Cambridge University Press
978-0-521-63030-6 - An Introduction to Star Formation
Derek Ward-Thompson and Anthony P. Whitworth
Excerpt
More information

Chapter 1
Introduction

1.1 About this book

It can be argued that astronomy is the oldest science. Since pre-historic times humans have gazed at the night sky and wondered about the nature and origin of stars. We now believe we understand a great deal about the nature of stars, but many aspects of the origin of stars remain the subject of intense study to this day.

In this book we aim to introduce the reader to the fundamentals of the subject of star formation. We describe the background physics underlying theories of star formation, and take the reader to the frontiers of current knowledge of this subject. However, we will make clear as we go along the points where we reach material that is less well established.

One of the most fundamental observations in astronomy is the fact that the night sky appears to be full of stars. Yet the processes which lead to the formation of those stars have taken astronomers many years to work out. Unlocking the mysteries of star formation has required the use of new techniques and the opening of new wavelength regimes to astronomy. We describe the chief physical processes which are believed to be important for star formation, and point out the role which each branch of observational astronomy has played in solving the various problems associated with star formation.

In this chapter we begin by introducing some of the main constituents of a galaxy, namely the stars, the medium between the stars and the gravitational and magnetic fields. We discuss their spatial distribution, and introduce the life-cycle of a star and the way in which the formation

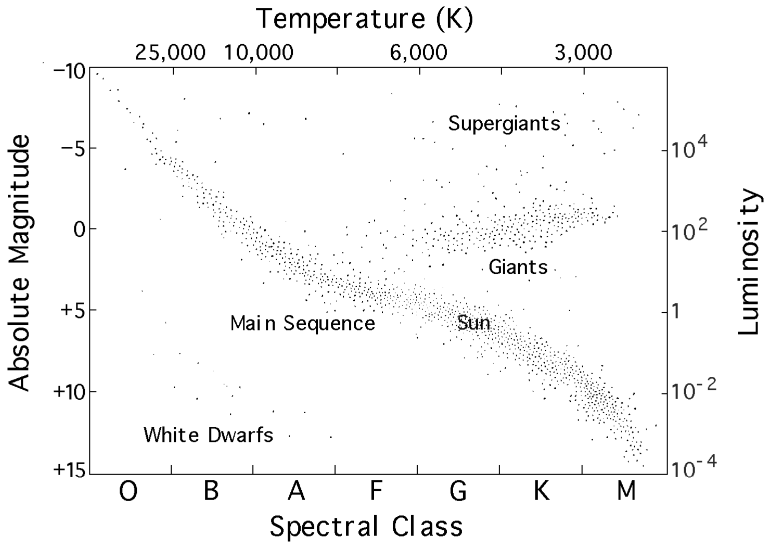

Fig. 1.1. A rough sketch of the Hertzsprung–Russell diagram, illustrating the main sequence, where a solar-type star spends the majority of its life.

of a star fits into this cycle. We introduce the sites where stars are formed and give a description of the initial mass function of stars, the explanation of which is one of the major challenges for any star-formation theory. We finish with a list of some of the chief objectives of star-formation theory.

1.2 The stellar life-cycle

The most important diagram for stellar evolution is known as the Hertzsprung–Russell (HR) diagram. The HR diagram plots the luminosity of a star against its colour. In this diagram it is seen that the majority of stars lie in a single strip along the diagram, known as the main sequence, wherein the brightest stars are also the bluest, while the faintest stars are the reddest.

Figure 1.1 sketches an HR diagram to illustrate the approximate position of the main sequence, and the position of the Sun, which is simply an ordinary main-sequence star. In the Universe today, a star of roughly solar mass spends the majority of its life-time as a main-sequence star, during which time its energy source is the fusion of hydrogen to helium, deep in the core of the star. A star that has not yet reached the main sequence on the HR diagram is known as a pre-main-sequence star. However, before a star reaches the main sequence it has to be formed from the material in interstellar space. This involves a series of stages of contraction and growth by accretion under the influence of gravity.

