Global Models of the ISM

The gas in the ISM exists in a number of **Thermal Phases** depending upon the local conditions of heating, ionization, etc. The classic work describing the "two-phase" ISM mode is by Field, Goldsmith, & Habing (1969, ApJ, 155, L149; usually known as "FGH"), who identified two stable phases, warm and cold, by considering the stability of clouds in local thermal pressure equilibrium against isobaric perturbations. Later work, notably by McKee & Ostriker (1977, ApJ, 218, 148) expanded this to consider the role of supernova remnants in contributing a hot, low-density "third phase" to the cold and warm phases identified by FGH.

Both models are interesting insofar as they illuminate the basic physics that must go into a successful global model of the ISM, but each fails in detail in ways that are also enlightening. Recent observations, especially EUVE observations of the local ISM, have thrown theorists back from their originally sound position. We have many observational and theoretical pieces that look right, but fitting them together into a coherent and self-consistent global model of the ISM has so far proven elusive.

Sources of Heating

A number of heating mechanisms are at work within different regions of the ISM. Which mechanism dominates depends on the conditions characteristic of that phase of the ISM.

Photoelectrons ejected from Dust Grains

UV and higher-energy photons can knock electrons out of small dust grains or large molecules, contributing thermal electrons to the mix. Originally thought to be negligible compared to cosmic rays, later studies have suggested that it is, in fact, the dominant source of heating in the diffuse ISM. The problem is that photoelectrons from large grains are ineffective at heating the ISM, so the most likely donors are small grains and large molecules like Polycyclic Aromatic Hydrocarbons (PAHs). The photons responsible for this heating come from the FUV part of the spectrum, below the hydrogen ionization threshold (13.6ev).

The total photoelectric heating rate is

$$nG_{pe} \approx 10^{-24}\,\varepsilon n G_0 \text{ erg cm}^{-3}\text{ s}^{-1}$$

Here ε is the heating efficiency, n is the gas density, and G_0 is the mean intensity of the ISRF. Theoretical estimates of the heating efficiency range from 0.003 in a dense photodissociation region to as high as 0.05 in regions where neutral PAHs and small dust grains are abundant. The maximum heating rate per nucleon (G/n_H) is ~5×10^{-26} erg/atom/sec. This heating rate is far in excess of that due to any other mechanism (cosmic rays or photoionization of neutral metals), and dominates in most of the diffuse ISM.

This mechanism becomes inefficient in regions where FUV photons from the ISRF, such as the cores of molecular clouds. In warmer regions or high ambient FUV radiation fields, grains can become charged and PAHs ionized, which means the escaping photoelectrons must also break free of Coulomb interactions as well as the material work functions, suppressing the heating rate.

Cosmic Rays

Low-energy cosmic rays (1-10 MeV protons) can heat the ISM via ionizing collisions:

$$H^+(1-10\text{ MeV}) + H \rightarrow H^+ + H^+ + e^-$$

where the ejected electron has a mean energy of ~35eV. This electron will very quickly thermalize via e-e collisions, although it is sufficiently energetic that it can cause secondary ionizations of hydrogen and helium can increase the effective heating rate.

The heating rate due to cosmic rays is

$$nG_{CR} = n\zeta_{CR}\left[1+\phi^{H}(\varepsilon, x_e)+\phi^{He}(\varepsilon, x_e)\right]\langle\varepsilon\rangle$$

ζ_{CR} is the cosmic-ray **primary ionization rate,** estimated to be ~2×10^{-16} s^{-1}, the []s include the secondary ionizations of H and He (the ϕ terms), and $\langle\varepsilon\rangle$ is the mean energy of the ejected electron.

Secondary ionization can contribute as much as 80% in regions where the electron fraction is small, such as in diffuse ISM clouds, but when the electron fraction is $x_e>10^{-3}$ there are so many electrons around that e-e collisions dominate and secondary ionization is suppressed.

In the cold neutral medium, the heating rate for low-degrees of ionization (mean energies of ~7eV) is

$$nG_{CR} \approx 3\times10^{-27} n\left[\frac{\zeta_{CR}}{2\times10^{-16}}\right] \text{ erg cm}^{-3}\ \text{s}^{-1}$$

Estimates of the cosmic ray heating rate as high as 10^{-15} s^{-1} will often be found in older papers (and ISM books), but all of these have been revised downward nearly 2 orders of magnitude with better observations. In general, ζ_{CR} is poorly constrained by observations because low-energy cosmic rays are easily deflected by magnetic fields in the ISM and solar system. The rate is instead inferred from observations of the ionization fraction in molecules deep inside cold neutral and molecular clouds where cosmic ray heating is expected to dominate.

