The interstellar environment of our galaxy

Katia M. Ferrière

Observatoire Midi-Pyrénées, 31400 Toulouse, France

(Published 5 December 2001)

This article reviews the current knowledge and understanding of the interstellar medium of our galaxy. The author first presents each of the three basic constituents—ordinary matter, cosmic rays, and magnetic fields—of the interstellar medium, with emphasis on their physical and chemical properties as inferred from a broad range of observations. The interaction of these interstellar constituents, both with each other and with stars, is then discussed in the framework of the general galactic ecosystem.

CONTENTS

I. Introduction 1031
II. Historical Background 1032
   A. Overall picture of the galaxy 1032
   B. The interstellar medium 1033
III. Interstellar Matter 1034
   A. General properties 1034
   B. Molecular gas 1035
   C. Neutral atomic gas 1037
   D. Warm ionized gas 1042
   E. Hot ionized gas 1045
   F. Dust 1046
IV. Interstellar Magnetic Fields and Cosmic Rays 1048
   A. Magnetic fields 1048
   B. Cosmic rays 1050
   C. Synchrotron radiation 1052
V. How Everything Fits Together 1054
   A. Role played by stars 1054
   B. Supernova parameters 1056
   C. Role played by cosmic rays and magnetic fields 1058
Acknowledgments 1059
Appendix 1059
References 1061

I. INTRODUCTION

The stars of our galaxy—traditionally referred to as “the Galaxy” with a capital G to distinguish it from the countless other galaxies—are embedded in an extremely tenuous medium, the so-called “interstellar medium” (ISM), which contains ordinary matter, relativistic charged particles known as cosmic rays, and magnetic fields. These three basic constituents have comparable pressures and are intimately bound together by electromagnetic forces. Cosmic rays and magnetic fields influence both the dynamics of the ordinary matter and its spatial distribution at all scales, providing, in particular, an efficient support against the gravitational force. Conversely, the weight of the ordinary matter confines magnetic fields and, hence, cosmic rays to the Galaxy, while its turbulent motions can be held responsible for the amplification of magnetic fields and for the acceleration of cosmic rays.

The ISM encloses but a small fraction of the total mass of the Galaxy. Moreover, it does not shine in the sky as visibly and brightly as stars do. Yet it plays a vital role in many of the physical and chemical processes taking place in the Galactic ecosystem.

The most important aspect of Galactic ecology is probably the cycle of matter from the ISM to stars and back to the ISM. In the first step of this cycle, new stars form out of a reservoir of interstellar material. This material, far from being uniformly spread throughout interstellar space, displays dramatic density and temperature contrasts, such that only the densest, coldest molecular regions can offer an environment favorable to star formation. In these privileged sites, pockets of interstellar gas, losing part of their magnetic support, tend to become gravitationally unstable and collapse into new stars.

Once locked in the interior of stars, the Galactic matter goes through a succession of thermonuclear reactions, which enrich it in heavy elements. A fraction of this matter eventually returns to the ISM, be it in a continuous manner via powerful stellar winds or in an instantaneous manner via supernova explosions (violent stellar outbursts resulting from a thermonuclear instability or from the sudden gravitational collapse of the core of some stars at the end of their lifetimes). In both cases, the injection of stellar mass into the ISM is accompanied by a strong release of energy, which, in addition to generating turbulent motions, contributes to maintaining the highly heterogeneous structure of the ISM and may, under certain circumstances, give birth to new molecular regions prone to star formation. This last step closes the loop of the partly self-induced ISM-star cycle.

Thus the interstellar medium is not merely a passive substrate within which stars evolve; it constitutes their direct partner in the Galactic ecosystem, continually exchanging matter and energy with them and controlling many of their properties. It is the spatial distribution of the interstellar material together with its thermal and chemical characteristics that determines where new stars form as well as their mass and luminosity spectra. These, in turn, govern the overall structure, the optical appearance, and the large-scale dynamics of the Galaxy. Hence understanding the present-day properties of our Galaxy and being able to predict its long-term evolution require a good knowledge of the dynamics, energetics, and chemistry of the ISM.

*Electronic address: ferriere@ast.obs-mip.fr
The purpose of the present review is precisely to describe the current status of this knowledge. Major advancements have been made over the last few years, thanks in large part to steady improvements in instrumentation, in observation and analysis techniques, and in computer power. A number of recent large-scale surveys in different wavelength bands have allowed astronomers to obtain complementary images with unprecedented spatial coverage, resolution, and sensitivity, often supplemented by valuable spectral information. These surveys have also stimulated numerous theoretical studies aimed at interpreting the observed phenomena within the context of a well-understood model and using them to place observational constraints on the various processes at work.

We shall start with a brief historical overview, emphasizing the main developments that paved the way to the modern view of our Galaxy (Sec. II). We shall then present the three basic constituents of the ISM, namely, ordinary matter (Sec. III), cosmic rays, and magnetic fields (Sec. IV). Finally, we shall discuss the interplay among these three constituents and their relations with stars (Sec. V).

II. HISTORICAL BACKGROUND

A. Overall picture of the galaxy

To a terrestrial observer, the Galaxy appears (on starry nights only) as a faint band of diffuse light stretching all the way around the sky; this is why it is also known as the Milky Way galaxy or simply the Milky Way. This denomination goes back to ancient Greek civilization and, in particular, to Claude Ptolemy (90–168), who provided one of the very first descriptions of the Milky Way, qualifying it as “a zone as white as milk.” However, the true nature of the Milky Way was not established until 1610, when Galileo Galilei (1564–1642), examining it for the first time through his telescope, discovered that it was actually composed of innumerable dim stars.

It became clear in the course of the eighteenth century that we see the Milky Way as a narrow band encircling us because the Galaxy has the shape of a flattened disk, in which we are deeply embedded. At the end of the century, William Herschel (1738–1822) undertook a systematic study of the distribution of stars across the sky. Since he relied on a couple of faulty assumptions—all stars have approximately the same intrinsic brightness and interstellar space is completely transparent to starlight—he erroneously concluded that the Galaxy is about five times more extended in its plane than in the perpendicular direction and that the Sun is located near the Galactic center. These conclusions were corroborated a good century later by Jacobus Kapteyn (1851–1922), based on the far more abundant stellar data available in his time. Going one step further, Kapteyn also estimated the spatial distribution of stars within the Galaxy together with its overall size; he thus obtained exponential scale lengths at half maximum of ~1.2 kpc in the radial direction and ~0.22 kpc in the vertical direction.

A totally different picture of the Galaxy emerged during World War I, from Harlow Shapley’s (1885–1972) observational work on globular clusters (compact, nearly spherical groupings of \(10^5–10^7\) stars). He noticed that, unlike ordinary stars, globular clusters do not spread uniformly along the Milky Way, but instead concentrate in the direction of the Sagittarius constellation. He further found that they have a roughly spherical distribution, the center of which, he argued, should approximately coincide with the center of the Galaxy. This is how his investigation led him to the radical conclusion that the Sun lies very far from the Galactic center, at a distance of about 15 kpc.

Strong support for Shapley’s picture came in the mid-1920s from the kinematical studies of Bertil Lindblad (1895–1965) and Jan Oort (1900–1992), which convincingly showed that the observed relative velocities of stars and globular clusters with respect to the Sun could be readily understood in the framework of a differentially rotating Galactic model with the Sun placed at the radial distance predicted by Shapley, whereas they were hard to reconcile with the total amount of mass inferred from Kapteyn’s model. Almost three decades later, radio-astronomical measurements of the spatial distribution of interstellar neutral hydrogen delivered the definitive proof that Shapley was correct about the off-center position of the Sun in the Galaxy, while demonstrating that he had overestimated its Galactocentric radius by almost a factor of 2.

We now know that our Galaxy comprises a thin disk with radius ~25–30 kpc and effective thickness ~400–600 pc, plus a spherical system itself composed of a bulge with radius ~2–3 kpc and a halo extending out to more than 30 kpc from the center (Binney and Merrifield, 1998, p. 606). The Sun resides in the Galactic disk, approximately 15 pc above the midplane (Cohen, 1995; Magnani et al., 1996) and 8.5 kpc away from the center (Kerr and Lynden-Bell, 1986).

The stars belonging to the disk rotate around the Galactic center in nearly circular orbits. Their angular rotation rate is a decreasing function of their radial distance. At the Sun’s orbit, the rotation velocity is \(\approx 220 \text{ km s}^{-1}\) (Kerr and Lynden-Bell, 1986), corresponding to a rotation period of about \(240 \times 10^6\) years. Disk stars also have a velocity dispersion of \(\sim 10–40 \text{ km s}^{-1}\).

1Because of the Earth’s annual revolution about the Sun, a nearby star seems to trace out an ellipse in the plane of the sky with respect to the very distant background stars. The parallax is defined as the angle under which the semimajor axis of this apparent ellipse is seen from Earth, and the parsec (pc) is, by definition, the length unit equal to the distance at which a star has a parallax of one second of arc. This length unit is the one most commonly used by astronomers. Nonetheless, in the context of individual galaxies, it often proves more convenient to employ the kiloparsec (kpc). For future reference, 1 kpc = 1000 pc = 3260 light years = \(3.09 \times 10^{16}\) km.
with a variety of shapes and sizes. It was soon realized that the spatial density of stars falls off in all directions over a continuum and at specific wavelengths. It was, incidentally, the presence of obscuring interstellar material that gave Herschel and Kapteyn the false impression that they were produced in a single cloud of “cold” \((T<1000\text{K})\) interstellar gas somewhere between \(\delta\) Orionis and the Earth. The subsequent detection of absorption lines with multiple narrow peaks in the spectrum of other stars provided further evidence for the existence of cold interstellar gas clumped into distinct clouds, the line multiplicity being naturally attributed to the presence of several intervening clouds with different line-of-sight velocities (see, for example, Beals, 1936; Adams, 1949).

Astronomers also suspected the existence of less conspicuous “diffuse clouds,” especially after Hartmann’s (1904) discovery of stationary absorption lines of once-ionized calcium (Ca II) in the spectrum of the spectroscopic binary \(\delta\) Orionis. As in any spectroscopic binary system, the spectral lines created by the two companion stars orbiting around each other undergo a periodically varying Doppler shift, resulting from the back-and-forth motion of the stars along the line of sight. Hence the stationary Ca II lines could not arise from \(\delta\) Orionis itself, but instead must have an interstellar origin. In addition, their single-peaked shape together with their narrow width strongly suggested that they were produced in a single cloud of “cold” \((T<1000\text{K})\) interstellar gas somewhere between \(\delta\) Orionis and the Earth. The more distant clusters to be larger, regardless of the considered direction. Since this tendency could not be considered real—otherwise the Sun would have a special position within the Galaxy—Trumpler was led to conclude that the light from remote clusters is gradually dimmed, as it propagates through interstellar space, by the obscuring action of a pervasive interstellar material. Here, too, the precise agent causing this general obscuration is interstellar dust.

The obscuration process due to interstellar dust is more effective at shorter wavelengths, which are closer to the typical grain sizes (see Sec. III.F), so that blue light is more severely dimmed than red light. In consequence, the light emitted by a faraway star appears to us.

---

*Mihalas and Binney, 1981, p. 423,* which causes them to execute small oscillations about a perfectly circular orbit, both in the Galactic plane (epicycles) and in the vertical direction. In contrast, the stars present in the bulge and in the halo rotate slowly and often have very eccentric trajectories.

Radio-astronomical observations of interstellar neutral hydrogen indicate that the Milky Way possesses a spiral structure, similar to that seen optically in numerous external galaxies. These “spiral galaxies” typically exhibit two spiral arms unwinding either directly from the central bulge or from both ends of a bar crossing the bulge diametrically. The exact spiral shape of our own Galaxy is difficult to determine from within; the best radio data to date point to a structure characterized by a bulge of intermediate size and a moderate winding of the arms (type Sbc in Hubble’s classification; Binney and Merrifield, 1998, p. 171), while recent infrared (IR) images of the Galactic center region clearly display the distinctive signature of a bar (Blitz and Spergel, 1991; Dwek *et al.*, 1995). Our position in the spiral pattern can be derived from local optical measurements, which give a quite accurate outline of the three closest arms; they locate the Sun between the inner Sagittarius arm and the outer Perseus arm, near the inner edge of the local Orion-Cygnus arm (Mihalas and Binney, 1981, p. 248).

---

**B. The interstellar medium**

The Milky Way system is not only made of stars; it also contains significant amounts of tenuous matter, inhomogeneously spread out throughout interstellar space. The interstellar matter, which exists in the form of gas (atoms, molecules, ions, and electrons) and dust (tiny solid particles), manifests itself primarily through obscuration, reddening, and polarization of starlight, through the formation of absorption lines in stellar spectra, and through various emission mechanisms (both over a continuum and at specific wavelengths). It was, incidentally, the presence of obscuring interstellar material that gave Herschel and Kapteyn the false impression that the spatial density of stars falls off in all directions away from us and thus led them to misplace the Sun near the center of the Galaxy. Shapley did not encounter the same problem with globular clusters, both because they are intrinsically much brighter and easier to recognize than individual stars and because most of them lie outside the thin layer of obscuring material.

Herschel, back in the late eighteenth century, had already noticed that some regions in the sky, particularly along the Milky Way, seemed devoid of stars. The first long-exposure photographs of the Milky Way taken by Edward Barnard (1857–1923) in the early days of astronomical photography revealed many more dark zones with a variety of shapes and sizes. It was soon realized that these apparent holes in the stellar distribution were due to the presence, along the line of sight, of discrete “clouds” of interstellar matter hiding the stars situated behind them. More specifically, it is the interstellar dust contained in these “dark clouds” that either absorbs or scatters the background starlight, the combination of these two processes being commonly called interstellar obscuration or extinction.

---

2In fact, the local Orion-Cygnus arm is probably a short spur rather than a major spiral arm like Sagittarius and Perseus (Blitz *et al.*, 1983).
TABLE I. Descriptive parameters of the different components of the interstellar gas, according to the references quoted in the main text. \( T \) is the temperature, \( n \) is the true (as opposed to space-averaged) number density of hydrogen nuclei near the Sun, \( \Sigma \) is the azimuthally averaged mass density per unit area at the solar circle, and \( M \) is the mass contained in the entire Milky Way. Both \( \Sigma \) and \( M \) include 70.4% hydrogen, 28.1% helium, and 1.5% heavier elements. All values were rescaled to \( R_\odot = 8.5 \) kpc, in accordance with footnote 3.

<table>
<thead>
<tr>
<th>Component</th>
<th>( T ) (K)</th>
<th>( n ) (cm(^{-3}))</th>
<th>( \Sigma ) (( M_\odot ) pc(^{-2}))</th>
<th>( M ) (( 10^4 )( M_\odot ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular</td>
<td>10–20</td>
<td>( 10^5–10^6 )</td>
<td>( \sim 2.5 )</td>
<td>( \sim 1.3^{a}–2.5^{b} )</td>
</tr>
<tr>
<td>Cold atomic</td>
<td>50–100</td>
<td>20–50</td>
<td>( \sim 3.5 )</td>
<td></td>
</tr>
<tr>
<td>Warm atomic</td>
<td>6000–10000</td>
<td>0.2–0.5</td>
<td>( \sim 3.5 )</td>
<td>( \sim 1.4 )</td>
</tr>
<tr>
<td>Warm ionized</td>
<td>( \sim 8000 )</td>
<td>0.2–0.5</td>
<td></td>
<td>( \geq 0.6 )</td>
</tr>
<tr>
<td>Hot ionized</td>
<td>( \sim 10^6 )</td>
<td>( 0.0065 )</td>
<td></td>
<td>( \geq 1.6 )</td>
</tr>
</tbody>
</table>

\( ^a \)Adapted from Bronfman et al., 1988.
\( ^b \)Adapted from Clemens et al., 1988.

redder than it actually is. This reddening effect can be measured by comparing the observed apparent color of the star to the theoretical color corresponding to its spectral type. By means of such measurements, Trumpler (1930) was able to show that the reddening of stars of a given spectral type increases with their distance from us, thereby providing conclusive proof that interstellar space is indeed pervaded by an obscuring, dust-bearing, interstellar material.

Another manifestation of interstellar dust, equally linked to its obscuration properties, is the linear polarization of starlight. The polarization effect, uncovered more than a few years to realize that the ISM is also filled with cosmic rays. Although the existence of cosmic rays outside the Earth’s atmosphere had been known since the balloon experiment conducted by Hess (1919), the Galactic origin of the majority of them and their widespread distribution throughout the Milky Way were not recognized until the observed Galactic radio emission was correctly identified with synchrotron radiation by cosmic-ray electrons gyrating about magnetic-field lines (Ginzburg and Syrovatskii, 1965).

III. INTERSTELLAR MATTER

A. General properties

Interstellar matter accounts for \( \sim 10–15 \% \) of the total mass of the Galactic disk. It tends to concentrate near the Galactic plane and along the spiral arms, while being very inhomogeneously distributed at small scales. Roughly half of the interstellar mass is confined to discrete clouds occupying only \( \sim 1–2 \% \) of the interstellar volume. These interstellar clouds can be divided into three types: the dark clouds, which are essentially made of very cold (\( T \sim 10–20 \) K) molecular gas and block off the light from background stars, the diffuse clouds, which consist of cold (\( T \sim 100 \) K) atomic gas and are almost transparent to the background starlight, except at a number of specific wavelengths where they give rise to absorption lines, and the translucent clouds, which contain molecular and atomic gases and have intermediate visual extinctions. The rest of the interstellar matter, spread out between the clouds, exists in three different forms: warm (mostly neutral) atomic, warm ionized, and hot ionized, where warm refers to a temperature of \( \sim 10^4 \) K and hot to a temperature of \( \sim 10^6 \) K (see Table I).

By terrestrial standards, interstellar matter is exceedingly tenuous: in the vicinity of the Sun, its density varies from \( \sim 1.5 \times 10^{-26} \) g cm\(^{-3}\) in the hot medium to \( \sim 2 \times 10^{-20}–2 \times 10^{-18} \) g cm\(^{-3}\) in the densest molecular regions, with an average of about \( 2.7 \times 10^{-24} \) g cm\(^{-3}\) (see the following subsections). This mass density, which corresponds to approximately one hydrogen atom per cubic centimeter, is over 20 orders of magnitude lower than in the Earth’s lower atmosphere.