Later in its life, when the majority of the hydrogen in its core has been processed by fusion into helium, a star leaves the main sequence, expands to become a giant or supergiant star, and undergoes various stages of losing mass. The most violent and best known of these mass-losing stages occurs for the most massive stars, and is known as a supernova explosion. However, lower-mass stars also undergo phases when they eject material. This ejected material can then form some of the ingredients for subsequent generations of stars.

Hence we see that stars undergo a life-cycle in which new generations of stars are formed in part from the debris of previous generations of stars. In fact almost all of the material which forms a new generation of stars, apart from hydrogen and helium, is the product of fusion in the centres of stars of previous generations. Thus the formation of stars is not the beginning of a linear process, but is an integral part of a cyclic process. The subject of this book is that part of the stellar life-cycle which occurs prior to the main sequence. In this book we start at the birth-place of stars, and follow the progress from there all the way to the point at which a star joins the main sequence.

1.3 The space between the stars

The space between the stars is known as the interstellar medium, or the ISM. The principal constituents of the interstellar medium are matter (gas, dust and cosmic rays), electromagnetic radiation, a gravitational field, and a magnetic field. The composition of the matter is typically 70% hydrogen, 28% helium, and 2% heavier elements, such as oxygen, carbon and nitrogen. Most of the matter – around 99% – is in the gas phase.

The gaseous ISM can be modelled with four phases: regions of very diffuse, hot, ionised gas which form a network of interconnecting stellar-wind bubbles and supernova remnants; a warm, partially ionised gas filling most of the rest of the volume of the Galactic disc; small clouds of cool, neutral, mainly atomic gas; and larger clouds of cold, dense, mainly molecular gas.

These larger clouds are known variously as ‘dense clouds’, ‘dark clouds’ or ‘molecular clouds’, depending on the context. Even in the centres of these clouds, the densities are only about 10^{12} molecules per cubic metre, so the term ‘dense cloud’ may be somewhat misleading. However, this density is still much greater than the average ISM, where there are typically only about 10^{6} atoms per cubic metre. The largest clouds of molecular gas are known as ‘giant molecular clouds’ (GMCs),

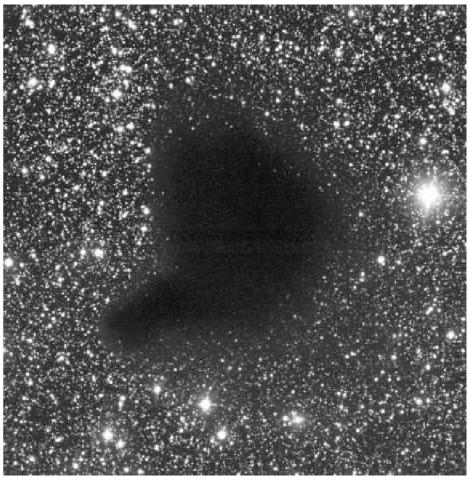

Fig. 1.2. An optical image of the dark cloud Barnard 68. Note how the background stars are not visible through the cloud.

or GMC complexes, and can contain masses up to a few million solar masses[†] and extend for tens of parsecs.[‡]

The term dark cloud arose because at optical wavelengths these clouds can be completely opaque. Figure 1.2 shows a picture of a dark cloud taken at visible wavelengths. Note how the light from the background stars, which can be seen across this field, is obscured by the material of the cloud in the centre of the image. However, modern astronomy utilises many more wavelengths than the optical. The infrared and radio wavelength regimes have become more important than the optical in studying star formation, as the regions in which stars form are not so opaque at these longer wavelengths.

The reason that clouds are dark is that they contain not only gas, but also dust, which is opaque to visible light. This dust consists of very small grains, i.e. particles of solid matter, less than a micron in size, and consisting mostly of silicates (sand), and carbon compounds, probably including graphite. Dust grains have similar sizes to the particles of cigarette smoke. By mass, the dust grains only represent about 1% of the total mass of the ISM, but this is still sufficient to block out much of the visible light.