Photoionization Heating

The primary source of ionizing photons in the diffuse ISM is Far UV photons from the Interstellar Radiation Field (ISRF), mostly 11-13.6eV photons leaking away from O and B stars.

$$X^{r} + h\nu_{ISRF} \rightarrow X^{(r+1)} + e^{-}$$

The mean energy imparted to the electron is the difference between the energy of the ISRF photon and the ionization potential of the species, and is usually a few eV. In an HI region only neutral metals are ionized (by definition there are few or no hydrogen ionizing photons in an HI region), with $C^0 \rightarrow C^+$ ionization dominating (11.3eV). The mean photoelectron energy is ~1eV, well below the ionization potential of hydrogen and other species, so secondary ionization is contributes negligibly to the heating rate. The C^+ heating rate is approximately

$$n_H G_{C^+} = 2.2\times10^{-22}\left(\frac{n_{C^0}}{n_C}\right)\left(\frac{n_C}{n_H}\right) nG_0 e^{-2.6A_V} \text{ erg cm}^{-3}\ \text{s}^{-1}$$

The first term is the fraction of neutral carbon, the second term is the total C/H abundance, G_0 is the mean ISRF intensity attenuated by dust parameterized by the total visual extinction A_V. For typical numbers in the diffuse ISM, the photoionization heating rate per nucleon due to $C^0 \rightarrow C^+$ is about $10^{-29}G_0$ erg/atom/sec, many orders of magnitude smaller than either photoelectric or cosmic ray heating.

Summary of Heating Processes

In the cold neutral clouds of the ISM where T=100K, $G_0 \approx 1$, and the ionization fraction is about 1.4×10^{-4}, photoelectric heating dominates over the entire range of densities, followed by cosmic ray heating down by about a factor of 10-100. Photoionization heating by Carbon and X-ray photoelectric heating are down by many orders of magnitude over the entire range.

In the warm neutral intercloud medium (T=8000K, and ionization fraction $\sim3\times10^{-3}$), photoelectric heating from grains dominates at densities larger than 0.1 cm^{-3}, but at lower densities for the heating becomes dominated by a combination of cosmic ray and X-ray heating, with negligible contributions from carbon photoionization (aided in part by the fact that at 8000K the carbon ionization fraction is high making photoionization heating inefficient).

In the cold cores of molecular clouds (T=10K and ionization fractions $<10^{-7}$), FUV photons from the ISRF cannot penetrate and photoelectric heating becomes negligible. Cosmic ray heating dominates over the entire range of densities, and other processes related to gas dynamics, particularly hydrodynamic turbulence, gravitational contraction, and collisions between warm dust grains and the atomic and molecular gas become important, but all at least an order of magnitude smaller than cosmic ray heating.

Sources of Cooling

Most cooling is by the particles in the gas emitting line and continuum radiation, some or all of which escapes the region and carries off energy. In a sense, the emergent spectrum of any phase of the ISM is the consequence of the gas in that phase trying to cool.

Collisionally Excited Line Emission

In cold regions, cooling is dominated by collisional excitation of C^+ by collisions with thermal electrons followed by emission of infrared fine-structure lines. As the temperature rises, other species begin to contribute collisionally excited lines to the cooling (see Figure I-13).

The ground state of C^+ has a fine-structure transition $^2P_{1/2}$–$^2P_{3/2}$ with excitation energy (in temperature units) of $\Delta E/k=92$ K, which emits a far-infrared photon with a wavelength of 158μm. The excited state is populated by collisions with electrons, or if the ionization fraction is low, it can be excited by collisions with either HI or H_2 molecules.