The chemical composition of interstellar matter is close to the “cosmic composition” inferred from abundance measurements in the Sun, in other disk stars, and in meteorites, namely, 90.8% by number (70.4% by mass) of hydrogen, 9.1% (28.1%) of helium, and 0.12% (1.5%) of heavier elements, customarily termed “metals” in the astrophysical community (Spitzer, 1978, p. 4). However, observations of interstellar absorption lines in the spectra of hot stars indicate that a significant fraction of these heavier elements is often missing or “depleted” from the gaseous phase of the ISM, being, in all likelihood, locked up in solid dust grains. Since the first systematic studies of interstellar elemental abundances along different sight lines (Morton et al., 1973; Rogerson et al., 1973), depletion factors have been seen to vary appreciably across the sky, presumably due to the wide fluctuations in environmental physical conditions. As a general rule, depletions tend to be more severe in regions with higher density and lower temperature (Jenkins, 1987; Van Steenberg and Shull, 1988); they also seem to depend weakly on the degree of ionization, insofar as they are somewhat less severe in the warm ion-
ized medium than in the warm neutral medium (Howk and Savage, 1999). On average, the most common “metals,” C, N, and O, are only depleted by factors of $\sim 1.2-3$, whereas refractory elements like Mg, Si, and Fe are depleted by factors of $\sim 10-100$ (Savage and Sembach, 1996). Altogether, about 0.5–1% of the interstellar matter by mass is in the form of dust rather than gas.

In the following subsections, we focus on the interstellar gas and successively describe the five different forms under which it can be found: molecular, cold atomic, warm atomic, warm ionized, and hot ionized. The last subsection is devoted to a description of the interstellar dust.

B. Molecular gas

The first interstellar molecules (CH, CH$^+$, and CN) were discovered in the late 1930s, through the optical absorption lines they produce in stellar spectra. However, it was not until 1970, when ultraviolet (UV) astronomy from above the Earth’s atmosphere had just opened a new window on the universe, that the most abundant interstellar molecule, H$_2$, was for the first time detected in the far-UV spectrum of a hot star (Carruthers, 1970). The next most abundant molecule, CO, was identified in a UV stellar spectrum the following year (Smith and Stecher, 1971).

These discoveries were succeeded in 1972 by the launch of a UV spectrometer on the Copernicus satellite, which prompted many observational studies on the interstellar molecular gas (see Spitzer and Jenkins, 1975, for a preliminary review of the Copernicus results). The Copernicus survey of H$_2$ absorption by Savage et al. (1977) provided the H$_2$ column density (number of H$_2$ molecules in a cylinder of unit cross section along the line of sight) between the Earth and 109 nearby hot stars and led to first estimates of the space-averaged (i.e., smoothed-out) density and temperature of the molecular gas near the Sun.

A wealth of additional information on the spatial distribution and physical properties of the molecular gas is expected from the Far Ultraviolet Spectroscopic Explorer (FUSE) satellite, which was launched in 1999 and will measure key absorption lines in the far-UV spectrum of hundreds of Galactic and extragalactic sources, with a much higher sensitivity than Copernicus (Moos et al., 2000); early release results on interstellar H$_2$ have been reported by Shull et al. (2000) and Snow et al. (2000).

Observations of optical and UV absorption lines, crucial as they are for our understanding of interstellar molecules, do not allow astronomers to probe the interior of dense molecular clouds, for the bright sources necessary to make absorption measurements are obscured by the interstellar dust present in the very regions to be probed. In order to explore the structure and large-scale distribution of the molecular gas, one may take advantage of the fact that radio waves are not subject to interstellar extinction and appeal to radio spectroscopy.

The H$_2$ molecule itself is not directly observable at radio wavelengths: because it possesses no permanent electric dipole moment and has a very small moment of inertia, all its permitted transitions lie outside the radio domain (Field et al., 1966). The CO molecule, for its part, has a $J=1\rightarrow 0$ rotational transition at a radio wavelength of 2.6 mm; the corresponding emission line, which was first observed a few months before the detection of CO in UV absorption (Wilson et al., 1970), has become the primary tracer of molecular interstellar gas (see, for example, Scoville and Sanders, 1987). The technique employed to deduce the molecular spatial distribution at Galactic scales from the profile of the CO 2.6-mm emission line relies on the Galactic rotation curve; a detailed description of the method is given in the Appendix.

The first large-scale surveys of CO 2.6-mm emission, carried out by Scoville and Solomon (1975) and by Burton et al. (1975), covered only a thin band along the Galactic equator over longitude intervals (as defined in Fig. 10 below) accessible from the northern terrestrial hemisphere. Nevertheless, they already showed that most of the molecular gas resides in a well-defined ring extending radially between 3.5 kpc and 7 kpc from the Galactic center, and they unveiled a strong molecular concentration in the region interior to 0.4 kpc.

Subsequent, more extensive CO surveys allowed a finer description of the molecular gas radial distribution and added information on its azimuthal and vertical distributions as well as on its small-scale structure. Dame et al. (1987) assembled data from 5 large and 11 more restricted surveys to construct a synoptic picture of the whole Milky Way. Although they made no attempt to convert the dependence on line-of-sight velocity (which may be considered an observable, being directly related to the measurable Doppler shift) into a dependence on heliocentric distance (by means of the Galactic rotation curve; see the Appendix), they were able to bring to light the spiral pattern of CO emission. Indeed, they found that CO concentrations in the longitude-velocity plane tend to follow the strips corresponding to the spiral arms observed in the 21-cm emission of neutral interstellar hydrogen (see Sec. III.C). Let us mention that, in the continuation of their 1987 work, Dame et al. (2001) produced a new composite CO survey of the entire Galactic disk with about 2–4 times better angular resolution and up to 10 times higher sensitivity per unit solid angle.

3These early surveys assumed $R_\odot=10$ kpc for the Galactocentric radius of the Sun, whereas the latest value recommended by the International Astronomical Union is 8.5 kpc (Kerr and Lynden-Bell, 1986). In consequence, we have scaled down the lengths inferred from these surveys by a factor of 0.85. Likewise, throughout this paper, all parameters taken from observational studies of the spatial distribution of interstellar constituents will systematically be rescaled to $R_\odot =8.5$ kpc. For reference, distances scale as $R_\odot^{-1}$, surface densities as $R_\odot^{-1}$, volume densities as $R_\odot^{-2}$, and masses as $R_\odot^{-1}$. 
The precise horizontal distribution of the molecular material in the first Galactic quadrant (quadrant I in Fig. 10 below) was investigated by Clemens et al. (1988). They contrived a means to overcome the near-far distance ambiguity for emission inside the solar circle (see the Appendix), which enabled them to draw a detailed face-on map of interstellar CO. This map is dominated by the molecular ring peaking at a Galactic radius \( R \approx 4.5 \) kpc and, to a lesser extent, by two discrete features closely associated with 21-cm spiral arms. Interarm regions, though, are not devoid of molecular gas: their inferred \( \text{H}_2 \) space-averaged density is, on average over the first quadrant, only a factor of \( \sim 3.6 \) lower than in the arms.

The situation is apparently different in the outer Galaxy \( (R > R_\odot) \). CO surveys of relatively extended portions of the sky have provided detailed images in which molecular concentrations clearly line up along the 21-cm spiral arms. However, the contrast in molecular surface density between spiral arms and interarm regions is much greater than in the inner Galaxy, with a mean value \( \sim 13:1 \) both in the longitude range \( 270^\circ - 300^\circ \) (Grabelsky et al., 1987) and in the longitude range \( 102^\circ - 142^\circ \) (Heyer, 1999). Such large density contrasts imply that, outside the solar circle, the bulk of the molecular gas belongs to the spiral arms.

Bronfman et al. (1988) combined two separate CO surveys of the first and fourth Galactic quadrants and fitted their data to an axisymmetric model of the space-averaged density of the molecular gas as a function of Galactic radius \( R \) and height \( Z \) over the radial range \( 2 \) kpc–\( R_\odot \). In Fig. 1, we display their result for the radial dependence of the \( \text{H}_2 \) column density, \( N_m(R) \), defined as the number of hydrogen nuclei tied up in \( \text{H}_2 \) molecules per unit area on the Galactic plane. For completeness, the \( y \) axis is labeled both in terms of \( N_m(R) \) and in terms of the corresponding mass density per unit area, \( \Sigma_m(R) = 1.42 m_p N_m(R) \), where \( m_p \) is the proton rest mass. Note that the strong central molecular peak, which falls outside the radial range explored by Bronfman et al., does not appear on the figure.

Also shown in Fig. 1 is the azimuthally averaged curve derived by Clemens et al. (1988) for the first Galactic quadrant. This curve differs markedly from that obtained by Bronfman et al. (1988), which, we recall, applies to the combined first and fourth quadrants: it lies everywhere higher and has a more pronounced maximum at the location of the molecular ring. Part of the difference reflects genuine large-scale departures from axisymmetry and can be attributed to the existence of molecular spiral arms, which, due to their unwinding shape, cross the first and fourth quadrants at different Galactic radii. However, the fact that the curve of Clemens et al. also differs from the fit of Bronfman et al. to the first quadrant alone reveals a true discrepancy between both studies, which Bronfman et al. explained in terms of differences in instrumental calibrations, in statistical treatments, and in the adopted CO/\( \text{H}_2 \) ratio.

Beyond \( R_\odot \), the \( \text{H}_2 \) column density averaged over azimuthal angle drops off rapidly outward. To our knowledge, there exists no quantitative estimate of its exact \( R \) dependence, except in restricted longitude intervals (see, for example, Grabelsky et al., 1987; Clemens et al., 1988), whose characteristics are not necessarily representative of average properties along full Galactic circles. The problem with partial azimuthal averages, already manifest in the inner Galaxy, probably becomes even worse beyond the solar circle, where arm-interarm density contrasts are more important.

In the vertical direction, the molecular gas appears to be strongly confined to the Galactic plane, and its space-averaged distribution can be approximated by a Gaussian. Upon averaging over azimuthal angle in the first Galactic quadrant, Clemens et al. (1988) found that the full width at half maximum (FWHM) of the molecular layer increased outward as \( R^{0.58} \) and had a value of \( 136 \pm 17 \) pc at the solar circle. They pointed out that the observed thickening of the molecular layer with increasing radius is consistent with a decreasing stellar mass surface density—which entails a decreasing gravitational pull toward the Galactic plane—together with an approximately constant \( \text{H}_2 \) velocity dispersion along \( Z \). At \( R_\odot \), they obtained for the space-averaged number density of hydrogen nuclei in molecular form

\[
\langle n_m(Z) \rangle = \langle n_m(0) \rangle \exp \left[-\frac{Z}{H_m} \right],
\]

with \( \langle n_m(0) \rangle = 0.58 \) cm\(^{-3} \) and \( H_m = 81 \) pc. The axisymmetric model of Bronfman et al. (1988), fitted to the combined first and fourth Galactic quadrant data, has a

---

**FIG. 1.** Column density \( N \) of interstellar hydrogen, defined as the number of hydrogen nuclei contained in a vertical cylinder of unit cross section through the Galactic disk and mass density per unit area of interstellar matter (including 70.4% hydrogen, 28.1% helium, and 1.5% heavier elements), \( \Sigma = 1.42 m_p N \), averaged over the Galactocentric azimuthal angle, as a function of Galactic radius \( R \) for the different gas components. Solid lines, contribution from molecular gas (thick line, from Bronfman et al., 1988; thin line, from Clemens et al., 1988); triple-dot-dashed line, contribution from cold+warm atomic gas (from Lockman, 1984; Dickey and Lockman, 1990; Diplas and Savage, 1991); dashed line, contribution from ionized gas outside the traditional HII regions (from Cordes et al., 1991, with a Gaussian radial scale length of 30 kpc for the thick component).
FIG. 2. Space-averaged number density \( \langle n \rangle \) of interstellar hydrogen nuclei and space-averaged mass density of interstellar matter (including 70.4% hydrogen, 28.1% helium, and 1.5% heavier elements), \( \langle \rho \rangle = 1.42 m_p \rho(n) \), as a function of Galactic height \( Z \) at the solar circle \( (R = R_\odot) \), for the different gas components: solid lines, contribution from molecular gas (thick line, from Bronfman \textit{et al.}, 1988; thin line, from Clemens \textit{et al.}, 1988); triple-dot-dashed line, contribution from cold+warm atomic gas (fromDickey and Lockman, 1990); dashed line, contribution from ionized gas outside the traditional \textsc{H} II regions (from Cordes \textit{et al.}, 1991, with a Gaussian radial scale length of 30 kpc for the thick component; in agreement with Reynolds, 1991, with an exponential scale height of 1 kpc).

FWHM of 120±18 pc, roughly independent of \( R \); at \( R_\odot \), the space-averaged density of molecular hydrogen is again given by Eq. (1), but with a midplane density \( \langle n_m \rangle(0) = 0.53 \text{ cm}^{-3} \) and a Gaussian scale height \( H_m = 71 \text{ pc} \). The vertical profile of \( \langle n_m \rangle \) or, equivalently, that of the space-averaged molecular mass density, \( \langle \rho_m \rangle = 1.42 m_p \langle n_m \rangle \), is drawn in Fig. 2, for both the studies of Bronfman \textit{et al.} and Clemens \textit{et al.}

High-resolution observations (see, for instance, Fig. 3) indicate that the molecular gas is contained in discrete clouds organized hierarchically from giant complexes (with a size of a few tens of parsecs, a mass of up to \( \times 10^9 M_\odot \)), and a mean hydrogen number density \( \sim 100–1000 \text{ cm}^{-3} \) down to small dense cores (with a size of a few tenths of a parsec, a mass \( \sim 0.3–10^3 M_\odot \)), and a mean hydrogen number density \( \sim 10^4–10^6 \text{ cm}^{-3} \) (Larson, 1981; Goldsmith, 1987). The majority of molecular clouds are sufficiently massive to be bound by self-gravity, and it can be verified that they approximately satisfy the virial balance equation, \( G M/R \sim \sigma^2 \), where \( M \), \( R \), and \( \sigma \) are the cloud mass, radius, and internal velocity dispersion, and \( G \) is the gravitational constant (Larson, 1981; Myers, 1987; but see also Maloney, 1990). They also roughly obey two empirical power-law relations, \( \sigma \sim R^{0.5} \) and \( M \sim \sigma^2 \), first obtained by Larson (1981) and confirmed by several subsequent studies (e.g., Solomon \textit{et al.}, 1987). Because of the large scatter in the observational data points, the normalization factors of these relations are ill defined. For reference, Solomon \textit{et al.} (1987) found \( \sigma_{1D} \approx 1 \text{ km s}^{-1} \) and \( M = 2000 M_\odot \) for \( R = 2 \text{ pc} \).

From measurements of the peak specific intensity of CO emission lines, it emerges that molecular clouds are, in general, extremely cold, with typical temperatures in the range 10–20 K (Goldsmith, 1987). Thermal speeds at these low temperatures \( (\sqrt{3kT/m_{H_2}} = 0.35–0.50 \text{ km s}^{-1}) \) are small compared to the measured internal velocity dispersions. This means that the total gas pressure inside molecular clouds has but a small contribution from its purely thermal component, the dominant contribution arising from internal turbulent motions. Moreover, in accordance with the notion that molecular clouds are gravitationally bound, the total gas pressure in their interior is much higher than in the intercloud medium.

\( \text{H}_2 \) molecules are believed to form by recombination of hydrogen atoms on the surface of interstellar dust grains (Hollenbach and Salpeter, 1971). The only regions where they can actually survive in vast numbers are the interiors of dark and translucent interstellar clouds (and possibly the deep interior of diffuse clouds), which are simultaneously shielded from radiative dissociation by external UV photons and cold enough to avoid collisional dissociation (Shull and Beckwith, 1982). The observed temperatures of molecular regions are easily explained as the result of thermal balance between heating by cosmic rays (and, at the cloud edges, collisions with photoelectrons from dust grains and with radiatively excited \text{H}_2 molecules) and cooling by molecular line emission (primarily CO), the rate of which increases steeply with increasing temperature (de Jong \textit{et al.}, 1980; Goldsmith, 1987; Hollenbach and Tielens, 1999). Collisions with dust grains also enter the thermal balance, either as a coolant or as a heat source, depending on the dust temperature with respect to that of the gas (Burke and Hollenbach, 1983).

The main descriptive parameters of the molecular interstellar gas are listed in Table I.

C. Neutral atomic gas

Neutral atomic hydrogen, usually denoted by \text{H} I (as opposed to \text{H} II for ionized hydrogen), is not directly observable at optical wavelengths. Under most interstellar conditions, particle collisions are so infrequent that nearly all hydrogen atoms have their electron in the ground energy level \( n = 1 \). It turns out that all the electronic transitions between the ground level and an excited state—forming the Lyman series—lie in the UV, with the Lyman \( \alpha \) (\( \text{L} \alpha \)) transition between the ground level and the first excited state \( n = 2 \) occurring at a wavelength of 1216 Å.

Since its initial detection from a rocket-borne spectrograph (Morton, 1967), the interstellar \( \text{L} \alpha \) line has been widely observed in absorption against background stars and used to study the \text{H} I distribution in the local ISM. The method consists of aiming at a great number of nearby hot stars distributed across the sky and ana-
lyzing their L\(\alpha\) absorption lines to deduce the H\textsc{i} column density between them and the Earth.

The early L\(\alpha\) survey undertaken by Savage and Jenkins (1972) and extended by Jenkins and Savage (1974) showed that H\textsc{i} is deficient in the immediate vicinity of the Sun, especially in the third Galactic quadrant (quadrant III in Fig. 10 below). We now understand the observed HI deficiency as a consequence of the Sun’s being located inside an H\textsc{i} cavity, known as the Local Bubble (Cox and Reynolds, 1987). This bubble is clearly asymmetric, not only in the Galactic plane, where it extends significantly farther in the longitude range 210°–250°, but also along the vertical, where it reaches much higher altitudes in the northern hemisphere (see, for example, Frisch and York, 1983). Based on a compilation of L\(\alpha\) absorption-line measurements by Fruscione \textit{et al.} (1994), Breitschwerdt \textit{et al.} (1996) estimated that the H\textsc{i} cavity has a radius in the plane \(\sim 60–100\) pc and a vertical extent from the plane \(\sim 120–180\) pc.

A much more detailed and accurate outline of the H\textsc{i} cavity was obtained by Sfeir \textit{et al.} (1999), who made use of the observational correlation between the column density of neutral sodium (Na\textsc{i}) inferred from its optical D-line doublet at 5890 Å and the H\textsc{i} column density inferred from the L\(\alpha\) line (Hobbs, 1974). They combined Na\textsc{i} column densities measured toward 456 stars with the improved stellar distances provided by the \textit{Hipparcos} satellite to draw contour maps of Na\textsc{i} absorption near the Sun. According to these maps, the Local Bubble has a radius in the plane varying between \(\sim 60\) pc toward Galactic longitude \(l = 0\) and \(\sim 250\) pc toward \(l = 235°\), it is elongated along the vertical and possibly open ended in the direction of the North Galactic Pole, and it is everywhere else surrounded by a dense wall of neutral gas.