The gas in the ISM is in a constantly changing chemical state. However, in the densest clouds most of the gas is normally molecular. This is firstly because the processes forming molecules in interstellar space – primarily two-body gas-phase reactions and catalysis on the surface of dust grains – proceed faster at higher density, and secondly

[†] 1 solar mass ($M_\odot$) = 2×10^{30} kg.
[‡] 1 parsec (pc) = 3×10^{16} m.

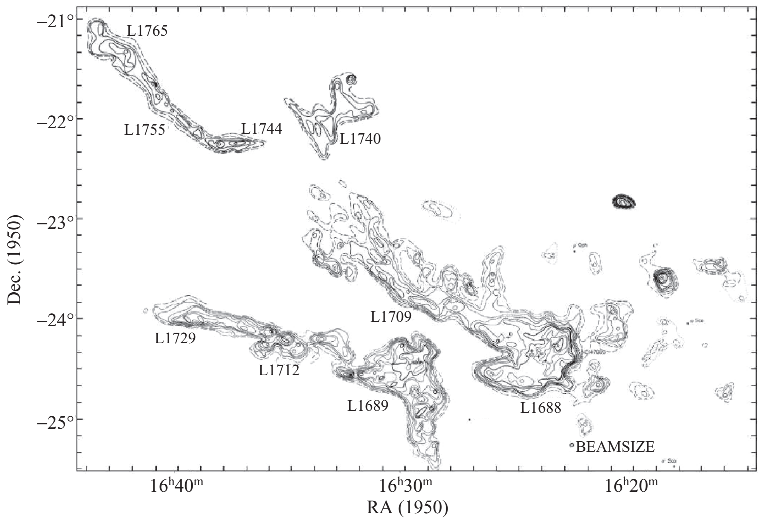

Fig. 1.3. A map of the Ophiuchus molecular cloud complex. The contours represent brightness of carbon monoxide (CO), which is taken as a tracer of the molecular gas as a whole.

because dust effectively shields the interior of a dense cloud from the ultraviolet (UV) radiation which destroys molecules.

Consequently the majority of the gas within the clouds is in molecular form, and the majority of a molecular cloud's mass is in the form of molecular hydrogen, H_2. Figure 1.3 shows a map of a molecular cloud in the constellation Ophiuchus. The sizes of molecular clouds range from $\sim$0.1 pc to $\sim$100 pc in diameter, and their densities range from 10^6 atoms per m^3 at their edges to greater than 10^{12} atoms per m^3 in their densest parts. GMCs are among the largest entities that we know of within galaxies.

In addition to gas and dust there are cosmic rays in the interstellar medium. Although these particles constitute a minute fraction of the rest mass in the interstellar medium, they travel at speeds approaching the speed of light, and consequently they have in total about as much kinetic energy as the rest of the matter put together.

The electromagnetic radiation in the interstellar medium comes from various sources. The principal sources of optical radiation are stars, but at longer wavelengths there is also continuum radiation from interstellar dust, line radiation from interstellar gas, and the cosmic microwave background (CMB – a relic of the Big Bang). Stellar radiation is dominated by optical and UV photons ($\lambda \sim$ 100–1000 nm), corresponding to stellar surface temperatures in the range 3000–30 000 K. The cosmic microwave background is characterised by millimetre and submillimetre photons ($\lambda \sim$ 0.3–3 mm $\equiv$ 3–30 $\times 10^{-4}$ m). Infrared continuum radiation from dust falls in between the stellar radiation and the cosmic microwave background, with $\lambda \sim$ 30–300 μm, corresponding to dust temperatures in the range $T_{\rm dust} \sim$ 10–100 K.