For example, the cooling rate for the [C II]158μm line excitation is (following Tielens' book):

$$n^2\Lambda_{[\mathrm{CII}]} = 3\times10^{-27} n^2 \left(\frac{A_C}{1.4\times10^{-4}}\right)\left(1+0.42\left(\frac{x}{10^{-3}}\right)\right)\left(\frac{T}{100K}\right)^{-1/2} e^{-92/T} \ \mathrm{erg\ cm^{-3}\ s^{-1}}$$

Where A_C is the carbon abundance relative to Hydrogen, x is the ionized fraction, and T is the kinetic temperature of the gas. Because the excitation is a 2-body collisional process, it has an n^2 density dependence and both an inverse square-root and exponential temperature dependence.

In HI regions with typical kinetic temperatures of 100 K, we expect the primary form of cooling to be by the excitation of far-infrared fine structure lines (transition whose energy is of order the mean thermal energy of electrons in kinetic equilibrium at 100K). Because H and He have first excited states with excitation energies >10–20eV, respectively, abundant metal ions (e.g., C^+ and Si^+) and neutral metals (e.g., O^0) will carry most of the radiative cooling load at these temperatures. The lines actually observed (from ISO and Kuiper) from HI regions are:

Species	Transition(s)	Wavelength	ΔE/k
C^+	$^2P_{1/2}$–$^2P_{3/2}$	158μm	92 K
Si^+	$^2P_{1/2}$–$^2P_{3/2}$	34.8μm	414 K
N^+	$^3P_0$–$^3P_1$	204μm	70K
	$^3P_2$–$^3P_1$	122μm	120K
O^0	$^3P_0$–$^3P_1$	147μm	98 K
	$^3P_2$–$^3P_1$	63.2μm	228 K

When the temperature climbs above ~8000 K, collisional excitation of ionized metals starts to become important, as is excitation of H into its first excited level (10.5eV).

Collisional Ionization & Excitation of H

At T≥ 15000 K collisional ionization is followed by recombination and emission of line radiation as the electron cascades into the ground state. H^0 can also be collisionally excited into bound state, followed by radiative de-excitation and emission of line radiation. In low-density regions, the gas is optically thin to this line radiation, so it acts as a source of cooling.

Other Cooling Sources

A number of other cooling mechanisms can contribute to the overall cooling, a few of interest are:

- Recombination cooling of H (minor losses).
- Thermal Bremsstrahlung radiation for $T>10^6$ K (cooling $\propto T^{1/2}$).
- Molecular cooling at low temperature in molecular regions (very complicated)
- Thermal and non-thermal emission from dust grains after collisions with atoms, molecules, or electrons.

Since all of these involve collisions, their cooling rates are all proportional to n^2.

Interstellar Cooling Curve

A representative interstellar cooling curve (adapted from Dalgarno & McCray 1972, ARAA, 10, 375) is shown in Figure I-13. The shaded regions are unstable according to the instability criterion above.

At low temperatures, cooling is dominated by collisional excitation of far-IR fine structure lines, principally [CII]158μm. These line cooling rates are very sensitive to the electron fraction of the gas, x_e, as if there are few free electrons in the gas to collide with C^+ and excite metal-line cooling, the cooling rate will diminish as the square of the density.

Around 10^4K, the rapid rise in cooling represents the sudden onset of Lyα cooling with contributions by metal ion collisional cooling (particularly strong highly-ionized cooling species responsible for the various bumps at the top of the cooling curve are illustrated). By 10^5K, cooling abruptly drops as hydrogen is now mostly ionized along with most of the metals. As the metals become more ionized, they have fewer electrons left to be collisionally excited, and these more tightly bound. Finally, above 10^7K, cooling begins to pick up again, more slowly, as free-free cooling becomes dominant.