Following Savage and Jenkins’s (1972) and Jenkins and Savage’s (1974) work, deeper, more reliable L\(\alpha\) absorption surveys with the \textit{Copernicus} satellite (Bohlin \textit{et al.}, 1978) and with the \textit{International Ultraviolet Explorer (IUE)} satellite (Shull and Van Steenberg, 1985) made it possible to proceed with methodical observations of the H\textsc{i} gas outside the Local Bubble and to gain a rough idea of its spatial distribution as a function of \(Z\) in the solar neighborhood. Unfortunately, the L\(\alpha\) line as

---

**FIG. 3.** High-resolution CO map of a 10.6°×8.4° portion of the sky centered on \((l = 112°, b = 1°)\). From the Five College Radio Astronomy Observatory (FCRAO) CO survey of the outer Galaxy (see Heyer \textit{et al.}, 1998). Figure courtesy of M. H. Heyer. [Color]
a diagnostic tool of H\textsc{i} is plagued by the same interstellar extinction problem as UV and optical molecular lines, which makes it unfit to map the H\textsc{i} distribution at Galactic scales. Here, too, one has to turn to radio astronomy.

The breakthrough event that opened the era of radioastronomical observations of interstellar H\textsc{i} was Ewen and Purcell's (1951) detection of the interstellar 21-cm line emission predicted seven years earlier by Hendrik van de Hulst. The existence of the 21-cm line results from the “hyperfine” structure of the hydrogen atom. In brief, the interaction between the magnetic moment of the electron and that of the proton leads to a splitting of the electronic ground level into two extremely close energy levels, in which the electron spin is either parallel or antiparallel (upper level) or antiparallel (lower level) to the proton spin. It is the “spin-flip” transition between these two energy levels that corresponds to the now famous 21-cm spin. It is the “spin-flip” transition between these two energy levels, in which the electron spin is either parallel or antiparallel (upper level) or antiparallel (lower level) to the proton spin. It is the “spin-flip” transition between these two energy levels that corresponds to the now famous 21-cm spin. It is the “spin-flip” transition between these two energy levels that corresponds to the now famous 21-cm spin.

In emission, the 21-cm line gives the total H\textsc{i} column density in the observed direction. Moreover, as explained in the Appendix, the contribution from each segment along the line of sight can be extracted from the shape of the line profile combined with the Galactic rotation curve. This is how 21-cm emission-line measurements covering the whole sky have been able to yield the H\textsc{i} space-averaged density as a function of position in the Galaxy.

H\textsc{i} maps projected onto the Galactic plane exhibit long arclike features organized into a spiral pattern (Oort et al., 1958; Mihalas and Binney, 1981, p. 528). Overall, this spiral pattern appears rather complex and fragmented, especially in the inner Galaxy ($R < R_\odot$) where the distance ambiguity problem (see the Appendix) further confuses the picture. Outside the solar circle, three major spiral arms clearly stand out in addition to the local minor Orion arm (Henderson et al., 1982; Kulkarni et al., 1982). According to Kulkarni et al. (1982), these arms have a roughly constant H\textsc{i} surface density, which is about four times greater than that in the interarm regions.

Radially, the H\textsc{i} gas extends out to at least 30 kpc from the Galactic center (Diplas and Savage, 1991). Its azimuthally averaged column density through the disk, $N_n(R)$, is characterized by a deep depression inside 3.5 kpc (Burton and Gordon, 1978), a relatively flat plateau through the solar circle (Lockman, 1984) and out to almost 14 kpc (Burton and te Lintel Hekkert, 1986; Diplas and Savage, 1991), and an exponential falloff beyond 14 kpc (Diplas and Savage, 1991). As a reminder, all the above lengths have been rescaled to R$_\odot = 8.5$ kpc. For the outer Galaxy, Diplas and Savage's (1991) study, based on a single large-scale H\textsc{i} survey, is limited to the longitude range $30^\circ - 250^\circ$, whereas both Henderson et al. (1982) and Burton and te Lintel Hekkert (1986) managed to achieve almost full longitudinal coverage by combining two complementary H\textsc{i} surveys. On the other hand, the H\textsc{i} data analyzed by Henderson et al. (1982) have a latitude cutoff at $|b| \leq 10^\circ$, which causes them to miss a significant fraction of the H\textsc{i} gas, particularly at large Galactic radii. Moreover, for technical reasons, the H\textsc{i} density contours of Henderson et al. (1982) and of Burton and te Lintel Hekkert (1986) become unreliable at a ten times higher density level than those of Diplas and Savage (1991). We therefore favor Diplas and Savage's (1991) results, which, parenthetically, can be verified as consistent with those of Henderson et al. (1982) out to a radius of $\sim 13$ kpc and with those of Burton and te Lintel Hekkert (1986) down to a density level of $\sim 10^{-2}$ cm$^{-3}$.

In Fig. 1, we plotted the composite function $N_n(R)$ constructed by smoothly connecting a constant function at Dickey and Lockman's (1990) best-estimate value of $N_n(R_\odot)$ (between 3.5 kpc and almost 14 kpc) to a vanishing function (at small R) and to an exponentially decreasing function averaged over the three directions whose best-fit parameters were tabulated by Diplas and Savage (1991; outside 14 kpc). The averaged exponential function should not be taken too seriously, both because the derived parameters are quite different in the three fitted directions, thereby suggesting that merely averaging over them does not lead to a trustworthy azimuthal average, and because the radial dependence of $N_n(R)$ outside the solar circle is very sensitive to the poorly known shape of the Galactic rotation curve. Despite the important uncertainties at large R, Fig. 1 underscores the stark contrast between the flat-topped profile of the H\textsc{i} gas and the peaked profile of the molecular gas.

The vertical structure of the H\textsc{i} distribution is roughly uniform for $3.5 \text{kpc} < R < R_\odot$ (Lockman, 1984). In this radial interval, the H\textsc{i} gas lies in a flat layer with a FWHM of 230 pc (almost twice the FWHM of the molecular gas at $R_\odot$), and its space-averaged number density can be approximated by the sum of two Gaussians and an exponential tail:

$$
\langle n_n \rangle (Z) = (0.57 \text{ cm}^{-3}) \left[ 0.70 \exp \left( - \frac{Z}{127 \text{ pc}} \right) \right] + 0.19 \exp \left[ - \frac{Z}{318 \text{ pc}} \right] + 0.11 \exp \left[ - \frac{|Z|}{403 \text{ pc}} \right]
$$

(Dickey and Lockman, 1990; see Fig. 2). The thickness of the H\textsc{i} layer drops to $\lesssim 100$ pc inside 3.5 kpc (Dickey and Lockman, 1990), and it grows more than linearly with $R$ outside $R_\odot$, reaching $\sim 3$ kpc at the outer Galactic boundary (Diplas and Savage, 1991). This substantial flaring, expected from the steep decrease in the vertical gravitational field, is accompanied by a general warping of the H\textsc{i} disk and by a regular scalloping of its outer edge (Henderson et al., 1982; Kulkarni et al., 1982).
whose physical origins are not well understood. The warp is such that the midplane of the H I layer lies above the Galactic equatorial plane in the first and second quadrants, with a maximum displacement of $\sim 4$ kpc, and below the Galactic plane in the third and fourth quadrants, with a maximum displacement of $\sim 2.15$ kpc (Dickey and Lockman, 1990; Diplas and Savage, 1991). The scalloping has an azimuthal wave number $n = 10$ and an amplitude comparable to that of the warp (Kulkarni et al., 1982).

The 21-cm absorption spectra generally look quite different from emission spectra taken in a nearby direction: while the emission spectra contain both distinct narrow peaks and much broader features, only the narrow peaks are present in the absorption spectra (see Fig. 4). The conventional interpretation of this difference is that the narrow peaks seen in emission and in absorption are produced by discrete cold ($T \leq 50–100$ K) H I clouds, whereas the broader features seen in emission only are due to a widespread H I gas that is too warm to give rise to detectable 21-cm absorption. The estimated temperature of the warm H I component is $\sim 6000–10000$ K (Dickey et al., 1978; Kulkarni and Heiles, 1987).

Comparisons between 21-cm emission and absorption measurements indicate that, in the vicinity of the Sun, the warm H I has roughly the same column density as the cold H I (Falgarone and Lequeux, 1973; Liszt, 1983) and about 1.5 times its scale height (Falgarone and Lequeux, 1973; Crovisier, 1978). The fraction of cold H I appears to remain approximately constant from $R_\odot$ to $\sim 5$ kpc and to drop by a factor of $\sim 2$ inside 5 kpc (Garwood and Dickey, 1989). Outside the solar circle, H I is probably mainly in the warm phase, as suggested by the fact that 21-cm emission profiles in the outer portions of external face-on spiral galaxies are almost perfectly Gaussian with a linewidth consistently in the range $6–9$ km s$^{-1}$ (Dickey, 1996).

High-resolution maps of the 21-cm emission sky strikingly show that the cold H I clouds are sheetlike or filamentary (Heiles, 1967; Verschuur, 1970; see also Fig. 5). Their true density can be estimated, for instance, by measuring the relative populations of the three fine-structure levels of the electronic ground state of interstellar neutral carbon (Jenkins et al., 1983). Typically, the hydrogen density in cold H I clouds is found to be $\approx 20–50$ cm$^{-3}$, i.e., some two orders of magnitude larger than in the warm intercloud H I (Kulkarni and Heiles, 1987). The fact that this density ratio is approximately the inverse of the temperature ratio (see Table I) supports the view that the cold and warm atomic phases of the ISM are in rough thermal pressure equilibrium.

The existence of two H I phases with comparable thermal pressures but with radically different temperatures and densities was predicted theoretically by Field et al. (1969; see also Goldsmith et al., 1969), who demon-

---

4The pure-absorption coefficient is independent of temperature, but the net absorption coefficient, corrected for stimulated emission, is inversely proportional to temperature.
strated that atomic interstellar gas heated by low-energy cosmic rays has two thermally stable phases: a cold dense phase, in which the primary cooling mechanism is the radiative deexcitation of collisionally excited fine-structure lines of metals, and a warm rarefied phase, resulting from the onset of $\text{H}\alpha$ cooling at about 8000 K.

Since the pioneering work of Field et al. (1969), other heating mechanisms have been put forward, such as photoelectric ejection off dust grains (Watson, 1972; Shull and Woods, 1985) and magnetohydrodynamic wave dissipation (Silk, 1975; Ferrière et al., 1988). The presence of these additional heating mechanisms does not alter the general conclusion that cold and warm interstellar HI may coexist in thermal pressure balance (see, for example, Wolfire et al., 1995); this is a direct consequence of the shape of the cooling curve, $\Lambda(T)$: fairly flat between a steep rise due to the [C II] 158-µm transition around 100 K and another steep rise due to $\text{H}\alpha$ around 8000 K.

Of course, the picture of a static ISM in strict equilibrium is excessively idealized. The observed atomic clouds have random motions characterized by a one-dimensional velocity dispersion $\approx 6.9 \text{ km s}^{-1}$ (Belfort and Crovisier, 1984). A sizable fraction of these clouds appear to be parts of expanding shells and supershells, with diameters ranging from a few tens of parsecs to $\sim 2$ kpc and with expansion velocities reaching a few tens of km s$^{-1}$ (Heiles, 1979, 1984). Whereas the most energetic supershells are thought to result from the impact of infalling high-velocity clouds onto the Galactic disk, the more modest HI shells are very likely created by stellar winds and supernova explosions, acting either individually or in groupings of up to a few thousands (Tenorio-Tagle and Bodenheimer, 1988). Many of the clouds that do not seem to belong to any expanding shell are probably fragments of old shells that have lost their identity (see Sec. V.A). Others may have directly condensed out of the warm neutral medium following a thermal instability induced by converging gas motions (Hennebelle and Pérault, 1999).

Let us note that the distinction between molecular and atomic clouds is not always clear-cut: some atomic clouds contain molecular cores (Dickey et al., 1981; Crovisier et al., 1984), while many molecular clouds possess an atomic halo (Wannier et al., 1983; Falgarone and Puget, 1985). Moreover, much of the molecular material and almost all of the atomic material in the ISM share an important property which ranks them amongst the “photodissociation regions,” namely, both are predominantly neutral and have their heating and chemistry

**FIG. 5.** High-resolution HI maps of (a) a $120^\circ \times 30^\circ$ portion of the sky centered on ($l = 80^\circ$, $b = -40^\circ$) at velocities between $-20$ and $20 \text{ km s}^{-1}$, from the Leiden-Dwingeloo Survey (Hartmann and Burton, 1997); figure courtesy of B. P. Wakker; (b) a $6^\circ \times 3^\circ$ portion of the sky centered on ($l = 260.5^\circ$, $b = 0^\circ$) at velocity 50 km s$^{-1}$, from the Southern Galactic Plane Survey; figure courtesy of J. M. Dickey.
largely regulated by stellar UV photons (Hollenbach and Tielens, 1999).

D. Warm ionized gas

O and B stars, the hottest and most massive stars in the Milky Way, emit a strong UV radiation, which, below a wavelength of 912 Å (corresponding to an energy of 13.6 eV), is sufficiently energetic to ionize hydrogen atoms. As a result, these stars are surrounded by a so-called “H II region” within which hydrogen is almost fully ionized. Given that the ionizing UV photons are promptly absorbed by neutral hydrogen, the transition between the H II region and the ambient ISM is rather abrupt. Inside the H II region, ions and free electrons keep recombining before being separated again by fresh UV photons from the central star. Thus the HII region keep recombining before being separated again by fresh UV photons from the central star. Therefore, the HII region grows until the rate of recombinations within it becomes large enough to balance the rate of photoionizations. In a uniform medium, this balance occurs when the radius of the H II region reaches the value of the Strömgren radius,

$$r_S = (30 \text{ pc}) \left( \frac{N_{\text{H}}}{n_H n_e} \right)^{1/3},$$

where \( N_{\text{H}} \) is the number of ionizing photons emitted per unit time by the central star, in \( 10^{48} \text{ s}^{-1} \) (e.g., \( N_{\text{H}} \approx 34 \) for an O5V star and \( N_{\text{H}} \approx 17 \) for a B0V star; Vacca et al., 1996), and \( n_H \) and \( n_e \) are the free-proton and free-electron number densities in the H II region, in \( \text{cm}^{-3} \) (Spitzer, 1978, p. 109).

The process of photoionization is accompanied by a net heating of the interstellar gas, as the ionizing photons transfer a fraction of their energy (the excess with respect to the ionization potential) to the ejected electrons. The equilibrium temperature, set by a balance between photoelectric heating and radiative cooling, has a typical value \( \approx 8000 \text{ K} \), depending on density and metallicity (Mallik, 1975; Osterbrock, 1989, p. 67). This theoretical estimate turns out to be in good agreement with observational determinations based on measurements of the radio continuum radiation (Osterbrock, 1989, p. 130) and on studies of emission-line ratios (Osterbrock, 1989, p. 123) from H II regions.

The radio continuum radiation of an H II region arises from the bremsstrahlung or “free-free” emission generated as free electrons are accelerated in the Coulomb field of positive ions (H+, He+, He2+). Emission lines, found at optical, infrared, and radio wavelengths, are primarily due to radiative recombination of hydrogen and helium ions with free electrons and to radiative de-excitation of collisionally excited ionized metals. Of special importance are the optical hydrogen Balmer lines produced by electronic transitions from an excited state \( n > 2 \) to the first excited state \( n = 2 \). Because each recombination of a free proton with a free electron into an excited hydrogen atom leads sooner or later to the emission of one Balmer photon, and because the rate per unit volume of recombinations into an excited hydrogen atom is proportional to \( n_{\text{H}^+} n_e n_e^2 \), the integrated intensity of the Balmer lines is directly proportional to the emission measure,

$$EM = \int n_e^2 \, ds,$$

where \( ds \) is the length element along the line of sight through the H II region. For future reference, let us specify that the hydrogen Balmer transition between the electronic energy levels \( n = 3 \) and \( n = 2 \) is usually referred to as the Hα transition and has a wavelength of 6563 Å.

The presence of warm ionized interstellar gas outside well-defined H II regions was first reported by Struve and Elvey (1938), who detected Hα and [O II] 3727 Å emission from extended zones in Cygnus and Cepheus. Systematic studies of this gas, through its optical emission lines, did not begin until over 30 years later, with, amongst others, an important Hα photographic survey by Sivan (1974) and more sensitive Hα spectroscopic scans by Roesler et al. (1978). What emerged from these observational studies is that diffuse Hα-emitting gas exterior to HII regions exists in all directions around us. More recent, high-resolution Hα maps of selected portions of the sky display a complex structure made of patches, filaments, and loops of enhanced Hα emission, superimposed on a fainter background (Reynolds, 1987; Reynolds et al., 1999a; see also Fig. 6). The new Wiscon-sin Hα Mapper (WHAM) survey (Reynolds et al., 1999a) will soon offer a complete and detailed view of the distribution and kinematics of the Hα-emitting gas over the entire sky north of \( -30^\circ \).

The temperature of the diffuse emitting gas, inferred from the width of the Hα and [S II] 6716 Å emission lines, is \( \approx 8000 \text{ K} \) (Reynolds, 1985a). This value has been confirmed by the recent WHAM observations of Hα, [S II] 6716 Å and [N II] 6583 Å, which, in addition, suggest a temperature rise at high \( |Z| \) (Haffner et al., 1999). At 8000 K, the observed Hα intensity along the Galactic equator translates into an effective emission measure in the range \( 9–23 \text{ cm}^{-6} \text{ pc} \) (Reynolds, 1983). Combined with a mean free path for absorption of Hα photons in the Galactic disk \( \approx 2 \text{ kpc} \) (Reynolds, 1985b), this range of emission measures implies a space-averaged electron density squared \( \langle n_e^2 \rangle \approx 4.5 \times 10^{-3} – 11.5 \times 10^{-3} \text{ cm}^{-6} \) in the diffuse ISM at low \( |Z| \). Across the Galactic disk, the emission measure is found to be \( \approx 4.5 \text{ cm}^{-6} \text{ pc} \) (Reynolds, 1984), which, together with the above electron density squared, yields an exponential scale height of \( \approx 390–1000 \text{ pc} \) for the diffuse Hα emission.