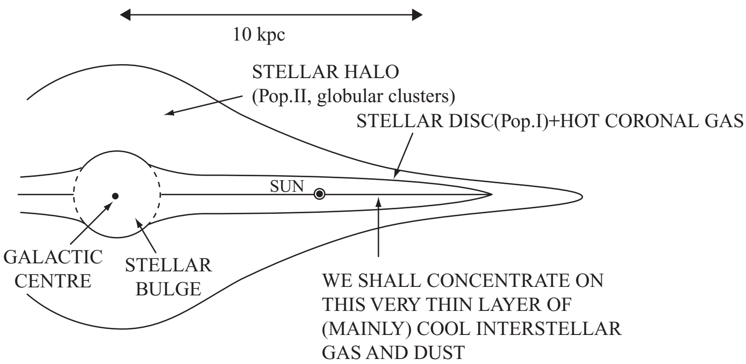

Fig. 1.4. Schematic view of the Galaxy from the side, showing the principal stellar components, the locations of the Galactic Centre and the Sun, and the location of the interstellar medium.

The gravitational field on large scales is the gravitational field of the Galaxy, which we believe to be dominated by stars and 'dark matter' (the precise nature of which is uncertain). Locally, the gravity of the interstellar gas may become important; this is particularly true in GMCs where the gas density is high and self-gravity may cause the interstellar gas to condense into new stars.

The interstellar medium is in a continual state of dynamical and chemical change, hence it is not in mechanical or chemical equilibrium. Additionally, the mean free paths for photons and cosmic rays in the interstellar medium are normally much longer than the typical distances over which the physical conditions change. Hence the interstellar medium is also not in local thermodynamic equilibrium (LTE). However, the mean free path of gas particles is short, so the gas is normally in thermal equilibrium, in the sense that there is a well-defined gas-kinetic temperature, T, characterising the distribution of gas particle velocities.

1.4 The distribution of the stars

We can identify several components of our Galaxy (the Milky Way).

If we could view it from the side we would see a distribution such as is shown in Figure 1.4. We start by describing how the stars are distributed. The oldest, most metal-poor stars† are known as Population II stars, and are distributed in a spheroidal (almost spherical) halo, which is at least 30 kpc across.‡ This component includes the globular clusters.

There is also a roughly spherical bulge near the centre of the Milky Way, about 3 kpc in radius. The stars in the bulge are also old, but they have higher metallicity than the halo stars. Finally the youngest, most

† In astronomy, a metal is defined as any element other than hydrogen or helium. Any star with significantly fewer metals than the Sun is referred to as metal-poor. Any star with significantly more metals than the Sun is referred to as metal-rich.

‡ 1 kpc $= 1000$ pc $= 3 \times 10^{19}$ m.

Fig. 1.5. The Orion Nebula, containing the Trapezium Cluster, as seen by the Hubble Space Telescope.

metal-rich stars (the Population I stars) are concentrated in a disc close to the midplane of the Milky Way. This stellar disc is at least 30 kpc across, and about 800 pc thick in the solar neighbourhood, although the more massive stars appear to be concentrated in a central layer only about 200 pc thick.

The Sun appears to be close to the midplane of the Milky Way (perhaps 10–20 pc above it), and about 8 kpc from the centre of the Milky Way. Most of the interstellar gas and dust is confined to an extremely thin layer, about 200 pc thick, inside the stellar disc. The radial extent of the gas disc is at least 20 kpc – i.e. at least 100 times its thickness – so the interstellar medium is like a very thin pancake.

There is some interstellar gas further from the midplane of the Milky Way. This is the coronal gas, very hot rarefied gas which is presumed to have escaped from the interiors of old supernova remnants, and which extends to about 3 kpc above and below the midplane. However, the total mass of the coronal gas is very small compared with the cooler interstellar gas near the Galactic midplane.

Stars and star clusters are observed to have a wide range of ages, from less than a million years for young stars like the Trapezium Cluster in Orion (see Figure 1.5), up to of order 10^{10} years for the oldest globular clusters. This implies that star formation has occurred since very early times – the age of the Universe is estimated to be $\sim 1.3 \pm 0.3 \times 10^{10}$ years – and is still continuing.