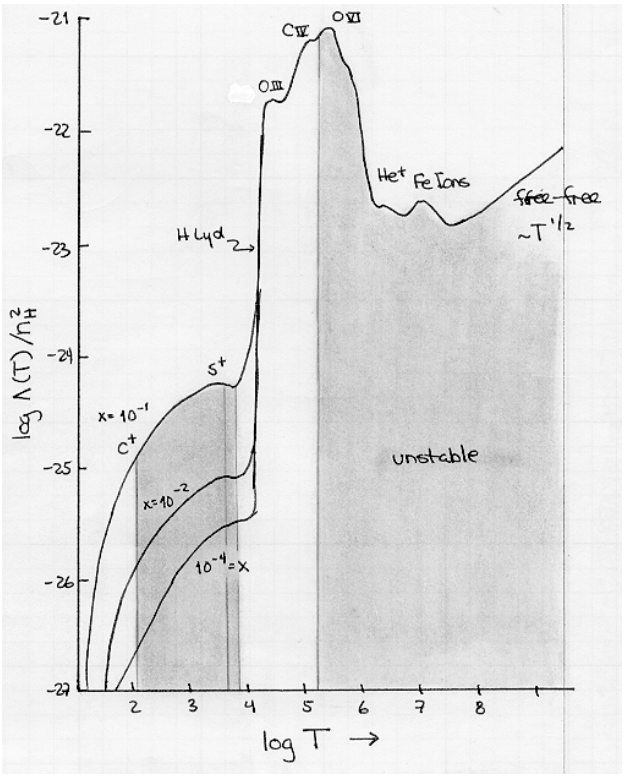

Figure I-13: Interstellar Cooling curve adapted from Dalgarno & McCray (1972).

Thermal Stability and Equilibrium – The Two-Phase ISM Model of Field, Goldsmith & Habing (1969)

Consider gas with density, *n*, a cooling rate per unit volume of $n^2\Lambda(T)$, and a heating rate per unit volume of *nG*. If the gases in different thermal phases are in thermal contact, and in *thermal* pressure equilibrium (i.e., ignoring non-thermal pressure sources like magnetic fields), then the thermal balance between heating and cooling is expressed in terms of a **Generalized Loss Function**, L:

$$L(n,T) = n^2\Lambda(T) - nG$$

where:

$$L > 0 \Rightarrow \text{ Net Cooling}$$
$$L < 0 \Rightarrow \text{Net Heating}$$
$$L = 0 \Rightarrow \text{ Equilibrium}$$

For a gas at constant (thermal) pressure nT in equilibrium with density ρ_0 and temperature T_0, in terms of the Generalized Loss Function, equilibrium occurs when

$$L(\rho_0, T_0) = 0$$

Is the gas stable? If either the density or the temperature of this gas is perturbed while holding some thermodynamic variable (like the pressure) fixed, the equilibrium will be **unstable** if:

$$\left(\frac{\partial L}{\partial S}\right)_X < 0$$

where X is the fixed variable (e.g., Pressure), and S is the entropy of the gas. For isobaric (constant pressure) perturbations in a perfect gas, this condition for thermal instability becomes:

$$\left(\frac{\partial L}{\partial T}\right)_P = \left(\frac{\partial L}{\partial T}\right)_\rho - \frac{\rho_0}{T_0}\left(\frac{\partial L}{\partial \rho}\right)_T < 0$$

For example, if heating is dominated by cosmic rays and cooling by collisions, the loss function is:

$$L = n^2\Lambda(T) - n_H\zeta_{CR} = 0$$

Since the cosmic-ray ionization rate ζ_{CR} is constant, the criterion for isobaric instability above is

$$\left(\frac{d\ln\Lambda}{d\ln T}\right)_P < 1$$

(The proof of this is left as a homework exercise). If we were to substitute photoelectric heating from dust grains, we'd get a similar criterion, as this cooling depends on the ISRF intensity which to a first approximation is about constant in the diffuse ISM.

At constant pressure (nT=constant) and constant heating, we can plot the equilibrium condition, $L(n,T)=0$ in the log(Λ(T)/T) versus log(T) plane as shown below.

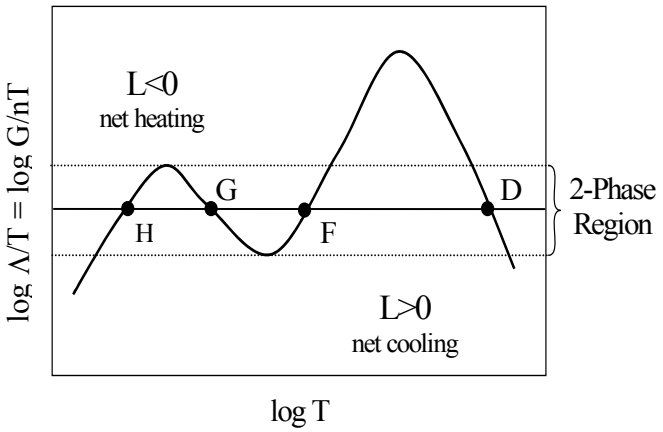

Figure I-14: Schematic Generalized Loss Function. The curve indicates the equilibrium condition, L=0. The horizontal line in the middle of the 2-phase region shows a line of constant pressure (nT).