Owing to the obscuration effect of interstellar dust, the region that can be probed with Hα and other optical emission lines is limited to a cylindrical volume of radius \( \approx 2–3 \text{ kpc} \) around the Sun. A totally different source of information on the warm ionized interstellar gas, unaffected by obscuration, are the signals from pulsars (rapidly spinning, magnetized neutron stars, which emit regularly spaced pulses of electromagnetic radiation). It is well known that electromagnetic waves traveling through an ionized medium interact with the free elec-
trons in such a manner that their group velocity decreases with increasing wavelength. The periodic pulses emitted by pulsars can each be decomposed into a spectrum of electromagnetic waves spanning a whole range of radio wavelengths, with the longer-wavelength waves propagating less rapidly through interstellar space and, hence, arriving slightly later at the observer. The resulting spread in arrival times, a measurable quantity, is directly proportional to the column density of free electrons between the pulsar and the observer, i.e., to the dispersion measure

$$\text{DM} = \int_0^L n_e \, ds,$$  \hspace{1cm} (4)

with $L$ the distance to the pulsar.

Following the discovery of the first pulsar (Hewish et al., 1968), astronomers devised methods, mostly based on 21-cm absorption measurements, to estimate pulsar distances. This enabled them to model the large-scale distribution of interstellar free electrons by applying Eq. (4) to pulsars with independent dispersion measures and distance estimates. The best-fit models contain a thin-disk component arising from localized H II regions plus a thick-disk component associated with the diffuse warm ionized medium (Manchester and Taylor, 1981; Harding and Harding, 1982; Vivekanand and Narayan, 1982). Near the Sun, the space-averaged density of free electrons can be approximated by

$$\langle n_e \rangle(Z) = (0.015 \text{ cm}^{-3}) \exp\left(-\frac{|Z|}{70 \text{ pc}}\right) + (0.025 \text{ cm}^{-3}) \exp\left(-\frac{|Z|}{900 \text{ pc}}\right),$$  \hspace{1cm} (5)

(Reynolds, 1991), where the contribution from H II regions (first term) is taken from Manchester and Taylor (1981), while the diffuse component (second term) relies on the midplane density deduced from a limited sample of low-latitude pulsars by Weisberg et al. (1980) and on the column densities toward newly discovered pulsars inside high-$|Z|$ globular clusters. As Reynolds (1991) himself admitted, the exponential scale height of the extended component in Eq. (5) may have been underestimated by up to a factor of 2, due to a probable deficiency inside the Local Bubble in which the Sun is located.

Not only do pulsar signals experience dispersion upon propagating through ionized regions, but they also get

FIG. 6. High-resolution Hα map of a 90°×70° portion of the sky centered on ($l=115°$, $b=0°$) at velocities between −60 and 40 km s⁻¹, from the WHAM survey. Figure courtesy of L. M. Haffner. [Color]
scattered by fluctuations in the free-electron density. A useful quantity in this context is the scattering measure (SM) defined as the line-of-sight integral of the spectral coefficient for a power-law spectrum of electron-density fluctuations. Scattering measures can be related to a number of observables, such as the angular broadening of a small-diameter source and the temporal broadening of pulsar pulses, and their observational determination furnishes additional relevant information, particularly valuable in the direction of the Galactic center and toward pulsars without an independent distance estimate.

Cordes et al. (1991) analyzed two distinct data sets, comprising distances, dispersion measures, and scattering measures of pulsars and other radio sources, to construct an axisymmetric model of the free-electron space-averaged density outside well-defined H II regions. Their model consists of the superposition of a thin (H_e = 150 pc), annular component centered on R = 4 kpc and a thick (H_e = 1 kpc), radially extended component with Gaussian scale length \( \simeq 20 \text{ kpc} \). The thin component, which is presumably linked to the molecular ring discussed in Sec. III.B, gives a very small contribution at the solar circle, whereas the thick component corresponds to the second term in Eq. (5) and approximately reduces to it at \( R_\odot \).

Taylor and Cordes (1993) refined the model of Cordes et al. (1991) by utilizing more extensive data sets and by allowing for departures from axisymmetry. They explicitly incorporated a contribution from spiral arms, based on the spiral pattern inferred from existing optical and radio observations of H II regions, as well as a contribution from the nearby Gum Nebula, which systematically enhances the dispersion measure of pulsars located behind it. In their final model, the free-electron space-averaged density in interarm regions is somewhat less than in the axisymmetric model of Cordes et al. (1991); for instance, at \( R_\odot \), \( \langle n_e \rangle (Z = 0) = 0.019 \text{ cm}^{-3} \) instead of \( \approx 0.025 \text{ cm}^{-3} \). Spiral arms add a contribution equal to \((0.08 \text{ cm}^{-3}) \text{sech}^2(Z/300 \text{ pc})\) along their axis, so that, upon azimuthal average, both models are in good agreement.

If we suppose that helium remains largely neutral in the warm ionized medium, as suggested by the weak measured He I recombination line emission (Reynolds and Tuft, 1995; Heiles et al., 1996; Tuft, 1997), and if we disregard the fact that helium is fully ionized in the hot medium (see Sec. III.E), we may identify the ionized-hydrogen space-averaged density with the free-electron space-averaged density. Resorting to the axisymmetric model of Cordes et al. (1991), we then obtain the curves drawn in Figs. 1 and 2 for the column density of ionized interstellar hydrogen as a function of \( R \) and for its space-averaged density at \( R_\odot \) as a function of \( Z \), respectively.

Furthermore, if we assume a clear-cut separation between neutral and ionized media, with hydrogen completely neutral in the former and completely ionized in the latter, and if we ignore the small fraction of free electrons arising in the hot medium, we can derive an estimate for the true density \( n_e \) and the volume filling factor \( \phi \) of the diffuse warm ionized medium near the Sun. Indeed, from emission-measure data we know that \( \langle n_e^2 \rangle = \phi n_e^2 = 4.5 \times 10^{-3} - 1.15 \times 10^{-3} \text{ cm}^{-6} \). On the other hand, dispersion-measure data yield \( \langle n_e \rangle = \phi n_e \approx 0.025 \text{ cm}^{-3} \). From this it follows that \( n_e \approx 0.18 - 0.46 \text{ cm}^{-3} \) and \( \phi = 5 - 14\% \) (repeating the reasoning of Kulkarni and Heiles, 1987).

The parameters of the diffuse warm ionized medium obtained in this subsection are tabulated in Table I. A comparison with the parameters of the cold and warm neutral phases of the ISM suggests that the thermal pressure in the warm ionized medium (\( \approx 2.1nkT \)) is on average roughly twice as high as in the neutral medium (\( \approx 1.1nkT \)).

Reynolds (1984) compared the interstellar hydrogen recombination rate inferred from measurements of the Galactic Hα emission to the ionizing power of known sources of ionizing radiation in the solar neighborhood, and he concluded that only O stars are potentially able to do, by themselves, the desired job of maintaining the warm ionized medium in an almost fully ionized state (and, at the same time, at a temperature of \( \approx 8000 \text{ K} \)). There exist, however, two inherent problems with O stars’ being the primary source of ionization. First, O stars are preferentially born in dense molecular clouds close to the Galactic plane, which makes it difficult for a sufficient fraction of their ionizing photons to escape their immediate vicinity and pervade the general ISM up to the high altitudes where warm ionized gas is found (Reynolds, 1984). Second, the observed emission-line spectrum of the diffuse ionized background differs markedly from that characteristic of the compact H II regions surrounding O stars. In particular, the intensity ratio of the [S II] 6716-Å line to Hα is larger by about a factor of 4 (Reynolds, 1985a). Moreover, the observed spatial variations of the [S II] 6716 Å/Hα and [N II] 6583 Å/Hα intensity ratios are difficult to explain by pure photoionization (Reynolds et al., 1999b).

The first problem can be overcome by taking into account the multicomponent nature and the vertical structure of the ISM. Adopting an ISM model consisting of a thin layer of small, opaque clouds plus a more extended low-density extracloud medium, Miller and Cox (1993) calculated the shape and size of the HII regions associated with the known O stars near the Sun; they showed that the most powerful of them were able to grow out of the cloud layer, up to high \( |Z| \), and argued that their dilute portions do, in fact, constitute the diffuse warm ionized medium. Another calculation by Dove and Shull (1994) suggests that the large, elongated cavities blown by associations of O and B stars provide natural channels for their ionizing photons to reach high-altitude regions (see also Dove et al., 2000).

To explain the high [S II]/Hα intensity ratio, Sivan et al. (1986) had to appeal to a combination of photoionization and weak-shock excitation, whereas Mathis (1986) and, later, Domgörjen and Mathis (1994) found that a very dilute ionizing radiation field representative of a plausible mixture of O stars could reproduce the observations. The spatially variable [S II]/Hα and [N II]/

Rev. Mod. Phys., Vol. 73, No. 4, October 2001
Ha intensity ratios, for their part, appear to require supplemental ionization/heating sources (Reynolds et al., 1999b), such as photoelectric ejection off dust grains (Reynolds and Cox, 1992), dissipation of interstellar plasma turbulence (Minter and Spangler, 1997), Coulomb encounters with Galactic cosmic rays (Valinia and Marshall, 1998), or magnetic reconnection (Birk et al., 1998).

E. Hot ionized gas

The notion that hot interstellar gas exists in the Milky Way dates back to Spitzer’s (1956) paper on a possible Galactic corona, made of hot rarefied gas, which would provide the necessary pressure to confine the observed high-altitude interstellar clouds. The presence of such a hot gas was born out almost two decades later by two independent types of observations: (1) the Copernicus satellite detected, in the spectrum of several bright stars, broad UV absorption lines of high-stage ions that form only at elevated temperatures (Jenkins and Meloy, 1974; York, 1974), and (2) the observed soft-x-ray background radiation was found to be most likely due to thermal emission from a hot interstellar plasma (Williamson et al., 1974).

Amongst the high-stage ions accessible to UV observations, OVI [five-times-ionized oxygen, with a doublet at (1032 Å, 1038 Å)] and NV [four-times-ionized nitrogen, with a doublet at (1239 Å, 1243 Å)] are the best tracers of hot collisionally ionized gas, insofar as their high ionization potential makes them difficult to produce by photoionization. Their degree of ionization together with the measured linewidths imply a temperature of a few $10^5$ K (York, 1974, 1977). In addition, the integrated line intensities, which directly yield the column density of the considered ion between the Earth and the target stars, shed some light on its spatial distribution in the vicinity of the Sun.

The Copernicus OVI data were analyzed by Jenkins (1978a, 1978b), who, fitting the inferred OVI column densities with an exponential along Z, arrived at a local scale height of $\sim 300$ pc. Unfortunately, this value is extremely uncertain, as most of the observed stars lie close to the Galactic plane. Hurwitz and Bowyer (1996) studied more recent OVI data from a far-UV observation program of high-latitude stars and derived an exponential scale height of $\sim 600$ pc, which again should be taken with caution, given the wide fluctuations between sight lines. Finally, Savage et al. (2000) measured OVI column densities toward 11 active galactic nuclei with the FUSE satellite. When coupling their measurements with a straight estimate of the OVI midplane density from Copernicus, they obtained an OVI exponential scale height of $2.7 \pm 0.4$ kpc, whereas by adopting the estimate of Shelton and Cox (1994) of the OVI midplane density—which takes our position inside the Local Bubble into account—they obtained a $35\%$ higher scale height.

For NV, Sembach and Savage (1992) used multiple IUE absorption spectra of a few bright halo stars to evaluate NV column densities in several directions, and from their limited sample they deduced an exponential scale height of $\sim 1.6$ kpc. With a larger set of stellar- and extragalactic-source absorption spectra from both the IUE and the Hubble Space Telescope (HST), Savage et al. (1997) came up with an NV exponential scale height of $3.9 \pm 1.4$ kpc. Note that these authors studied other high-stage ions likely to be related to the hot gas; by way of reference, they derived exponential scale heights of $4.4 \pm 0.6$ kpc for CIV and $5.1 \pm 0.7$ kpc for SiIV.

The soft-x-ray background radiation around 0.25 keV appears to arise predominantly from the Local Bubble (Cox and Reynolds, 1987). The temperature of the emitting gas, deduced from the relative intensities of three adjacent energy bands, is $\approx 10^5$ K (McCaman and Sanders, 1990). Its average density can be inferred from the observed intensity of the soft-x-ray flux, provided that the emission path lengths are known. Snowden et al. (1990) estimated these path lengths by assuming that the x-ray-emitting region coincides with the local H I cavity and that the x-ray emissivity is uniform throughout the cavity. The cavity’s shape was then determined from the x-ray intensity distribution across the sky, and its scale was adjusted so as to match at best 21-cm emission measurements of H I column densities. The resulting electron density is $\approx 0.0037$–$0.0047$ cm$^{-3}$, which, if hydrogen and helium are fully ionized, corresponds to a hydrogen density $\approx 0.0031$–$0.0039$ cm$^{-3}$.

After the 1990 launch of an x-ray telescope on the Röntgen Satellite (ROSAT), it became increasingly clear that a significant fraction of the 0.25-keV x-ray flux observed in a number of directions originates outside the Local Bubble, the best evidence being the deep shadows cast in the soft-x-ray background by interstellar clouds in the Draco Nebula (Burrows and Mendenhall, 1991; Snowden et al., 1991), in the region of Ursa Major (Snowden et al., 1994), and in five other selected areas (Wang and Yu, 1995). The discovery of these shadows led Snowden et al. (1998) to reanalyze the 0.25-keV x-ray radiation, this time based on the high-resolution maps from the ROSAT all-sky survey, and model it with the superposition of an unabsorbed contribution from the Local Bubble, an absorbed contribution from the Galactic halo, and an absorbed isotropic contribution from extragalactic space. Unlike Snowden et al. (1990), they calibrated the bubble’s scale with the help of the estimated distance to a shadowing molecular cloud. In this manner, they found a Local Bubble similar in shape to that of the previous model but somewhat smaller, and they derived a hydrogen density $\approx 0.0065$ cm$^{-3}$.

If the values obtained for the temperature and the density of the hot gas in the Local Bubble are representative of the hot ionized phase of the ISM near the Sun, the thermal pressure in the hot ionized phase ($\approx 2.3 nkT$) exceeds that in the neutral phases ($\approx 1.1 nkT$) by a factor of $\sim 3$–15 (see Table I).

The 0.25-keV x-ray emission from the Galactic halo varies considerably over the sky. The northern halo exhibits strong intensity enhancements superimposed on a relatively uniform background, thereby suggesting that the emitting gas has a patchy distribution. In contrast,
the southern halo is characterized by intensity gradients toward low latitudes, roughly consistent with a plane-parallel distribution (Snowden \textit{et al.}, 1998).

It is very likely that 0.25-keV x-ray-emitting regions exist throughout the Milky Way, but because their radiation is efficiently absorbed by the intervening cool interstellar gas, the majority of them must escape detection. On the other hand, a number of bright features have been observed in the intermediate energy band 0.5–1.0 keV, which is less affected by photoelectric absorption. Most of these features were shown to be associated either with individual supernova remnants (produced by isolated supernova explosions) or with “superbubbles” (produced by the joint action of stellar winds and supernova explosions in a group of massive stars), and their x-ray radiation was attributed to thermal emission from a hot plasma at a temperature of a few $10^6$ K (Aschenbach, 1988; McCammon and Sanders, 1990).

Present-day observations are too limited to enable astronomers to map the large-scale distribution of hot interstellar gas in our Galaxy. Nevertheless, some qualitative conclusions can be drawn from the observational situation in external spiral galaxies. Their x-ray radiation is believed to arise from a combination of unresolved discrete sources and diffuse thermal emission from hot interstellar gas (Fabbiano, 1989). Despite the difficulty of separating both contributions and correcting for line-of-sight obscuration by cool interstellar gas, Cui \textit{et al.} (1996) measured the soft-x-ray intensity of several high-latitude face-on spirals—for which obscuration effects are minimal—and derived radial profiles for the emission measure of their hot component. The derived profiles unambiguously show that the amount of hot gas decreases radially outward. However, the large scatter among galaxies as well as the uncertainties involved in deducing column densities from emission measures and emission measures from measured intensities preclude any quantitative assessment.

It is now widely accepted that the hot interstellar gas is generated by supernova explosions and, to a lesser extent, by the generally powerful winds from the progenitor stars (see, for example, McCray and Snow, 1979; Spitzer, 1990). Supernova explosions drive rapidly propagating shock waves through the ISM, which sweep out cavities filled with hot rarefied gas and surrounded by a cold dense shell of collapsed interstellar matter. Cox and Smith (1974) pointed out that the hot gas inside the cavities had a sufficiently long radiative cooling time to be able to persist for millions of years; they further argued that, for a Galactic supernova frequency of one every 50 years, the hot cavities would overlap and form a network of interconnecting tunnels. Elaborating on this idea, McKee and Ostriker (1977) developed a self-consistent model of the local ISM, in which ~70% of interstellar space turns out to be filled with hot gas. However, Slavin and Cox (1993) demonstrated that the interstellar magnetic pressure substantially reduces the hot gas filling factor (see also McKee, 1990). Another important issue is the tendency for supernovae to be clustered and to create superbubbles rather than individual supernova remnants (McCray and Snow, 1979; Heiles, 1987). Adopting a first crude model to describe the shape and temporal evolution of superbubbles, Heiles (1990) estimated that they cover ~17% of the Galactic disk area near the Sun. With a more realistic model, including the latest observational data on ISM parameters and allowing for nearly equal numbers of isolated and clustered supernovae, Ferrière (1998b) obtained a local hot gas filling factor of ~20%, with an overwhelming contribution from superbubbles.

\section*{F. Dust}

The most visible manifestation of interstellar dust, already mentioned in Sec. II.B, is the obscuration and reddening of starlight as a result of absorption and scattering. The dust column density between a star and the Earth can be determined observationally from the “color excess” of the star, defined as the difference between its measured color index and the intrinsic color index given by its spectral type. On the whole, the dust column density appears fairly closely correlated, not with distance, but with hydrogen column density (Jenkins and Savage, 1974; Bohlin, 1975). This is a first indication that interstellar dust tends to follow the inhomogeneous, patchy distribution of the interstellar gas.

Extinction curves, which portray the wavelength dependence of interstellar extinction toward individual stars, provide crucial clues to the nature and size of the obscuring dust grains. These curves contain several spectral features, which may be regarded as signatures of specific radiative transitions. The fact that the relative strength of these features varies from star to star strongly suggests that they are produced by different types of grains (see, for example, Meyer and Savage, 1981; Witt \textit{et al.}, 1984). The most prominent feature, a bump at a UV wavelength $\sim$2175 Å, is traditionally attributed to graphite particles (Gilr, 1972; Mathis \textit{et al.}, 1977), although other carbon-containing compounds have also been put forward (see Mathis, 1987, for a review). The IR bands at 9.7 μm and 18 μm may be imputed to amorphous silicates (Knacke and Thomson, 1973; Draine and Lee, 1984). The carrier of a set of five mid-IR emission lines between 3.3 and 11.3 μm has been identified with polycyclic aromatic hydrocarbons (PAH’s; Duley and Williams, 1981; Léger and Puget, 1984). And other weaker features attest to the presence of additional species, including amorphous carbon and organic refractory material (Tielens and Allamandola, 1987).