Star formation converts diffuse interstellar gas clouds into star clusters. The rate at which interstellar gas is converted into stars has a profound effect on the overall dynamics of a galaxy, and hence on its

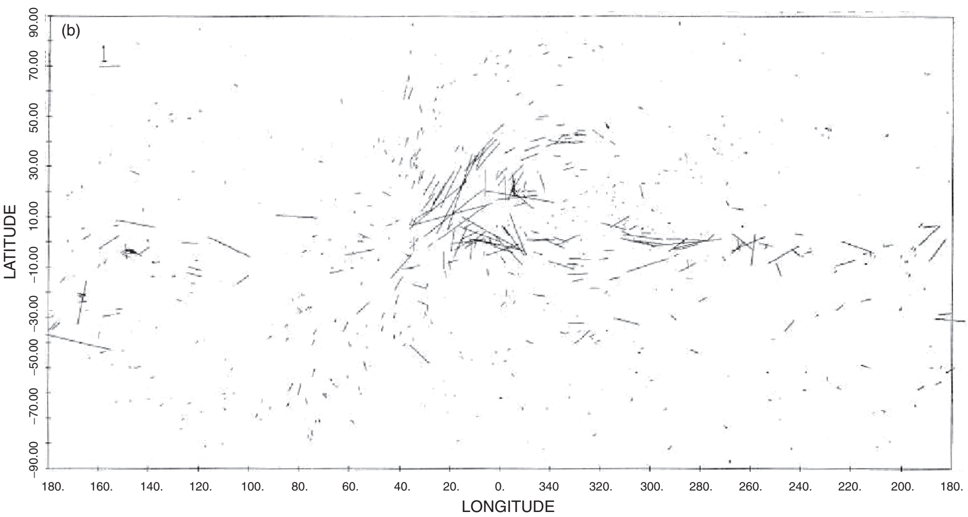

Fig. 1.6. The polarisation of starlight projected onto Galactic coordinates. This is believed to be tracing the large-scale magnetic field in the interstellar medium of our Galaxy.

formation, structure and evolution. Thus, the star formation within a galaxy is crucial to that galaxy's overall evolution.

1.5 The magnetic field

The origin of the interstellar magnetic field is uncertain. Usually it is assumed to be generated by a galactic-scale dynamo, and amplified locally by gas-dynamical processes, but the details of the underlying mechanisms are not well understood. The strength of the magnetic field is typically $|\mathbf{B}| \sim 3\,\mu\mathrm{gauss}^{\dagger}$ and so its energy density is $|\mathbf{B}|^2/4\pi \sim 10^{-13}\ \mathrm{J\,m^{-3}}$, which is comparable with the energy densities of the gas, of the radiation field, and of the cosmic rays. Therefore the magnetic field is likely to play an important role in determining the structure and evolution of the interstellar medium.

A number of observations lead astronomers to believe that magnetic fields are ubiquitous in the ISM. This is not a totally surprising result given that magnetic fields are caused by moving charges and that the ISM appears to be in constant motion. One of the main pieces of evidence we have for magnetic fields comes from observations of the polarisation of starlight. When such polarisation measurements were first carried out, astronomers were surprised to find that the light from stars was partially plane polarised in a pattern across the sky (see Figure 1.6). Furthermore, this pattern appears to trace out large-scale structures along the plane of the Galaxy.

If the polarisation were a property purely of the stars themselves it would be very difficult to explain this apparent structure. However,

† 1 gauss $= 10^{-4}$ tesla, so 1 μgauss $= 10^{-10}$ tesla.

astronomers believe that the polarisation is not a property of the stars themselves but of the ISM between the stars and us, and that it is caused by an interstellar magnetic field.

To understand how this mechanism works we must realise that interstellar dust grains are not spherically symmetrical, but typically have very complex shapes. We usually try to simplify our consideration of their shapes by thinking of them either as needle-like cylinders or prolate spheroids (rugby ball or American football shaped). In this way we can think of their asymmetry simply in terms of a 'long axis' and a 'short axis'.