A line of constant G/nT intersects the L=0 curve at 4 points, labeled H, G, F, and D as shown in Figure I-14 above.

At points H and F:

$$\left(\frac{\partial L}{\partial T}\right)_P > 0$$

which implies *stable* phases at T=100 K (cold) and 8000 K (warm).

At points G and D:

$$\left(\frac{\partial L}{\partial T}\right)_P < 0$$

which implies *unstable* phases at these points. Let's examine what is happening at point G and its neighbors (H and F) more closely to see why it is unstable. This is shown in Figure I-15.

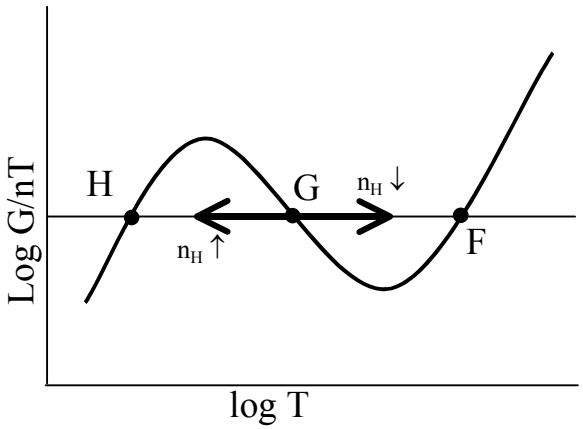

Figure I-15: Region around point G, showing perturbations in density from the equilibrium point.

To maintain constant pressure, nT, this requires that n vary like T^{-1}. For perturbations from equilibrium about point **G**,

1. **If n *increases* ⇒ T must *decrease*** (cloud collapses).

   The isobaric perturbation to higher density and lower temperature drives the cloud at point G into the region of net cooling (L>0). This causes T to decrease further, which at constant pressure means the density must increase, sending the cloud into a runaway collapse caused by runaway cooling.

2. **If n *decreases* ⇒ T must *increase*** (cloud expands).

   The isobaric perturbation to lower density and higher temperature drives the cloud at point G into the region of net heating (L<0). This causes T to increase further, which at constant pressure means the density must decrease further, sending the cloud into a runaway expansion caused by runaway heating.

By contrast, at the stable points H and F, isobaric perturbations in density that lead to net heating (cooling) produce a compensatory response in the temperature. For example, compression leading to cooling drives the cloud into the region of net heating, which drives it back to the original temperature.

The parameters of each of these phases of the original FGH model are as follows:

Phase	n_H (cm^{-3})	T (K)	$x=n_e/n_H$	Identification
F	≤0.1	≥8000	0.1	Warm Intercloud Medium
G	0.2	8000	—	**UNSTABLE**
H	≥10	≤300	0.001	Cold Neutral Clouds

In general, the agreement was pretty good considering the state of the observational data in 1969. Recent observational work, however, suggests that non-thermal pressures from magnetic fields, turbulence, etc. may in fact dominate over thermal pressure, and the assumption of pressure equilibrium is not relevant for the real ISM. While illustrative of the basic physics at play, the FGH model is not the whole story.

The Hot ("Third") Phase

The history of the third phase is more varied. In the 1950s, Spitzer predicted a hot "coronal" gas that was in pressure equilibrium with the cold and warm ionized gas seen locally. McCray & Buff (1972, ApJ, 175, L65) suggested a hot phase heated by cosmic rays. Cox & Smith (1974, ApJ, 189, L105) suggested instead that Supernova Remnants could produce this very hot phase. Maps of the local ISM show it to have numerous "loops" and "bubbles" (see schematic next page). The Sun resides in one such "Local Bubble", a hot (~10^6K) low-density (0.005 cm^{-3}) region characteristic of this phase.

Based on a thermal stability argument like that presented above, it is clear that this hot phase cannot be stable. However, it is such a low-density medium that the cooling time is very long, ~10^9 years, much longer than the mean time between supernovae (about 1 per 50 years), thus new explosions would continually reheat the ISM before it could cool by very much.