The spectral shape of extinction curves may also be used to extract information on the size distribution of interstellar dust grains, the upper end of which may be further constrained by the cosmically available elemental abundances. Mathis \textit{et al.} (1977) were able to reproduce the standard extinction curve of the diffuse ISM in the wavelength interval 0.1–1 μm, with a mixture of spherical, uncoated graphite particles and adjustable amounts of other substances (mainly silicates), all distributed in size according to a power law of index...
≈−3.5, \( N(a) \, da \propto a^{−3.5} \, da \), over a radius range \( a =0.005−1 \, \mu m \) for the graphite particles and \( a =0.025−0.25 \, \mu m \) for the other substances. A similar size distribution, albeit with more structure between 0.02 \( \mu m \) and 0.2 \( \mu m \) and with a smooth falloff between 0.2 \( \mu m \) and 1 \( \mu m \), was obtained by Kim et al. (1994), who relied on extinction data extended to the wavelength interval 0.1−5 \( \mu m \) and employed a more objective statistical treatment. In addition, Kim et al. (1994) confirmed earlier claims that dense interstellar clouds have a flatter size distribution than the diffuse ISM, with a greater proportion of large grains.

The energy contained in the stellar photons absorbed by the interstellar dust grains heats them up to a temperature of over 15 K and is then mostly reemitted in the IR. Since the advent of IR astronomy in the early 1970s, it has been possible to observe the dust thermal emissions of interstellar dust. Our knowledge in the 1970s, it has been possible to observe the dust thermal emission and utilize it as an additional diagnostic tool to study the spatial distribution, composition, and physical properties of interstellar dust. Our knowledge in the field took two major steps forward with the experiments carried out by the Infrared Astronomy Satellite (IRAS) in 1983 and the Cosmic Background Explorer (COBE) in 1989–1990. Both satellites returned all-sky maps of the diffuse IR emission, the former with a finer spatial resolution and the latter with a broader spectral coverage and better absolute calibration.

Analysis of IRAS maps showed that away from localized heating sources there exists a good correlation both between the 100-\( \mu m \) IR intensity outside molecular clouds and the \( H I \) column density, and between the 100-\( \mu m \) IR intensity from nearby molecular clouds and their \( H_2 \) column density (Boulanger and Pé rault, 1988). Infrared emission measured by COBE at longer wavelengths is also well correlated with the \( H I \) gas, and the excess emission observed in some regions can be explained by a fraction of the interstellar hydrogen’s being in molecular or ionized form (Boulanger et al., 1996). The best discussion to date of dust-gas correlations in the ISM was presented by Schlegel et al. (1998), who produced a composite all-sky map of 100-\( \mu m \) dust emission combining the IRAS spatial resolution with the COBE quality calibration.

The interstellar dust emission spectrum covers a broad wavelength interval, from the near IR to the millimeter regime. With the help of preliminary COBE data, Wright et al. (1991) estimated for the first time the complete dust emission spectrum beyond 100 \( \mu m \). Boulanger et al. (1996) later found that its long-wavelength portion is well fitted by a Planck function with temperature ≈17.5 K times an emissivity \( \propto \lambda^{−2} \), although the emission peak near 150 \( \mu m \) and, more particularly, the excess emission below ∼100 \( \mu m \) are incompatible with a single-temperature fit.

Independently of this observational finding, Draine and Anderson (1985) had already shown that dust grains heated by ambient starlight exhibit a whole range of temperatures, the width of which increases with decreasing grain radius and becomes appreciable at less than 0.01 \( \mu m \). This spread in temperatures results from the discrete nature of the heating process: because the typical energy of stellar photons is not small compared to the thermal energy of dust grains, each photon absorption causes a finite jump in the grain temperature, which subsequently cascades back down through successive emissions of IR photons.

Allowing for a realistic spread in grain temperatures makes it possible to reproduce the observed dust emission spectrum. For example, Dwek et al. (1997) obtained a very good fit up to a wavelength ≈500 \( \mu m \), with a dust model consisting of spherical graphite and silicate grains and planar PAH molecules, exposed to the local interstellar radiation field. The graphite and silicate grains follow a slight variant of the power-law size distribution proposed by Mathis et al. (1977; see above), while the PAH molecules are distributed in radius according to a power law of index ≈−3 over a range \( a =4−10 \, \AA \) assuming that the radius \( a \) of a PAH, is related to its number of carbon atoms, \( N_c \), through \( a = (0.913 \, \AA) \sqrt{N_c} \) (Désert et al., 1990). In the model of Dwek et al. (1997), the far-IR emission beyond 140 \( \mu m \) arises predominantly from large graphite and silicate grains, whose temperature reaches ∼17–20 K and ∼15–18 K, respectively, whereas the excess emission below 100 \( \mu m \) is due to small particles, which are stochastically heated to temperatures ∼20 K (up to ∼500 K for the PAHs).

The characteristics of interstellar dust grains, inferred from their IR emission spectrum, vary among the different ISM phases. While the equilibrium dust temperature (temperature of the large grains) is relatively uniform at ∼17.5 K over most of the sky, large-scale far-IR maps reveal the presence of cold (∼15 K) dust, statistically associated with molecular clouds (Lagache et al., 1998). Low dust temperatures (as low as 12 K) have also been observed in several dense condensations within nearby interstellar clouds (Ristorcelli et al., 1998; Bernard et al., 1999). In general, low dust temperatures appear to be on a par with low abundances of small dust particles (Lagache et al., 1998). The dust emission associated with the warm ionized medium was first isolated by Lagache et al. (1999), whose preliminary analysis yielded a best-fit dust temperature ∼29 K for this medium. Since then, the new results from the WHAM survey have permitted a more trustworthy separation of the contribution from the warm ionized medium to the observed dust emission, which led Lagache et al. (2000) to conclude that, in fact, the dust temperature and abundances in the warm ionized medium do not differ significantly from those in \( H I \) regions.

To make the link with the interstellar depletion factors discussed in Sec. III.A, let us note that in the dust emission model of Dwek et al. (1997), silicate grains take up virtually all the cosmically available Mg, Si, and Fe, plus ∼15% of the available O, while graphite grains take up ∼62% of the available C, and PAH’s take up another ∼18%. For comparison, the extinction model of Kim et al. (1994), which does not include PAH’s, requires ∼95% of the available Si to be locked up in silicate grains and ∼75% of the available C in graphite grains. Hence direct studies of interstellar dust, based on its IR
emission, on the one hand, and on its extinction properties, on the other hand, lead to similar conclusions. These conclusions, however, are only partially supported by interstellar depletion data, according to which the heavy refractory elements are indeed highly depleted, but at least one-third of the carbon remains in the gaseous phase (see Sec. III.A).

The origin of interstellar dust is not yet completely understood. What seems firmly established is that a fraction of the dust grains form in the cool outer atmosphere of red-giant and supergiant stars and in planetary nebulae, where the temperature and pressure conditions are conducive to condensation of carbonaceous species (graphite and amorphous carbon) and silicates; the newly formed particles are then expelled into interstellar space by the stellar radiation pressure (Woolf and Ney, 1969; Salpeter, 1976; Draine, 1990). However, two independent lines of evidence converge to suggest that dust grains are also produced in the ISM itself: On the theoretical side, the high-velocity shock waves driven by supernova explosions were estimated to destroy dust grains at a much faster rate than the injection rate from stars, so that other sources are necessary to maintain the existing grain population (Dwek and Scalo, 1980; Seab and Shull, 1983; Jones et al., 1994). On the observational side, the measured interstellar elemental depletions tend to be higher in dense clouds (see Sec. III.A), which could be explained by these cold entities being the sites of additional grain formation (Seab, 1987).

Although dust represents but a small fraction of the interstellar mass, it plays a key role in the chemical and energetic balances of the ISM. A first way in which dust participates in the ISM chemistry is by providing direct sinks and sources for the interstellar gas: in cold dense clouds dust grains accrete particles from the gas phase, whereas in the warmer intercloud medium they shed their volatile mantle, which then returns to the gas phase; in addition, their refractory core is partially vaporized by each passing shock wave (Seab, 1987). Another important chemical process to which dust grains contribute significantly is the formation and maintenance of molecular hydrogen. Not only do they serve as catalysts by allowing hydrogen atoms to recombine on their surface (Hollenbach and Salpeter, 1971; Duley and Williams, 1993; Katz et al., 1999), but they also help to shield the resulting H₂ molecules from photodissociation by the ambient UV radiation field (Shull and Beckwith, 1982). Regarding the energetic balance of the ISM, dust makes a significant contribution to both heating and cooling, through ejection of energetic photoelectrons (Watson, 1972; Bakes and Tielens, 1994) and through collisions with gas particles (Burke and Hollenbach, 1983), respectively.

A. Magnetic fields

The presence of interstellar magnetic fields in our Galaxy was first revealed by the observational discovery of linear polarization of starlight (Hall, 1949; Hiltner, 1949a, 1949b). Davis and Greenstein (1951) explained this polarization in terms of selective extinction by elongated dust grains that are at least partially aligned by an interstellar magnetic field. They showed that rapidly spinning paramagnetic grains tend to make their spin axis coincide with their short axis and orient the latter along the magnetic field. Since dust grains preferentially block the component of light with polarization vector E parallel to their long axis, the light that passes through is linearly polarized in the direction of the magnetic field. Applying their theory to the observed polarization of starlight, Davis and Greenstein (1951) concluded that the interstellar magnetic field is locally parallel to the Galactic plane.

The first large-scale polarization database was put together by Mathewson and Ford (1970), for a total of almost 7000 stars distributed over both celestial hemispheres and located within a few kpc from the Sun. Their compilation provides a comprehensive map of the magnetic-field direction on the plane of the sky, which not only substantiates Davis and Greenstein’s (1951) conclusions but also indicates that the local magnetic field is nearly azimuthal, pointing toward a Galactic longitude ≈80°, i.e., having an inclination angle ≈10°. A recent, thorough reanalysis of Mathewson and Ford’s (1970) polarization data gives for the local magnetic field an inclination angle =7.2° and a radius of curvature =8.8 kpc (Heiles, 1996); this inclination angle of magnetic-field lines is somewhat less than the standard pitch angle of the Galactic spiral pattern inferred from optical data (=12°; Georgelin and Georgelin, 1976) and from radio data (=13°; Beuermann et al., 1985). Although considered the best reference for over a quarter of a century, Mathewson and Ford’s (1970) catalog has now been superseded by the more complete compilation produced by Heiles (2000).

Stellar polarimetry acquaints us solely with the direction of the interstellar magnetic field. To determine its strength, one has to resort to one out of three methods, based on the Zeeman splitting of the H I 21-cm line or other radio lines, the Faraday rotation of linearly polarized radio signals, or the radio synchrotron emission from relativistic electrons.

The Zeeman splitting of a given atomic or molecular line occurs in the presence of an external magnetic field, whose interaction with the magnetic moment of the valence electrons causes a subdivision of certain electronic energy levels. The amplitude of the Zeeman splitting, Δν, is directly proportional to the magnetic-field strength B so that, in principle, it suffices to measure Δν in order to obtain B in the region of interest. Unfortunately, in the case of the widely used H I 21-cm line, Δν

5Another possibility would be that low-density regions have statistically undergone more grain destruction by fast shock waves (Shull, 1986).
is usually so small compared to the linewidth that it cannot be measured in practice. The way to circumvent this problem is to observe instead the difference between the two circularly polarized components of the 21-cm radiation, which directly yields the line-of-sight component of the magnetic field, $B_\parallel$.

The first successful implementation of this technique was by Verschuur (1969), who reported magnetic-field strengths of a few $\mu$G for several nearby HI clouds. Since Verschuur’s early detections, a vast body of Zeeman-splitting observations has built up, both for the HI 21-cm line and for several centimeter-wavelength lines of the OH molecule. From a compilation of these observations by Troland and Heiles (1986), it emerges that interstellar magnetic fields have typical strengths of a few $\mu$G in regions with gas density $n \approx 1 - 100$ cm$^{-3}$, and that they display only a slight tendency for $B$ to increase with increasing $n$; this tendency is a little more pronounced in the higher density range $n \approx 10^2 - 10^4$ cm$^{-3}$, where field strengths may reach up to a few tens of $\mu$G (see also Myers et al., 1995; Crutcher, 1999). For comparison, the dipole magnetic field of the Earth has a strength of 0.31 G at the equator, the solar magnetic field has a strength of $\sim 1$ G in quiet regions and $\sim 10^3$ G in sunspots, and pulsars have typical field strengths of $\sim 10^8 - 10^{15}$ G (see, for example, Asseo and Sol, 1987).

Zeeman-splitting measurements are plainly biased toward interstellar regions with high HI column densities and narrow 21-cm linewidths, i.e., toward cold neutral clouds. In contrast, Faraday-rotation measurements sample ionized regions. As a reminder of basic plasma physics, a linearly polarized electromagnetic wave propagating along the magnetic field of an ionized medium can be decomposed into two circularly polarized modes: a right-hand mode, whose $\mathbf{E}$ vector rotates about the magnetic field in the same sense as the free electrons gyrate around it, and a left-hand mode, whose $\mathbf{E}$ vector rotates in the opposite direction. As a result of the interaction between the $\mathbf{E}$ vector and the free electrons, the right-hand mode travels faster than the left-hand mode; consequently the plane of linear polarization experiences a rotation—known as Faraday rotation—as the wave propagates. The angle by which the polarization plane rotates is equal to the wavelength squared times the rotation measure,

$$\text{RM} = C \int_{0}^{L} n_e B_\parallel ds,$$

where the numerical constant is given by $C = 0.81$ rad m$^{-2}$ when the free-electron density $n_e$ is expressed in cm$^{-3}$, the line-of-sight component of the magnetic field, $B_\parallel$, is in $\mu$G, and the path length $L$ is in pc (Gardner and Whiteoak, 1966). Observationally, the rotation measure of a given source can be determined by measuring the polarization position angle of the incoming radiation at two or more wavelengths.

The sources of linearly polarized waves used to carry out Faraday-rotation measurements in our Galaxy are either pulsars or extragalactic radio continuum sources. Pulsars present a double advantage in this context: first, they lie within the Galaxy and have estimable distances, and second, their rotation measure [Eq. (6)] divided by their dispersion measure [Eq. (4)] directly yields the $n_e$-weighted average value of $B_\parallel$ along their line of sight.

Rand and Kulkarni (1989) analyzed 116 pulsars closer than 3 kpc to the Sun and concluded that the local interstellar magnetic field has a uniform (or regular) component $\approx 1.6 \mu$G and a random (or irregular) component of $\sim 5 \mu$G. By including pulsars deeper into the inner Galaxy and making a more careful selection in the pulsar sample, Rand and Lyne (1994) arrived at a local uniform field strength $\approx 1.4 \mu$G. Moreover, they found that the uniform field strength increases smoothly toward the Galactic center, reaching at least 4.2 $\mu$G at $R = 4$ kpc (see their Fig. 6), which implies an exponential scale length $\approx 4.1$ kpc. They also showed that a concentric-ring model, in which the uniform magnetic field is purely azimuthal, leads to field reversals at 0.4 kpc and at 3 kpc inward from the solar circle, with a maximum field strength between field reversals of 2.1 $\mu$G.

Han and Qiao (1994) obtained the same value, $\approx 1.4 \mu$G, for the local uniform field strength, but they argued that a bisymmetric magnetic-field structure with a pitch angle of $8.2^\circ$ and an amplitude of 1.8 $\mu$G gives a better fit to the pulsar data than the concentric-ring model. Han et al. (1999) found evidence for at least one field reversal in the outer Galaxy and possibly a third reversal inside the solar circle. In line with most astrophysicists working on this problem, they claimed that the observed field reversals together with the measured pitch angle lend credence to the bisymmetric field picture. Yet it can be shown that field reversals are also consistent with an axisymmetric magnetic configuration, and there are reasons, both observational (Vallée, 1996) and theoretical (Ferri`ere and Schmitt, 2000), to favor the axisymmetric field picture.

Vertically, interstellar magnetic fields exist well beyond the pulsar zone, as indicated by the fact that the rotation measures of extragalactic radio sources have the same sign and are systematically larger in absolute value than the rotation measures of pulsars in a nearby direction (Simard-Normandin and Kronberg, 1980). In principle, the scale height of the uniform field can be gathered from the observed rotation measures of high-latitude extragalactic sources supplemented by a model of the free-electron density. In practice, though, the inferred scale height turns out to be very sensitive to the parameters of the free-electron distribution, which, regrettably, are poorly constrained. For reference, from a sample of more than 600 extragalactic sources, Inoue and Tabara (1981) obtained a magnetic scale height of $\sim 1.4$ kpc. This value is about one order of magnitude greater than the scale height deduced from pulsar rotation measures (see Thomson and Nelson, 1980; Han and Qiao, 1993). Despite the substantial uncertainties in both estimates, the huge discrepancy could be indicative of the existence of two magnetic layers with very different scale heights (Han and Qiao, 1994), a view supported by synchrotron-emission data (see Sec. IV.C).
Regarding the magnetic-field parity with respect to the Galactic midplane, present-day observations do not convey a very clear picture. The rotation-measure vertical distribution for extragalactic radio sources and for pulsars with Galactic latitudes $|b| > 8^\circ$ appears to be approximately antisymmetric about the midplane in the inner (first and fourth) quadrants and roughly symmetric in the outer (second and third) quadrants (Oren and Wolfe, 1995, Fig. 3; Han et al., 1997, Figs. 1 and 2). Although the observed antisymmetry in the inner Galaxy has often been attributed to nearby anomalies like the North Polar Spur, Han et al. (1997) argue for a genuine property of the uniform magnetic field away from the midplane. Meanwhile, rotation measures of low-latitude pulsars point to a symmetric distribution near the midplane, at all longitudes (Rand and Lyne, 1994, Fig. 4).

To conclude this subsection, let us reemphasize the two important limitations of Faraday-rotation studies (see Heiles, 1995). First, rotation measures essentially probe the warm ionized medium, which presumably contains most of the interstellar free electrons but occupies only a modest fraction of the interstellar volume (see Sec. III.D), so that the inferred magnetic-field strengths are not necessarily representative of the ISM in general. Second, rotation measures furnish the line-of-sight component of the Galactic magnetic field, not its total strength. A more adequate method of determining interstellar magnetic-field strengths draws on the Galactic radio synchrotron emission. Since the synchrotron emissivity also depends on the density and spectrum of relativistic electrons, we shall return to this method in Sec. IV.C, after discussing in detail the observed properties of interstellar cosmic rays.