Put simply, the long axis of a grain extinguishes the background starlight more efficiently than the short axis. Hence if a large ensemble of non-spherical dust grains is aligned preferentially in one direction, the background starlight is preferentially extinguished along one axis, causing the transmitted light to be partially plane polarised.

The manner of the alignment is still a matter of debate, and many theories have been put forward to explain it. However, most theories predict that the grains spin around the direction of the magnetic field, with their long axis perpendicular to the field. The original version of this mechanism is known as the Davis–Greenstein effect. The basis of this effect is as follows:

The dust grains in a molecular cloud are in random motion and undergo frequent collisions. These collisions set the grains spinning. The grains are typically composed of silicate material which is paramagnetic in nature. If the cloud is threaded by a magnetic field, then the presence of this external field causes an induced internal field within the paramagnetic dust grain material, whose strength depends on the magnetic susceptibility of the material. Normally these two fields would be parallel to one another, but because the grain is spinning, the internal field cannot respond quickly enough, so it always 'lags' behind the external field direction. This causes a net torque which tends to cause the grain to spin with its long axis perpendicular to the external magnetic field direction. Further collisions of course serve to misalign the grains once more.

This mechanism may not be exactly correct in practice as it requires a field strength about an order of magnitude higher than that which is measured (see below). However, a number of alternatives have been proposed. One plausible mechanism is super-paramagnetic alignment, which requires inclusions of a ferromagnetic substance within the dust grain material to increase the alignment efficiency. Another possibility is suprathermal alignment in which molecules being ejected from the grain surface help to spin up the grain and hence shorten the alignment time-scale and increase the mechanism's efficiency. Whichever of these mechanisms ultimately proves correct, all agree that the grains become preferentially aligned perpendicular to the magnetic field direction. Hence the background starlight is preferentially extinguished in this direction. The transmitted light is therefore partially plane polarised parallel to the magnetic field direction.

Linear polarisation allows us to measure the orientation of the magnetic field in the plane of the sky. To measure the strength of the magnetic field we use the Zeeman effect, which relies on the splitting of degenerate atomic or molecular energy levels in the presence of a magnetic field. The amount of the splitting is proportional to the field strength, thus allowing the field strength to be measured.

Consider the simplest case of a hydrogen atom. The electronic ground state (s state) has principal quantum number $n = 1$ and angular momentum quantum number $l = 0$. The first excited state (p state) has $n = 2$ and $l = 1$. The magnetic quantum number m_l must obey the relation $|m_l| \leq l$. Hence the allowed values for m_l in the p state are 0, $+1$ and -1, and the state is said to have triple degeneracy. In the presence of a magnetic field the degeneracy is lifted and the p state becomes a triplet. Hence the spectral line of the transition between p and s states becomes a triplet. Each level is shifted in energy E by an amount

$$E = \mu_{\mathrm{B}} m_l B, \tag{1.1}$$

where the constant μ_{B} is known as the Bohr magneton and has a value of 9.27×10^{-24} J/T, and B is the magnitude of the magnetic field strength along the line of sight to the observer.

Hence by observing such a multiple line whose values of m_l are known, we can measure the magnetic field strength in one direction. Typical values that have been measured in the ISM are $\sim 10^{-10}$–10^{-9} T. These values are so small that the typical splitting is too small for most spectrometers to measure. However, the two levels have opposite circular polarisations and can therefore be split by a polarimeter, which is sensitive to circular polarisation. This means therefore that only the field strength along the line of sight is measured and various geometric assumptions have to be made to infer the three-dimensional magnetic field configuration. In Chapter 4 we will discuss the effects that magnetic fields have on the dynamics of the ISM.

1.6 Star formation in a galactic context

Star formation converts diffuse interstellar gas clouds, which undergo highly dissipative collisions and are therefore very inelastic, into star clusters, which are effectively collisionless and therefore much more elastic. What this means is that when two clouds of interstellar gas collide, the mean free path for collisions between the individual gas