Cox & Smith asked how many hot, low-density supernova bubble are needed to fill the ISM? They defined "fill" in terms of a porosity factor, q, which if $q>1$ the bubbles all overlap and merge into a pervasive, hot, low-density component of the ISM. Cox & Smith estimated that

$$q > 0.1 S_{-13}$$

where S_{-13} is the supernova rate in units of 10^{-13} SN pc^{-3} yr^{-1} (essentially a SN rate/volume). In 1974, this was thought to be of order unity in these units, so the porosity was ~0.1

McKee & Ostriker (1977, ApJ, 218, 148) re-examined the Cox & Smith model and suggested an improved estimate of the porosity:

$$q = 0.5 S_{-13} E_{51}^{1.28} n_0^{-0.14} \tilde{p}_{04}^{-1.3}$$

where:

S_{-13} = SN rate/volume in 10^{-13} SN pc^{-3} yr^{-1}
E_{51} = SN blast energy in 10^{51} ergs
n_0 = ambient ISM density in cm^{-3}
$\tilde{p}_{04} = (p_0 / k) \times 10^{-4}$ cm^{-3} K, and
$p_0 = (n_e + n)_0 k T_0$ pressure of the ISM in 10^4 cm^{-3} K

For their assumptions, S_{-13}=1, E_{51}=1, $n_0 \leq 0.3$ cm^{-3}, $T_0=10^4$ K, and there is purely thermal pressure, hence $\tilde{p}_{04} \approx n_0$. Under these assumptions, they derived a porosity of q>3. This implies that the supernova bubbles overlap and that interactions between bubbles are important in ISM physics.

The subsequent McKee & Ostriker (MO for short) picture is of a 3-phase ISM, as follows:

The **third phase** is a hot (10^6 K), low-density (~0.002 cm^{-3}) intercloud medium composed of overlapping supernova bubbles filling most of the ISM (filling factor of f=0.7–0.8).

The **second phase** is the "warm medium" composed of both ionized and neutral components, both at temperatures of 8000 K and with densities of 0.1–1 cm^{-3}. The ionization fraction varies between x<0.15 ("neutral") to x≥0.68 ("ionized").

The **first phase** is the "Cold Neutral Medium", consisting of cold (T=80 K), dense (n≥40 cm^{-3}) HI clouds with an ionization fraction of x≤0.001. This phase fills only a small volume of the ISM (f≈0.05), but it contains most of the *mass* of the ISM.

The basic assumptions of the MO 3-phase model are:

1. **Supernovae** heat and disrupt the entire ISM (porosity, q, is >1)

2. **Local thermal Pressure Balance** between the different phases
3. **Mass Exchange** occurs between the hot and cold phases:
   - Supernova blast waves run over cold clouds, ablating them and adding their mass to the hot bubble
   - Old supernova bubbles cool and the shells condense back into cold clouds.

The important difference with the FGH 2-phase model is the predominant role of collisional heating due to supernovae in the MO picture. FGH assumed that ionization from the ISRF was the determining factor in the ionization balance of the cold and warm phases.

The problem is, neither picture is complete or correct.

Assumption #1: Supernovae disrupt & dominate ISM physics.

Slavin & Cox (1993, ApJ, 417, 187) re-examined supernova models and porosity estimates, and found

$$q = 0.176 S_{-13} E_{51}^{1.17} n_0^{-0.61} \tilde{p}_{04}^{-1.06}$$

Using more recent estimates of *local* SN rates and blast energies yields numbers like $S_{-13} \approx 0.4$, $E_{51} \approx 0.75$ (midrange of 0.5–1).

A better estimate of the density for the arm neutral and ionized mediums gives $n_0 \approx 0.1–0.2$ cm^{-3}, and if you now include an estimate of the magnetic field pressure contribution:

$$\tilde{p}_{04} = 10^{-4}\left[\frac{B_0^2}{8\pi k} + (n_e + n_p)_0 T_0\right] \text{ cm}^{-3}\text{ K}$$

$$\approx 9000 \text{ cm}^{-3}\text{ K for } T_0 = 10^4 \text{ K \& } B_0 = 5\mu G$$

Then the porosity is $q \approx 0.18$, or 18% filling factor, which is way down from the $q > 3$ computed by McKee & Ostriker.

In sum, this makes supernovae important enough to be interesting, but maybe not dominant.

Assumption #2: Local thermal pressure balance between phases.