B. Cosmic rays

The Earth is continually bombarded by highly energetic, electrically charged particles from space. Since their extraterrestrial origin was established by a balloon experiment (Hess, 1919), they have been referred to as cosmic rays, even though it was later realized that they are in fact material particles rather than photons (Bothe and Kohlhörster, 1929). Measurements from instrumented balloons and satellites have shown that cosmic rays comprise protons, $\sim 10\%$ helium nuclei, $\sim 1\%$ heavier nuclei, $\sim 2\%$ electrons, and smaller amounts of positrons and antiprotons (Blandford and Eichler, 1987; Bloemen, 1987). They have typical velocities close to the speed of light and span a wide range of kinetic energies $E$. While the majority of cosmic rays with $E \lesssim 0.1$ GeV/nucleon originate in the Sun, the more energetic ones emanate mainly from the ISM.

The observed cosmic-ray energy spectrum is strongly modulated by the irregular, fluctuating magnetic field of the solar wind (Gleeson and Axford, 1968). Because solar modulation effects are difficult to disentangle, the actual shape of the interstellar cosmic-ray energy spectrum is uncertain below $\sim 1$ GeV/nucleon. At higher energies, its nuclear component can be approximated by a piecewise continuous power law, $f(E) \propto E^{-\gamma}$, with $\gamma = 2.5$ for $E = 2-10$ GeV/nucleon and $\gamma = 2.7$ for $E = 10 - 10^5$ GeV/nucleon (Simpson, 1983a; Webber, 1983a). The differential spectral index increases to $\gamma = 3.1$ at $E = 3 \times 10^5$ GeV/nucleon and decreases again above $10^8$ GeV/nucleon (Hillas, 1984). The electron spectrum runs parallel to the proton spectrum between 2 GeV and 10 GeV and steepens to a differential index $\sim -3.3$ at a few tens of GeV (Webber, 1983b).

The energy density of interstellar cosmic rays measured at Earth varies over the course of the 11-year magnetic cycle of the Sun, from $0.98$ eV cm$^{-3}$ at sunspot minimum to $0.78$ eV cm$^{-3}$ at sunspot maximum. Webber (1987) extrapolated these values to the boundary of the heliosphere (a cavity in the ISM carved out by the solar wind) by using the solar modulation model of Gleeson and Axford (1968) together with an estimate for their modulation parameter based on the heliospheric cosmic-ray intensity gradient measured by the Voyager and Pioneer spacecraft out to $\sim 30$ AU from the Sun. Proceeding in this manner, he obtained an interstellar cosmic-ray energy density $\sim 1.5$ eV cm$^{-3}$ just outside the heliosphere. Eleven years later, with Voyager and Pioneer cosmic-ray data available out to $\sim 60$ AU and with more sophisticated solar modulation models (Potgieter, 1995) at his disposal, Webber (1998) updated the value of the interstellar cosmic-ray energy density to $\sim 1.8$ eV cm$^{-3}$. At the time of this writing (October, 2000), Voyager 1 is at nearly 80 AU from the Sun, probably quite close to the heliospheric termination shock, and Webber (2000), who has periodically reexamined the situation, maintains the value of $1.8$ eV cm$^{-3}$.

If all cosmic rays were ultrarelativistic, the cosmic-ray pressure would simply be one-third of their energy density, i.e., $\sim 9.6 \times 10^{-15}$ dyn cm$^{-2}$. However, as reemphasized by Boulares and Cox (1990), the bulk of the cosmic-ray energy density is due to mildly relativistic protons with kinetic energies of a few GeV, so the cosmic-ray pressure must lie somewhere between one-third and two-thirds of their energy density. The correct pressure integral reads

$$P_{CR} = \frac{1}{3} \int \frac{E + 2E_0}{E + E_0} E n(E) dE,$$

where $n(E)$ is the cosmic-ray differential number density and $E_0$ is the rest energy per nucleon (Ip and Axford, 1985). Applying this expression to Webber’s (1998) “demodulated” cosmic-ray spectrum gives for the interstellar cosmic-ray pressure just outside the heliosphere ($P_{CR}) \sim 12.8 \times 10^{-15}$ dyn cm$^{-2}$.

We know from absorption-line studies toward the nearest (closer than a few parsecs) stars that the solar system is not directly surrounded by the hot tenuous gas of the Local Bubble, but that it is first immersed in a warm interstellar cloud, the Local Cloud, with a tem-

---

6The astronomical unit (AU) is the length unit equal to the distance between the Earth and the Sun. Its numerical value is given by $1$ AU = $1.50 \times 10^8$ km = $4.85 \times 10^{-6}$ pc.
perature ~6700−7600 K, a H\textsc{i} density \approx 0.18−0.28 \text{ cm}^{-3}, and an electron density \geq 0.10 \text{ cm}^{-3} (Frisch, 1995; Lallement et al., 1996; Redfield and Linzky, 2000). It is therefore reasonable to surmise that the aforementioned value of the cosmic-ray pressure is representative of the warm phase of the ISM near the Sun. The space-averaged value of $P_{\text{CR}}$ near the Sun is probably somewhat less, for magnetic and cosmic-ray pressures are both expected to be lower in the hot interstellar phase than in the rest of the ISM. Indeed, supernova explosions, believed to constitute the main source of hot interstellar gas, expel magnetic-field lines and cosmic rays out of the hot cavities they create, and it is only during their contraction phase that these hot cavities are gradually replenished with field lines and cosmic rays from the surrounding ISM. Ferriere (1998a) estimated that the space-averaged magnetic and cosmic-ray pressures near the Sun are only a few percent lower than their counterparts in the warm Local Cloud. In view of the many uncertainties involved in the above derivation, we shall ignore this small difference and ascribe the value $(P_{\text{CR}})_{\odot} \approx 1.2 \times 10^{-13} \text{ dyn cm}^{-2}$ to the space-averaged vicinity of the Sun.

The interaction of cosmic rays with interstellar matter and photons gives rise to $\gamma$-ray radiation through various mechanisms, including (1) the production of $\pi^0$ mesons, which rapidly decay into two $\gamma$ photons, (2) the Coulomb acceleration of cosmic-ray electrons by the nuclei and electrons of the interstellar gas, which leads to $\gamma$-ray bremsstrahlung emission, and (3) the scattering of high-energy cosmic-ray electrons on ambient soft photons, which results in “inverse Compton emission,” again at $\gamma$-ray wavelengths (Bloemen, 1989). Owing to technical difficulties, $\gamma$-ray astronomy had a slow start, with several isolated balloon and satellite experiments, but since the successful launch of the second Small Astronomy Satellite (SAS-2) in 1972 (see Fichtel et al., 1975) and the Cosmic-Ray Satellite (COS-B) in 1975 (see Bloemen, 1989), rapid progress has been made in observing the diffuse $\gamma$-ray background as a means of tracing the Galactic distribution of cosmic rays. With a broader energy range ($\approx 50$ MeV−6 GeV) than SAS-2 and with far better counting statistics due to its longer lifetime, the COS-B mission has been especially useful in this regard.

Correlation studies of COS-B $\gamma$-ray intensity maps with H\textsc{i} and CO maps show that the $\gamma$-ray emissivity per H atom decreases away from the Galactic center, this decrease being more pronounced at smaller $\gamma$-ray energies (see, for example, Bloemen et al., 1986; Strong et al., 1988). The energy-dependence of the emissivity gradient was at first interpreted as evidence that cosmic-ray electrons have a steeper radial gradient than cosmic-ray nuclei (Bloemen et al., 1986). However, followup work (Strong et al., 1988) favored an alternative explanation, according to which molecular clouds, which are more numerous at small $R$, contain proportionally more cosmic-ray electrons than the diffuse ISM. In the second scenario, cosmic-ray nuclei and electrons outside molecular clouds have the same radial dependence, with an exponential scale length $\approx 13$ kpc beyond 3.5 kpc (after rescaling to $R_{\odot} = 8.5$ kpc).

The available COS-B $\gamma$-ray maps convey little precise information on the vertical distribution of interstellar cosmic rays, first because the interstellar gas, which governs the $\pi^0$ decay and bremsstrahlung components of the $\gamma$-ray emission, tends to be confined close to the Galactic plane, and second because the observed $\gamma$-ray radiation at Galactic latitudes $|b| \geq 10^\circ$ is probably contaminated by extragalactic sources. The lower energy band 1−30 MeV surveyed from the Compton Gamma Ray Observatory (CGRO) is comparatively more sensitive to the inverse Compton component of the $\gamma$-ray emission. Attempts to isolate this component suggest that it has significantly wider latitudinal extent than the interstellar gas (Strong et al., 1996), which, if proven correct, implies that at least a fraction of the cosmic-ray electrons are distributed in a thick disk.

It is possible to derive a more quantitative estimate of the cosmic-ray vertical scale height on a totally different basis. In order to describe the method employed, we first need to discuss the sources of cosmic rays, their propagation through the ISM, and their escape from the Galaxy.

The measured elemental composition of Galactic cosmic rays points to two different kinds of cosmic-ray sources, both related to stars. On the one hand, the similarity with the elemental composition of solar energetic particles, notably concerning an anticorrelation between abundance and first-ionization potential for elements heavier than helium, suggests that Galactic cosmic rays originate in unevolved late-type stars and are injected into the surrounding ISM via flares out of their corona (Meyer, 1985). On the other hand, the relative overabundance of iron (Simpson, 1983b) argues for the formation of cosmic rays in very evolved early-type stars, with a release into the ISM upon the terminal supernova explosion.

Regardless of the exact injection sites, the injected cosmic rays—the so-called primaries—are probably further accelerated upon traveling through the ISM and repeatedly scattering off moving irregularities in the interstellar magnetic field. This acceleration process may occur either stochastically in the turbulent ISM (second-order Fermi acceleration; Fermi, 1949, 1954; Jokipii, 1978) or systematically at supernova shock waves, where the converging upstream and downstream flows scatter cosmic rays back and forth across the shock front (first-order Fermi acceleration; Axford et al., 1978; Bell, 1978; Blandford and Ostriker, 1978).

In spite of their relativistic velocities, cosmic rays cannot freely travel through interstellar space: they are trapped by the interstellar magnetic field, which constrains their motion both perpendicular and parallel to its direction. In the perpendicular direction, cosmic rays are forced to gyrate about magnetic-field lines along a circular orbit of radius
where \( q \) is the cosmic-ray electric charge, \( Z \) is the atomic number (taken as \(-1\) for electrons), and \( \theta = \arcsin(v/v_c) \) is the cosmic-ray pitch angle, defined as the angle between its velocity and the magnetic field. In the parallel direction, cosmic rays excite resonant Alfvén waves, which in turn scatter them and limit their streaming motion to a slow diffusion at a velocity barely larger than the Alfvén speed (Kulsrud and Pearce, 1969; Wentzel, 1969). The literature abounds with theoretical models of cosmic-ray propagation and confinement in our Galaxy. We shall not discuss these models here, but we refer the interested reader to the comprehensive review paper by Cesarsky (1980).

What about the final fate of Galactic cosmic rays? A first possibility is that they end up losing all their energy in inelastic collisions with interstellar gas particles (Rasmussen and Peters, 1975). A more attractive alternative is that they eventually escape from the Galaxy, either by streaming along magnetic-field lines that connect with a weak extragalactic magnetic field (see Piddington, 1972), or by diffusing across field lines to the edge of the Galaxy, where they have a finite probability of “leaking out” following magnetic reconnection [for example, in a magnetic bubble that detaches itself from the Galactic magnetic field (Jokipii and Parker, 1969)], or else by being convected away, together with the interstellar gas and field lines, in a Galactic wind (Jokipii, 1976).

Let us now return to the question of how Galactic cosmic rays are distributed along the vertical. When primary cosmic-ray nuclei—mainly C, N, and O—collide with interstellar hydrogen, they can break up into lighter secondary nuclei—such as Li, Be, and B. The measured abundance of stable secondaries indicates that the mean column density of interstellar matter traversed by cosmic rays is energy dependent and reaches a maximum value \( 9 \text{ g cm}^{-2} \) at a cosmic-ray energy \( \approx 1 \text{ GeV/nucleon} \) (García-Munoz et al., 1987). Furthermore, the mean cosmic-ray lifetime can be deduced from the measured abundance of unstable secondaries—typically \(^{10}\text{Be}\)—interpreted in the framework of a cosmic-ray propagation model. If the volume occupied by cosmic rays is modeled as a homogeneous “leaky box,” the mean cosmic-ray lifetime works out to be \( \approx 15 \text{ Myr} \) at 0.38 GeV/nucleon, and the average ISM density in the cosmic-ray box amounts to \( 0.24 \text{ cm}^{-3} \) (Simpson and García-Munoz, 1988); since this is about one-fifth of the local ISM density, the cosmic-ray box must be about five times thicker than the disk of interstellar gas. In halo diffusion models, the mean cosmic-ray lifetime is longer and the effective thickness of the cosmic-ray confinement region is larger than in the homogeneous leaky-box model. For instance, Bloemen et al. (1993) derived an effective cosmic-ray density \( \approx 3 \text{ kpc} \) at 1 GeV/nucleon. Likewise, Webber et al. (1992) found that the thickness of the cosmic-ray halo at 1 GeV/nucleon must lie between 2 kpc and 3–4 kpc in the absence of Galactic wind convection and that its maximum allowed value decreases with increasing wind velocity.

### C. Synchrotron radiation

The rapid spiraling motion of cosmic-ray electrons about magnetic-field lines generates nonthermal radiation, termed synchrotron radiation, over a broad range of radio frequencies. The synchrotron emissivity at frequency \( \nu \) due to a power-law energy spectrum of relativistic electrons, \( f(E) = K \nu^{-\gamma} \), is given by

\[
\varepsilon_{\nu} = \mathcal{F}(\gamma) K_{\nu} B_{\nu}^{(1+\gamma)/2} \nu^{-(\gamma-1)/2},
\]

where \( \mathcal{F}(\gamma) \) is a known function of the electron spectral index and, as before, \( B \) is the magnetic-field strength (Ginzburg and Syrovatskii, 1965). For the following, let us note that electrons with energy \( E \) emit the most power at frequency

\[
\nu = (16 \text{ MHz}) B(\mu \text{G}) E_{\text{GeV}}.
\]

(Beuermann et al. 1985) modeled the Galactic synchrotron emissivity, based on the all-sky radio continuum map at 408 MHz of Haslam et al. (1981a, 1981b). Their model is restricted to the radial interval 3.5–17 kpc (after rescaling to \( R_\odot = 8.5 \text{ kpc} \)); it possesses a spiral structure similar to that observed at optical wavelengths, and it contains a thin disk of equivalent half-thickness \( H_\nu(R) = (157 \text{ pc}) \exp[(R-R_\odot)/R_\odot] \) plus a thick disk of equivalent half-thickness \( H_\nu(R) = (1530 \text{ pc}) \exp[(R-R_\odot)/R_\odot] \). The modeled synchrotron emissivity, averaged over the Galactic azimuth, can be written as

\[
\varepsilon_{\nu}(R,Z) = (8.2 \text{ K kpc}^{-1}) \times \left[ 0.46 \exp \left( -\frac{R-R_\odot}{2.8 \text{ kpc}} \right) \left( \text{sech} \left( \frac{Z}{255 \text{ pc}} \right)^{n(R)} \right) + 0.54 \exp \left( -\frac{R-R_\odot}{3.3 \text{ kpc}} \right) \left( \text{sech} \left( \frac{Z}{255 \text{ pc}} \right)^{b(R)} \right) \right],
\]

where the exponents \( n(R) \) and \( b(R) \) take on the values imposed by the above half-thicknesses; in particular, at the solar circle, \( n(R_\odot) = 4.60 \) and \( b(R_\odot) = 0.187 \).

Since the synchrotron emissivity [Eq. (7)] depends on both the magnetic-field strength and the spectrum of cosmic-ray electrons, it is necessary to know one of these two quantities in order to deduce the other one from Eq. (9). Unfortunately, as explained in Secs. IV.A and IV.B, neither quantity has been reliably determined so far. One is thus constrained to take another approach and appeal to a double assumption often made in this context, namely, the assumption that the density of cosmic-ray electrons is proportional to cosmic-ray pressure and that cosmic-ray pressure is proportional to magnetic pressure. The second part of this double assumption may be justified by minimum-energy-type arguments (see, for example, Beck et al., 1996). The first part is certainly not strictly verified throughout the electron spectrum, insofar as cosmic-ray electrons suffer...
more severe radiation losses (bremsstrahlung, inverse Compton radiation, synchrotron radiation) than cosmic-ray nuclei. However, it is probably reasonable in the 2–10-GeV energy range in which cosmic-ray electrons generate most of the synchrotron emission at 408 MHz (see below) and cosmic-ray protons contribute most of the cosmic-ray pressure (see Sec. IV.B), since in that energy range the electron spectrum parallels the proton spectrum. Under these conditions, Eq. (7) implies that magnetic pressure, \( P_M = B^2/8\pi \), and cosmic-ray pressure, \( P_{CR} \), are both proportional to \( \varepsilon_e^{4/(\gamma + 5)} \).

The last needed piece of information is the value of \( P_M \) and \( P_{CR} \) at a given point in the Galaxy, say, at the Sun. The cosmic-ray pressure at the Sun was estimated in Sec. IV.B to be \( (P_{CR})_0 = 12.8 \times 10^{-13} \text{ dyn cm}^{-2} \). As far as magnetic pressure is concerned, there are estimates from Zeeman-splitting and Faraday-rotation measurements, but these estimates refer only to the line-of-sight component of the magnetic field and they are biased toward cold neutral and warm ionized regions, respectively. However, the local ISM constitutes a unique place in the Galaxy for which it is possible to determine the cosmic-ray electron spectrum independently of other physical quantities and, consequently, to deduce the magnetic-field strength from the value of the synchrotron emissivity, \( (\varepsilon_e)_0 = 8.2 \text{ K kpc}^{-1} \) [see Eq. (9)].

The cosmic-ray electron spectrum in the Local Cloud harboring the solar system can be directly measured at Earth down to an energy \( E \sim 10 \text{ GeV} \). Below this energy, it becomes increasingly deformed by solar modulation effects, but its shape is reflected in the shape of the synchrotron emission spectrum, which is well established throughout the radio-frequency range 5 MHz–10 GHz. Webber (1983b) was thus able to construct a composite cosmic-ray electron spectrum valid down to \( E \sim 0.3 \text{ GeV} \), by matching the un-normalized spectrum inferred from the synchrotron-emission spectral shape to the normalized spectrum directly measured at Earth.

The cosmic-ray electron spectrum derived by Webber (1983b) is not an exact power law, so that the values of \( \gamma \) and \( K_\varepsilon \) appearing in Eq. (7) vary weakly with electron energy. The relevant energy here is that corresponding to \( \nu = 408 \text{ MHz} \) (the radio frequency at which the Galactic synchrotron emissivity has been modeled) and \( B = 5 \mu \text{ G} \) (the estimated magnetic-field strength near the Sun; see below), or, according to Eq. (8), \( E = 2.3 \text{ GeV} \). At this energy, Webber’s (1983b) composite spectrum yields \( \gamma = 2.5 \) and \( K_\varepsilon = (210 \text{ m}^2 \text{s}^{-1} \text{sr}^{-1} \text{GeV}^{-1}) \times (2.3 \text{ GeV})^{-3} \). Strictly speaking, the last value applies to the warm Local Cloud and probably constitutes a slight overestimate for the space-averaged vicinity of the Sun. We will nonetheless consider it as the space-averaged value of \( K_\varepsilon \) near the Sun, for the same reasons as those invoked in Sec. IV.B, when identifying the space-averaged value of \( P_{CR} \) near the Sun with its value in the Local Cloud.

We now introduce the above values of \( \nu, \gamma, \) and \( K_\varepsilon \), together with \( (\varepsilon_e)_0 = 8.2 \text{ K kpc}^{-1} \) [see Eq. (9)], into Eq. (7), whereupon we obtain for the local magnetic-field strength \( B_0 = 5 \mu \text{ G} \). The ensuing local magnetic pressure is \( (P_M)_0 = 10.0 \times 10^{-13} \text{ dyn cm}^{-2} \).

In summary, we have derived the local values of the interstellar magnetic and cosmic-ray pressures, and we have argued that their large-scale spatial distribution follows that of the synchrotron emissivity to the power \( 4/(\gamma + 5) = 0.53 \). Thus, if we denote by \( F_\varepsilon(R,Z) \) the expression within curly brackets in Eq. (9), we may write

\[
P_M(R,Z) = (10.0 \times 10^{-13} \text{ dyn cm}^{-2}) \left[ F_\varepsilon(R,Z) \right]^{0.53} \tag{10}
\]

and

\[
P_{CR}(R,Z) = (12.8 \times 10^{-13} \text{ dyn cm}^{-2}) \left[ F_\varepsilon(R,Z) \right]^{0.53} \tag{11}
\]

For comparison, the thermal pressure inside the Local Cloud probably lies in the range \( 2.8–6.8 \times 10^{-13} \text{ dyn cm}^{-2} \) (Frisch, 1995; see Sec. IV.B), while that in the Local Bubble was recently estimated at \( \approx 21 \times 10^{-13} \text{ dyn cm}^{-2} \) (Snowden et al., 1998; see Sec. III.E). At the solar circle \( R = R_\odot \), Eqs. (10) and (11) lead to the vertical profiles drawn in Fig. 7.

According to Eq. (10), the interstellar magnetic-field strength, \( B \approx \sqrt{P_M} \), has a radial scale length \( \approx 12 \text{ kpc} \) and a vertical scale height at the Sun \( \approx 4.5 \text{ kpc} \), which are both significantly greater than the values found for the uniform field component from Faraday-rotation measurements (see Sec. IV.A). One can argue that, vertically, the interstellar magnetic field becomes less regular away from the midplane (Boulares and Cox, 1990). Radially, the discrepancy is more intriguing, as the outward increase in the degree of linear polarization of the synchrotron radiation from external galaxies indicates that the regular field component falls off with radius less rapidly than the total field strength (Heiles, 1995). A possible explanation, offered by Heiles (1995), is that the ionized regions sampled by Faraday rotation measures both have a weaker magnetic field than the neutral regions and occupy an outward-decreasing fraction of the interstellar volume.
For Galactic cosmic rays, Eq. (11) predicts a radial scale length $= 6$ kpc, which is barely half that gathered from $\gamma$-ray observations (see Sec. IV.B). Although part of the difference is likely due to uncertainties in the modeled $\gamma$-ray and radio emissivities and in the fraction of these emissivities truly attributable to cosmic rays, an important source of disagreement could reside in $\gamma$-ray observations’ being systematically biased toward the dense ISM phases. In this regard, it is noteworthy that the cosmic-ray scale length inferred from synchrotron measurements is easier to reconcile with the short scale length of the presumed cosmic-ray sources. Finally, the cosmic-ray equivalent scale height in Eq. (11) is $= 2$ kpc at the solar circle, consistent with the maximum value $= 3$ kpc allowed by the diffusion-convection models of Bloemen et al. (1993) and Webber et al. (1992).

V. HOW EVERYTHING FITS TOGETHER

Now that we have reviewed the different constituents of the ISM, let us look at their interactions as well as their relations with stars.

A. Role played by stars

As already alluded to in the previous sections, stars affect the interstellar matter essentially through their radiation field, their wind, and, in some cases, their terminal supernova explosion. Globally, the massive, luminous O and B stars are by far the dominant players, even though they represent but a minor fraction of the stellar population (Abbott, 1982; Van Buren, 1985). Low-mass stars appear on the scene only for short periods of time, during which they have important outflows or winds.

Stellar radiation photons, above all the energetic UV photons from O and B stars, have a threefold direct impact on the interstellar matter. (1) They dissociate $H_2$ molecules (provided $\lambda < 1120 \\AA$) at the surface of molecular clouds (Federman et al., 1979). More generally, they dissociate molecules such as $H_2$, CO, OH, $O_2$, $H_2O$, etc. in photodissociation regions (Hollenbach and Tielens, 1999). (2) They ionize the immediate vicinity of O and B stars, thereby creating compact $H_\text{II}$ regions, and they ionize more remote diffuse areas, which together constitute the warm ionized medium (see Sec. III.D). In neutral regions, they ionize elements such as C, Mg, Si, and S, whose ionization potential lies below the 13.6-eV threshold of hydrogen (Kulkarni and Heiles, 1987). (3) They heat up the interstellar regions that they ionize to a temperature of $\sim 8000$ K, by imparting an excess energy to the liberated photoelectrons (see Sec. III.D). They also contribute to the heating of neutral regions, mainly through the ejection of photoelectrons from dust grains and through the radiative excitation of $H_2$ molecules followed by collisional deexcitation (see Secs. III.B, III.C, and III.F). As a side effect of ionization and heating by stellar photons, the traditional $H_\text{II}$ regions reach high thermal pressures, which cause them to expand into the ambient ISM.

Stellar winds are emitted by stars of all masses. Low-mass stars are involved only for limited periods in their lifetime. Early on, just before joining the main sequence (prolonged phase of hydrogen burning in the stellar core), they experience energetic, more or less collimated outflows (Lada, 1985). Toward the end of their life, after they have moved off the main sequence, they successively pass through the red-giant, asymptotic-giant-branch, and planetary-nebula stages, during which they again lose mass at a very fast rate (Salpeter, 1976; Knapp et al., 1990). High-mass stars suffer rapid mass loss throughout their lifetime (Conti, 1978; Bieging, 1990). Their wind becomes increasingly powerful over the course of the main-sequence phase (see, for example, Schaller et al., 1992), and, if their initial mass exceeds $\sim 32 M_\odot$, the wind reaches a climax during a brief post-main-sequence Wolf-Rayet phase (Abbott and Conti, 1987).

Supernovae come in two types. Type-I supernovae arise from old, degenerate low-mass stars, which supposedly are accreting from a companion and undergo a thermonuclear instability upon accumulation of a critical mass. Type-II supernovae arise from young stars with initial mass $\geq 8 M_\odot$, whose cores collapse gravitationally once they have exhausted all the star’s fuel (Woosley and Weaver, 1986). Both types of supernovae release an amount of energy $\approx 10^{51}$ ergs (Chevalier, 1977).

To a large extent, stellar winds and supernova explosions act in qualitatively similar ways, although supernova explosions are more sudden and usually far more spectacular. First of all, both constitute an important source of matter for the ISM. Since this matter has been enriched in heavy elements by the thermonuclear reactions taking place inside the stars, the metallicity of the ISM is gradually enhanced. The main contributors to the injection of mass into the ISM are the old red-giant, asymptotic-giant-branch, and planetary-nebula stars (Salpeter, 1976; Knapp et al., 1990). As was mentioned in Sec. III.F, these cool stars are also outstanding in that their wind carries away newly formed dust in addition to the regular gas.

Second, stellar winds and supernova explosions forge the structure of the ISM and are largely responsible for both its multiphase nature and its turbulent state. Here the main contributors are the young, massive O and B stars (Abbott, 1982; Van Buren, 1985). To start with, the wind from a massive star blows a cavity of hot gas in the surrounding ISM and compresses the swept-up interstellar gas into a rapidly expanding circumstellar shell (Castor et al., 1975; Weaver et al., 1977). If it is initially more massive than $\approx 8 M_\odot$, the star explodes at the end of its lifetime (between $\sim 3$ Myr for a $120 M_\odot$ O3 star and $\sim 38$ Myr for a $8 M_\odot$ B3 star; Schaller et al., 1992), and the shock wave driven by the explosion pursues, in an amplified fashion, the action of the wind, sweeping up a lot more interstellar matter into the expanding shell and greatly enlarging the hot cavity enclosed by the shell. The compressed swept-up gas, at the elevated postshock pressure, radiates efficiently, cools down, and collapses, so that the shell soon becomes cold and dense (Woltjer,
1972; Chevalier, 1977; Cioffi et al., 1988). Part of it may even turn molecular after typically ~1 Myr (McCray and Kafatos, 1987).

If the shell collides with a comparatively massive interstellar cloud or, finally, when the shock expansion velocity slows to roughly the external “signal speed” (a generalized sound speed, based on the total pressure, i.e., the gas+magnetic+cosmic ray pressure, rather than the purely thermal pressure), the shell begins to break up and lose its identity. The resulting shell fragments keep moving independently of each other and start mixing with the interstellar clouds; at this point, the shell is said to merge with the ambient ISM. Meanwhile, the hot rarefied gas from the interior cavity comes into contact with the ambient interstellar gas, mixes with it, and cools down—through thermal conduction followed by radiation—to a temperature of ~10^4 K. Ultimately, what an isolated massive star leaves behind is a cavity of hot rarefied gas surrounded by an increasingly thick layer of warm gas, plus several fragments of cold dense matter moving at velocities ~10 km s^-1. These fragments, be they atomic or molecular, appear to us as interstellar clouds.

The majority of O and B stars are not isolated, but grouped in clusters and associations (see the catalog of Garmany and Stencel, 1992), so that their winds and supernova explosions act collectively to engender superbubbles (McCray and Snow, 1979; Heiles, 1987). A superbubble behaves qualitatively like an individual supernova remnant, except that it has a continuous supply of energy. For the first 3 Myr at least, this energy supply is exclusively due to stellar winds, whose cumulative power rises rapidly with time. Supernovae start exploding after ~3 Myr, and within ~2 Myr they overpower the winds. From then on, the successive supernova explosions continue to inject energy into the superbubble at a slowly decreasing rate, depending on the initial mass function of the progenitor stars, until ~40 Myr (McCray and Kafatos, 1987; Ferrière, 1995). The time dependence of the energy deposition rate is in fact complicated by the likely spread in star formation times, with deep implications for the superbubble’s growth (Shull and Saken, 1995). Altogether, stellar winds account for a fraction comprising between ~12% (Ferrière, 1995) and ~17% (Abbott, 1982) of the total energy input.

Type-I supernovae are less frequent than their Type-II counterparts (see Sec. V.B). All of them are uncorrelated in space, and they have basically the same repercussions on the ISM as isolated Type-II supernovae.

In the local ISM (see Ferrière, 1998b), the remnant of a typical isolated supernova grows for ~1.5 Myr, and reaches a maximum radius ~50 pc. An “average superbubble,” produced by 30 clustered Type-II supernovae (see Sec. V.B), grows for ~13 Myr to a radius varying from ~200 pc in the Galactic plane to ~300 pc in the vertical direction. This vertical elongation is a direct consequence of ISM stratification: because the interstellar density and pressure fall off away from the midplane, superbubbles encounter less resistance and therefore manage to expand farther vertically than horizontally (see Tomisaka and Ikeuchi, 1986; Mac Low and McCray, 1988).

The hot gas created by supernova explosions and stellar winds in the Galactic disk rises into the halo under the effect of its buoyancy. In the course of its upward motion, it cools down, almost adiabatically at first, then through radiative losses, and eventually condenses into cold neutral clouds. Once formed, these clouds fall back ballistically toward the Galactic plane. The existence of such a convective cycle of interstellar matter between the disk and the halo was first suggested by Shapiro and Field (1976), who dubbed it the “Galactic fountain.” Detailed models of the Galactic fountain with specific observationally verifiable predictions were developed by Bregman (1980) and many other authors (see Bregman, 1996 for a review). Norman and Ikeuchi (1989) proposed a slight variant in which the upward flow of hot gas concentrates in the “chimneys” formed by elongated superbubbles that have broken out of the Galactic disk. An attractive aspect of the Galactic fountain is that it furnishes a natural explanation for the observed existence of high-velocity clouds (H I clouds whose measured velocity ≥90 km s^-1 is too large to be solely due to the large-scale differential rotation) as well as for their measured velocity distribution, characterized by a notable up-down asymmetry in favor of infalling clouds. For the interested reader, the various possibilities for the origin of high-velocity clouds are reviewed by Wakker and van Woerden (1997).

Let us now inquire into the long-term evolution of the cold shell fragments produced by supernova explosions. Some of them remain mostly atomic and are observed as diffuse atomic clouds, moving randomly at velocities ~10 km s^-1. Others, typically those arising from old superbubbles, become largely molecular, at least away from their surface. These molecular fragments are responsive to self-gravity, which, past a critical threshold, drives them unstable to gravitational collapse. The collapse of individual fragments or pieces thereof eventually leads to the formation of new stars (see, for example, Shu et al., 1987), which, if sufficiently massive, may in turn initiate a new cycle of matter and energy through the ISM.

The idea of sequential OB star formation has been discussed by several authors, including Mueller and Arnett (1976), Elmegreen and Lada (1977), and McCray and Kafatos (1987). In particular, it has been suggested that, once triggered in a molecular-cloud complex, OB star formation could propagate and spread throughout the entire complex (Elmegreen and Lada, 1977) and even possibly to more distant areas (Elmegreen, 1987). However, as pointed out by McCray and Kafatos (1987), self-induced star formation probably represents but a secondary process, the primary triggering mechanism remaining, in all likelihood, gas compression at the shock waves associated with the Galactic spiral arms (Lin and Shu, 1964).

On the other hand, the process of OB star formation in the Galaxy is, to some extent, self-regulated. Norman
and Silk (1980) had already argued that outflows from pre-main-sequence low-mass stars in a molecular cloud provide a continuous dynamic input, which maintains the turbulent pressure at an adequate level to support the cloud against gravitational collapse and therefore limits further star formation. Franco and Shore (1984) applied a modified version of this argument to OB stars: through their radiation field, powerful wind, and supernova explosion, OB stars inject energy and momentum into their surrounding medium, but because they do so much more vigorously than low-mass stars, they rapidly disrupt their parent molecular cloud and bring local star formation to a halt. Other, milder regulatory mechanisms have been advocated, involving either photoionization (McKee, 1989) or grain photoelectric heating (Parravano, 1988).

Beside their obvious impact on the interstellar matter, stars are equally vital for Galactic cosmic rays and magnetic fields. As pointed out in Sec. IV.B, stars are the likely birthplaces of most Galactic cosmic rays, and the shock waves sent by supernova explosions constitute important sites of further cosmic-ray acceleration. Likewise, as we shall touch upon in Sec. V.C, interstellar magnetic fields could have their beginnings in stellar interiors; while this possibility remains to be proven, there is now little doubt that, once a tiny magnetic field has been created, the turbulent motions generated by supernova explosions rapidly amplify it.

B. Supernova parameters

Given the particular importance of supernovae, we considered it necessary to devote a full subsection to a description of their parameters.

The Galactic frequency of both types of supernovae can be estimated by monitoring their rate of occurrence in a large number of external galaxies similar to the Milky Way. From an eight-year observation program of 855 Shapley-Ames galaxies, Evans et al. (1989) derived average frequencies of $0.2h^2$ SNu for Type-I supernovae and $1.7h^2$ SNu for Type-II supernovae in Sbc-Sd galaxies, where $h$ is the Hubble constant in units of 100 km s$^{-1}$ Mpc$^{-1}$ and 1 supernova unit (SNu) represents 1 supernova per $10^{10}(L_B)_{\odot}$ per 100 yr. More recently, Cappellaro et al. (1997) combined five independent supernova searches [including that of Evans et al. (1989)], and from the resulting sample of 7773 galaxies they derived average frequencies of $0.41h^2$ SNu for Type-I supernovae and $1.69h^2$ SNu for Type-II supernovae in Sbc-Sc galaxies.

The value of the Hubble constant, which gives the present expansion rate of the universe, was under heavy debate for over half a century, until various kinds of observations made in the last few years finally converged to a narrow range $\approx 60–70$ km s$^{-1}$ Mpc$^{-1}$ (Branch, 1998; Primack, 2000; Freedman et al., 2001). If we choose the median value of this range, corresponding to $h = 0.65$, and assume that the Milky Way is an average Sbc galaxy with a blue luminosity of $2.3 \times 10^{10}(L_B)_{\odot}$ (van den Bergh, 1988; van der Kruit, 1989), we find that the results of Cappellaro et al. (1997) translate into a Type-I supernova frequency

$$f_I \approx \frac{1}{250 \text{ yr}}$$

and a Type-II supernova frequency

$$f_{II} \approx \frac{1}{60 \text{ yr}}$$

in our Galaxy.

The corresponding total supernova frequency in our Galaxy is $\approx 1/(48 \text{ yr})$, in reasonably good agreement with the evidence from historical supernovae. Only five Galactic supernovae brighter than zeroth magnitude were recorded in the last millennium, but it is clear that many more supernovae occurred without being detected from Earth, mainly because they remained obscured by the interstellar dust. Tamman et al. (1994) extrapolated from the five recorded events, with the help of a detailed model of the Galaxy accounting for obscuration by dust, and they concluded that a Galactic supernova frequency $\approx 1/(26 \text{ yr})$—with a large uncertainty due to small-number statistics—could reproduce the historical observations.