EUVE observations of the local ISM have found a "shadow region" that permitted measurements of the hot component near the local warm medium (Bowyer et al. 1995, Nature, 375, 212). From this they measured the properties of the warm and hot components:

Warm Component: $(P/k)_{warm} = 730 \pm 30$ cm^{-3} K

Hot (7×10^5 K) Component: $(P/k)_{hot} = 19000$ cm^{-3} K

This demonstrates a factor of 26 thermal pressure mismatch between two regions in contact with each other. This suggests that pressure balance is *not* dominated by thermal pressure, but that other pressures (e.g., magnetic fields, cosmic rays, etc.) must be important.

Assumption #3: Significant mass transfer between the hot and cold phases.

Even modest magnetic fields could inhibit blast heating of cold clouds, shutting down mass exchange and thermal contact. We see little direct observational evidence in our own galaxy (or others) for mass exchange between supernova blast waves running over cold clouds.

The bottom line is that there is as yet no viable general model of the ISM, although we a lot of the observational and conceptual pieces. It is clear that thermal pressure equilibrium is not a good tool, and we must consider other pressure sources, the main contenders of which (magnetic fields, turbulence, and cosmic rays) all have comparable strengths, so no clear leading source emerges. The ISM is far more dynamic than the earlier models assumed, making it a fruitful area for new research.

Summary: The Five Thermal Phases of the ISM

Overall, interstellar gas can be in one of 5 thermal phases, roughly in order from coolest to hottest:

Molecular Clouds: H_2

These are H_2 clouds at temperatures of 10-20K, and densities $>10^3$ cm^{-3}. Molecular clouds comprise ~30% of the mass of the ISM, but occupy only ~0.05% of its volume. Most molecular clouds are gravitationally bound. The densest cores are likely unstable and sites of new star formation. The main tracers are mm-wavelength molecular emission lines (primarily CO).

Cold Neutral Medium (CNM): HI absorption

Cold neutral hydrogen (HI) gas is distributed in sheets and filaments occupying ~1-4% of the ISM with temperatures of ~80-100K and densities of ~50 cm^{-3}. The main tracers are UV and optical absorption lines seen towards bright stars or quasars. The CNM is approximately in pressure equilibrium with its surroundings.

Warm Neutral Medium (WNM): HI emission

Warm neutral atomic hydrogen occupies ~30% of the volume of the ISM, and is located mainly in photodissociation regions on the boundaries of HII regions and molecular clouds. It has characteristic temperatures of ~8000K and densities of ~0.5 cm^{-3}, and is traced by HI 21cm emission lines. It is often called the "warm intercloud medium" in some older papers.

Warm Ionized Medium (WIM): HII emission

Diffuse gas with temperatures of 6000-12000K, and densities ~0.1cm^{-3} occupying about ~25% of the volume of the ISM. While primarily photoionized (it requires about 1/6th of all of the ionizing photons emitted by the Galaxy's O and B stars), there is some evidence of shock or collisional ionization high above the plane of the Galaxy. It is traced by low-surface brightness Hα λ6563Å emission. Nearly 90% of the H^+ in the Galaxy resides in the WIM, with the remaining 10% in the bright high-density HII regions that occupy a tiny fraction of the ISM.

Hot Ionized Medium (HIM): X-ray and OIV absorption

Hot, low-density gas heated by supernovae, with temperatures $>10^6$ K and very low densities of <0.003 cm^{-3} occupying ~50% of the ISM. The vertical scale height of this gas is ~3kpc, so it is sometimes referred to in the literature as the hot "**corona**" of the galaxy. This hot gas is often buoyant and appears as bubbles and fountains high above the disk. Its primary tracers are absorption lines seen towards hot stars in the far-UV (e.g., OIV, NV, and CIV) in gas with T$\approx 10^5$ K, and diffuse soft X-ray emission from gas hotter than 10^6 K.

Dust

Mixed in with all but the very hottest phases is interstellar dust. Dust grains are an important source of interstellar extinction, gas-phase element depletion, sites of interstellar chemistry, etc. This is solid-phase rather than gas-phase material. Dust grains range in size from a few microns down to macromolecular scales (clumps of 50–100 atoms or less). We will treat dust in detail in a later chapter. While we will spend a lot of our time talking about gas-phase processes, it is important to keep dust in mind, since it plays a role often disproportionate to its share of the mass of the ISM.