The spatial distribution of supernovae is even more uncertain than their frequency. In external spiral galaxies, both types tend to concentrate in the spiral arms. Like the bright massive stars, Type-II supernovae are tightly confined to the arms, whereas Type-I supernovae have a more spread-out distribution, similar to that of the general stellar population (McMillan and Ciardullo, 1996). Radially, supernovae appear to be distributed with an exponential scale length of $\sim 2.4–3.8$ kpc for Type-I supernovae and $\sim 3.0–5.5$ kpc for Type-II supernovae, where the first and second values of each range refer to the regions $R < 5.8$ kpc and $R > 5.8$ kpc, respectively, and both values have been rescaled to a Hubble constant of 65 km s$^{-1}$ Mpc$^{-1}$ (Bartunov et al., 1992; see also van den Bergh, 1997).

It is also possible to infer the spatial distribution of supernovae in our Galaxy from that of related objects. For instance, we may reasonably suppose that Type-I supernovae follow the distribution of old disk stars, with an exponential scale length $\approx 4.5$ kpc along $R$ and an exponential scale height $\approx 325$ pc along $Z$ (Freeman, 1987). Adopting the Galactic frequency given by Eq. (12), we may then write the Galactic Type-I supernova rate per unit area as

$$\sigma_I(R) = (4.8 \text{ kpc}^{-2} \text{Myr}^{-1}) \exp \left( - \frac{R - R_{\odot}}{4.5 \text{ kpc}} \right)$$

and their rate per unit volume at the solar circle as
\[ R_t(Z) = (7.3 \text{kpc}^{-3} \text{Myr}^{-1}) \exp \left( -\frac{|Z|}{325 \text{ pc}} \right). \]  

For Type-II supernovae, we may use either H II regions, which are produced by their luminous progenitor stars, or pulsars, which are the likely leftovers of core-collapse explosions. McKee and Williams (1997) found that Galactic giant H II regions are approximately distributed in a truncated exponential disk with a radial scale length \( \approx 3.3 \text{ kpc} \) over the radial range \( R = 3-11 \text{ kpc} \). Galactic pulsars, for their part, were shown to be radially distributed according to a rising Gaussian with a scale length \( \approx 2.1 \text{ kpc} \) for \( R < 3.7 \text{ kpc} \) and a standard Gaussian with a scale length \( \approx 6.8 \text{ kpc} \) for \( R > 3.7 \text{ kpc} \) (Narayan, 1987; Johnston, 1994). Their vertical distribution at birth can be approximated by the superposition of a thin Gaussian disk with a scale height \( \approx 212 \text{ pc} \) and a thick Gaussian disk with three times the same scale height, containing, respectively, 55% and 45% of the pulsar population (Narayan and Ostriker, 1990). In view of Eq. (13), the pulsar model leads to a Galactic Type-II supernova rate per unit area \( \sigma_{\text{II}}(R) = (27 \text{kpc}^{-2} \text{Myr}^{-1}) \times \begin{cases} 3.55 \times \exp \left( \frac{R}{21.5 \text{kpc}} \right)^2, & R < 3.7 \text{kpc} \\ \exp \left( \frac{R^2 - R^3}{(6.8 \text{kpc})^2} \right), & R > 3.7 \text{kpc} \end{cases} \) and a Type-II supernova rate per unit volume at \( R_\odot \) \[ R_{\text{II}}(Z) = (50 \text{kpc}^{-3} \text{Myr}^{-1}) \times \begin{cases} 0.79 \times \exp \left( -\frac{Z}{212 \text{ pc}} \right)^2, & \text{for } Z < 3.7 \text{kpc} \\ 0.21 \times \exp \left( -\frac{Z}{636 \text{ pc}} \right)^2, & \text{for } Z > 3.7 \text{kpc} \end{cases} \] Note that the mean height predicted by Eq. (17) is significantly larger than the mean height of OB stars (\( \approx 90 \text{ pc} \); Miller and Scalo, 1979), consistent with the fact that these stars are usually born close to the midplane and progressively increase their average distance from it as they grow older. The Galactic supernova rates per unit area and per unit volume are plotted in Figs. 8 and 9, respectively, for both types of supernovae.

Before closing this subsection, let us estimate the fraction of Type-II supernovae that are clustered and the way they are distributed amongst different clusters. In the catalog of 195 Galactic O stars compiled by Gies (1987), 71% lie in groups and 29% lie in the field. For the O stars in groups, we may use the radial peculiar velocities tabulated by Gies together with the assumption of isotropy in peculiar-velocity space to reconstruct the distribution in total peculiar velocity. This distribution clearly possesses an excess of high-velocity stars, amounting to \( \approx 15\% \) of the group stars, i.e., \( \approx 11\% \) of all O stars. According to Gies's interpretation, these high-velocity stars were recently ejected from their native cluster and will end up in the field. From this, we conclude that \( \approx 60\% \) of the O stars were born and will remain in groupings, while \( \approx 40\% \) of them will die in the field.

---

\footnote{In principle, one could also rely on the pulsar birthrate to evaluate the Type-II supernova frequency in our Galaxy. Unfortunately, the pulsar birthrate is still poorly known. It is, however, reassuring that its estimated value \( \approx 1/(30-120 \text{ yr}) \) (Lyne et al., 1985) or \( \approx 1/(100 \text{ yr}) \) (Narayan and Ostriker, 1990) is compatible with Eq. (13).}
Can these figures be extended to all Type-II supernova progenitors? Humphreys and McElroy (1984) compiled a list of all Galactic luminous stars known to date and found that 47% of them are grouped. Since their list contains on average older stars than Gies’s (1987), it is not surprising that a larger fraction of their stars appear in the field. Indeed, stars born in a group may after some time be observed in the field, either because they have been ejected from the group or because the group has dispersed. If we accept that ≈11% of the O stars are observed as group members but will end up in the field as a result of ejection, and if we assume that OB associations disperse into the field after ∼15 Myr (Blauuw, 1964; Mihalas and Binney, 1981, p. 164), we find that Humphreys and McElroy’s (1984) compilation is consistent with 60% of all Type-II supernovae’s being clustered.

Clustered Type-II supernovae are very unevenly divided between superbubbles. In other words, the number of clustered supernovae contributing to one superbubble, \( N \), is extremely variable. The distribution of \( N \) can be deduced from the observed luminosity distribution of HII regions, either in external galaxies (Kennicutt et al., 1989; Heiles, 1990) or in our own Galaxy (McKee and Williams, 1997). Both kinds of observations suggest a power-law distribution in \( N^{-2} \). Moreover, relying on local observations of OB stars and stellar clusters, Ferrière (1995) estimated that \( N \) averages to \( \approx 30 \) (in agreement with an earlier estimate by Heiles, 1987) and varies roughly between 4 and \( \approx 7000 \).

C. Role played by cosmic rays and magnetic fields

Interstellar cosmic rays, those extremely energetic and electrically charged particles pervading interstellar space, impinge on interstellar matter in three important ways. (1) They contribute to its ionization, through direct collisions with interstellar gas particles (Spitzer and Tomasko, 1968). (2) They constitute a triple source of heating, arising from the excess energy carried away by the electrons released in cosmic-ray ionization (Field et al., 1969), from Coulomb encounters with charged particles of the ordinary gas (Field et al., 1969), and from the damping of Alfvén waves excited by cosmic rays’ streaming along magnetic-field lines (Wentzel, 1971). (3) They are dynamically coupled to the interstellar matter via the intermediary of the interstellar magnetic field, more specifically, by their gyromotion in the perpendicular direction and by their scattering off of self-excited Alfvén waves in the parallel direction (see Sec. IV.B). In consequence, they exert their full pressure on the interstellar matter, thereby affecting its dynamics.

The interstellar magnetic field acts on interstellar matter through the Lorentz force. Of course, the field acts directly on the charged particles only, but its effect is then transmitted to the neutrals by ion-neutral collisions (Spitzer, 1958). Apart from the densest parts of molecular clouds, whose degree of ionization is exceedingly low \( (x \sim 10^{-5} – 10^{-6}; \text{Shu et al., 1987}) \), virtually all interstellar regions are sufficiently ionized \( (x \sim 4 \times 10^{-4} – 10^{-3}) \) in cold atomic clouds and \( x \sim 0.007 – 0.05 \) in the warm atomic medium; Kulkarni and Heiles, 1987) for their neutral component to remain tightly coupled to the charged component and hence to the magnetic field.

At large scales, the interstellar magnetic field helps to support ordinary matter against its own weight in the Galactic gravitational potential, and it confines cosmic rays to the Galactic disk. In this manner, both magnetic fields and cosmic rays partake in the overall hydrostatic balance of the ISM and influence its stability. Boulares and Cox (1990) were the first authors to fully appreciate the importance of magnetic fields and cosmic rays in the hydrostatic balance. By the same token, they managed to solve the long-standing problem of apparent mismatch between the total interstellar pressure at a given point and the integrated weight of overlying interstellar material: by adopting higher magnetic and cosmic-ray pressures than previously estimated, they were able to bring the total pressure at low \( |Z| \) into agreement with the integrated weight, and by including the magnetic tension force at high \( |Z| \), they could explain why the weight integral falls off faster than the total pressure.

As magnetic and cosmic-ray pressures inflate the gaseous disk, they tend to make it unstable to a generalized Rayleigh-Taylor instability, now known in the astrophysical community as the Parker instability (Parker, 1966; see also Zweibel, 1987). When this instability develops, magnetic-field lines ripple, and the interstellar matter slides down along them toward the magnetic troughs, where it accumulates. This whole process, it has been suggested, could give birth to new molecular-cloud complexes and ultimately trigger star formation (Mouschovias et al., 1974; Elmegreen, 1982).

At smaller scales, the interstellar magnetic field affects all kinds of turbulent motions in the ISM. Of special importance is its impact on supernova remnants and superbubbles (see Tomisaka, 1990; Ferrière et al., 1991; Slavin and Cox, 1992). First, the background magnetic pressure acting on their surrounding shells directly opposes their expansion. Second, the magnetic tension in the swept field lines gives rise to an inward restoring force, while the associated magnetic pressure prevents the shells from fully collapsing and, therefore, keeps them relatively thick. Third, the enhanced external “signal speed” causes the shells to merge earlier than they would in an unmagnetized medium. All three effects conspire to lower the filling factor of hot cavities (Ferrière et al., 1991; Slavin and Cox, 1993).

The interstellar magnetic field also constrains the random motions of interstellar clouds that are not, or are no longer, parts of coherent shells. The basic physical idea is the following: Interstellar clouds are magnetically connected to their environment, namely, to the intercloud medium and possibly to neighboring clouds, through the magnetic-field lines that thread them. When a given cloud moves relative to its environment, these field lines get deformed, and the resulting magnetic tension force modifies the cloud’s motion, transferring part of its momentum to its environment (Elmegreen, 1981). Likewise, angular momentum can be transferred from a...
cloud to its environment by magnetic torques (Mouschovias, 1979; Mouschovias and Paleologou, 1979). The latter mechanism is particularly relevant to the star formation process, as it allows the contracting protostellar cores to get rid of angular momentum (Mouschovias and Morton, 1985).

Finally, magnetic fields play a crucial role in the support of molecular clouds against their self-gravity and in the eventual gravitational collapse of protostellar cores. The magnetic support of molecular clouds is essentially provided by magnetic pressure gradients in the directions perpendicular to the average field and, presumably, by nonlinear Alfvén waves in the parallel direction (see the review paper of Shu et al., 1987). In the case of protostellar cores, magnetic support is insufficient to prevent their ultimate gravitational collapse. This is generally because their degree of ionization is so low that neutrals are not perfectly tied to magnetic-field lines, which enables them to drift inwards under the pull of their self-gravity and eventually form stars (Nakano, 1979; Mestel, 1985).

The last issue we would like to discuss pertains to the very origin of interstellar magnetic fields. The most likely scenario to date involves a hydromagnetic dynamo, based on the concept that the motions of a conducting fluid embedded in a magnetic field generate electric currents, which, under favorable conditions, can amplify the original magnetic field. In the Galaxy, the dynamo process appeals to a combination of large-scale differential rotation and small-scale turbulent motions (Steenbeck et al., 1966; Parker, 1971; Vainshtein and Ruzmaikin, 1971). Remember that the interstellar gas is tied to magnetic-field lines, or, stated more appropriately here, magnetic-field lines are “frozen into” the interstellar gas. Accordingly, the large-scale differential rotation stretches field lines in the azimuthal direction about the Galactic center, thereby amplifying the azimuthal component of the large-scale magnetic field. Meanwhile, its radial component is amplified by small-scale turbulent motions through a mechanism called the “alpha effect,” the principle of which can be described as follows: turbulent motions, taking place in a rotating medium, are acted upon by the Coriolis force, which imparts to them a preferred sense of rotation; in consequence, the small-scale magnetic loops that they produce are preferentially twisted in a given sense, and the net result is the creation of a mean magnetic field component in the direction perpendicular to the prevailing field. In addition to being responsible for the alpha effect, turbulent motions also contribute to the vertical escape of magnetic-field lines and to their spatial diffusion.

It is important to realize that the operation of the Galactic dynamo requires a seed magnetic field to initiate the amplification process. Several possibilities have been advanced regarding the nature and origin of this seed field (see Rees, 1987 for a review). Briefly, the seed field could be an extragalactic magnetic field already present in the universe prior to galaxy formation, or it could arise in the protogalaxy as a result of charge separation due to electrons interacting with the microwave background photons, or else it could originate in the first generation of stars and be expelled into the ISM by their winds and/or supernova explosions.

In the preceding sections, we have emphasized the major role played by stars in various aspects of the ISM, notably in its extremely heterogeneous character, its continual agitation, its partial ionization, its heating, and the formation and acceleration of cosmic rays. The present section has shown that stars are also key agents in the generation and amplification of interstellar magnetic fields. Not only do they constitute potential candidates for triggering the Galactic dynamo, but also, and more importantly, they very likely drive much of the turbulence necessary for the alpha effect and hence for dynamo action.

ACKNOWLEDGMENTS

I wish to thank C. Heiles and J. M. Shull for their careful reading of the paper and for their numerous suggestions, which greatly helped improve its content. I am also grateful to T. M. Dame, J. M. Dickey, L. M. Haffner, M. H. Heyer, and B. P. Wakker, who kindly provided me with the maps and spectra appearing in Figs. 3, 4, 5, 6, and 12.

APPENDIX

The purpose of this appendix is to explain how the large-scale spatial distribution of an interstellar gas component can be mapped by means of one of its radio emission lines (say, the 21-cm emission line of H1), once a model of Galactic rotation has been adopted (see Fig. 10). Radio spectra including the chosen emission line are taken in a large number of directions scanning the sky. For each direction, the emission line spreads out over a range of wavelengths around the rest wavelength \( \lambda_0 \), as illustrated in Fig. 11(a) (with \( \lambda_0 = 21 \text{ cm} \) in the case of H1).

If the interstellar gas is transparent or, in more astrophysical terms, optically thin for the considered line (as is often assumed for the H1 21-cm emission line), photons emitted toward the Earth travel all the way to the receiver without being reabsorbed by the gas, so that they automatically contribute to the line intensity. Therefore the specific intensity \( I_\lambda \) at any given wavelength \( \lambda \) is directly proportional to the amount of material producing the emission at that wavelength. The latter can be traced back to a given location in the Galaxy by noting that the Doppler shift \( (\lambda - \lambda_0)/\lambda_0 \) is straightforwardly related to the line-of-sight velocity with respect to the Earth, which, in turn, can be converted into a distance from us with the help of the Galactic rotation curve. In fact, inside the solar circle, there exist two dis-

---

9Beside the direct, stabilizing effect of magnetic fields, there exists an indirect, possibly destabilizing effect, due to the distortion of the equilibrium configuration (Zweibel, 1990).
tances corresponding to a given line-of-sight velocity (for instance, 1 and 1\' in Fig. 10 have the same line-of-sight velocity), but this twofold distance ambiguity may, under some conditions, be resolved with appropriate pattern recognition techniques.

If the considered line is optically thick (like the CO 2.6-mm emission line), most of the emitted photons are reabsorbed before reaching the receiver, and the line intensity saturates at a level that is a function of the gas temperature, independent of the amount of emitting material. The case of the CO 2.6-mm emission line, however, is a little more subtle. An idealized CO emission spectrum is sketched in Fig. 11(b). Each feature in the spectrum can be attributed to a clump of emitting material in the observed direction, whose line-of-sight velocity is that deduced from the Doppler shift of the feature.

FIG. 10. Schematic view of our Galaxy seen from above. GC indicates the position of the Galactic center and \( \odot \) the position of the Sun. \( l \) is the Galactic longitude defined with respect to the Sun and measured counterclockwise from the direction to the Galactic center. I, II, III, and IV denote the first, second, third, and fourth Galactic quadrants. The interstellar gas is assumed to rotate clockwise about the Galactic center at a rate decreasing away from it. In a frame rotating with the Sun, the gas interior to the solar orbit rotates clockwise, while the exterior gas rotates counterclockwise, as indicated by the thick arrows.

FIG. 11. Schematic profile of an emission line with rest wavelength \( \lambda_0 \) obtained in the direction \( l \) drawn in Fig. 10. \( I_\lambda \) is the specific intensity at wavelength \( \lambda \). \( \lambda_1 \) corresponds to point 1 or 1\' in Fig. 10, while \( \lambda_2 \) corresponds to point 2. (a) Generic case, directly applicable to the H I 21-cm emission line (adapted from Shu, 1982, p. 268); (b) case of the CO 2.6-mm emission line (adapted from Scoville and Sanders, 1987).

FIG. 12. Real emission spectra, for which the Doppler shift \((\lambda - \lambda_0)/\lambda_0\) on the \( x \) axis has already been converted to line-of-sight velocity: (a) typical H I emission spectra toward \((l = 90^\circ, b = 0^\circ)\) and \((l = 180^\circ, b = 10^\circ)\), from the Leiden-Dwingeloo Survey (courtesy of B. P. Wakker); (b) typical CO emission spectrum toward \((l = 20.7^\circ, b = 0^\circ)\), obtained with the 1.2-m telescope at the Harvard-Smithsonian Center for Astrophysics (courtesy of T. M. Dame).
The fact that individual features are well separated indicates that the probability of overlap is small and that every clump along the line of sight is fully represented in the spectrum. Once averaged over a small area in the plane of the sky, the CO emission spectrum has a profile similar to that drawn in Fig. 11(a), where the contribution from every emitting clump is fully taken into account. Thus, even though individual features are optically thick, the average emission line is effectively optically thin, and its specific intensity as a function of wavelength provides a direct measure of the amount of CO as a function of line-of-sight velocity, which, as before, translates into a distance from us—with a twofold ambiguity inside the solar circle. For a more detailed discussion on the supposedly linear relationship between the CO line intensity and its column density, the reader is referred to the review paper by Scoville and Sanders (1987).

A few real emission spectra from recent Galactic surveys are displayed in Fig. 12.

REFERENCES


