HOW DO STARS FORM?

11.1 Introduction

The last decades have seen great strides forward in our understanding of the processes that govern the formation and evolution of stars. Progress has been made on both observational and theoretical fronts, driven by improvements in instrumentation (primarily infrared) and in computational resources. Indeed, it is now possible to 'follow' the process of star formation in sophisticated numerical simulations of increasing spatial resolution.

Figure 11.1: At a distance of 7.4 Mpc from the Milky Way, M101 (the Pinwheel galaxy) is a 'grand-design' spiral galaxy seen nearly face on ($i = 18°$). This picture is a composite of three images, each taken with one of the NASA's 'Great Observatories'. Visible light (yellow) captured by the Hubble Space Telescope traces the general stellar population. Infrared light (red) recorded with the Spitzer Space Telescope is emitted by hot dust in molecular clouds. X-ray light (blue) seen by the Chandra X-ray Observatory marks the locations of high-energy gas in the vicinity of young, massive stars. Note the close correspondence between X-ray emitting regions of recent star formation and dust clouds.

Despite such advances, star formation is still considered to be a poorly-understood problem in astrophysics. The reason is that we do not yet have a *predictive* theory of star formation; for example, given a set of initial conditions, we do not yet have the means to predict reliably important properties, such as the star formation efficiency (that is, what fraction of the gas is turned into stars), or the Initial Mass Function which describes the relative numbers of stars of different masses. On the other hand, once a star has formed and nuclear burning starts, all the uncertain details of the star formation process are no longer relevant to its evolution. In this lecture, we explore some of the basic ideas concerning the way stars form.

Observationally, we know that there is a strong spatial association between clusters of newly formed stars and interstellar gas. Both are concentrated in spiral arms in external galaxies (see Figure 11.1 for an example); in the Milky Way, regions of recent star formation such as the Great Carina Nebula (Figure 11.2) are also some of most spectacular visualisations of the interstellar medium.

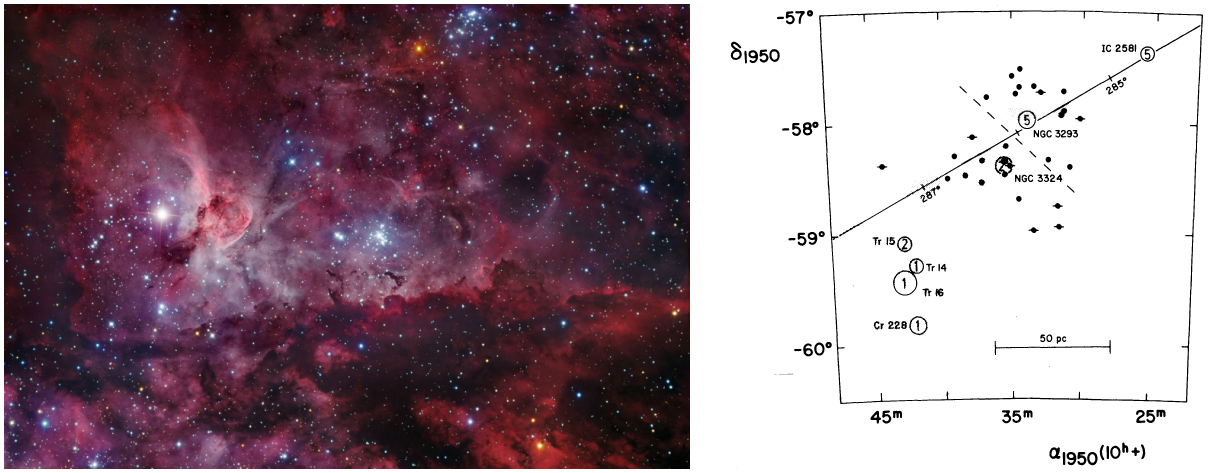

Figure 11.2: Located $\sim 2.5\,\mathrm{kpc}$ from the Sun, the Giant Carina Nebula is one of the Galaxy's largest star-forming regions. It is rich is clusters of newly formed O and B stars, harbouring a significant fraction of the most massive stars known in the Milky Way, including several O3 V stars and the LBV η Carinae. The picture on the left highlights the interstellar matter in the nebula, some glowing in the light of newly-formed stars, while other regions are dense clouds of molecules and dust. The sketch on the right, reproduced from Turner et al. (1980), shows the locations and ages (in millions of years) of the most prominent star clusters. Star formation is propagating from the north-west to the south-east of the region at a speed of $\sim 30\,\mathrm{km\ s^{-1}}$, attesting to the interaction between massive stars and the ambient interstellar medium. The continuous line is the Galactic equator with intervals of Galactic longitude marked.

The interstellar medium is a complex environment, worthy of its own lecture course. For our present purposes, suffice to say that among the various components of the ISM, it is the Giant Molecular Clouds that are associated with sites of star formation. Images taken at near-infrared wavelengths ($\lambda = 1$–$2.2\,\mu$m), show very young clusters and protostellar objects still embedded deep into such clouds, before the ultraviolet radiation emitted by the most massive stars evaporates the cloud away.

Giant molecular clouds, of which there are thousands in the Milky Way, are enormous complexes of dust and gas, sufficiently dense to be self-shielding from the diffuse interstellar ultraviolet radiation field. Typical parameters are temperatures $T \sim 15\,\mathrm{K}$, volume densities $n = 100$–$300\,\mathrm{cm}^{-3}$, and masses $M = 10^5$–$10^6 M_\odot$. The low temperatures and high densities (and the presence of solid particles known as interstellar dust which acts as a catalyst) favour the formation of molecules; thus most of hydrogen in GMCs is in the form of H_2 rather than H I. Dust makes up about 1% of the material, and it is the dust that renders the clouds very opaque at visual wavelengths.

With typical sizes of 10s of parsecs, GMCs show considerable structure on a variety of scales. On scales of 10 pc, $M \sim 10^4 M_\odot$, densities can be $n \sim 500\,\mathrm{cm}^{-3}$ and the extinction at visible wavelengths is $A_V \sim 5$ mag; such regions are sometimes referred to as Dark Cloud Complexes. Smaller clumps on scales of 1–2 pc with $M \sim 30 M_\odot$ can have $n \sim 1000\,\mathrm{cm}^{-3}$, $A_V \sim 10$ mag, and $T \sim 10\,\mathrm{K}$. Dense Cores on scales of 0.1 pc, can have $M \sim 10 M_\odot$, $n \sim 10\,000\,\mathrm{cm}^{-3}$, and $A_V > 10$ mag. Finally, in some localised regions of GMCs observations have revealed 'Hot Cores' with $T = 100$–$300\,\mathrm{K}$, and extreme values of density and extinction, as high as $n \sim 10^9\,\mathrm{cm}^{-3}$ and $A_V \sim 50$; masses can be in the range 10–$3000 M_\odot$. It is in these hot cores that near-IR imaging[1] has revealed the presence of embedded young O and B-type stars, strongly suggesting a causal connection. As we shall see presently, giant molecular clouds are indeed unstable to gravitational collapse.

[1]Recalling that $A_V = 10$ mag corresponds to a dimming of the visible light by a factor of 10 000, it can be easily appreciated that these dense regions are opaque to light of visible wavelengths. Longer wavelength light, in the near-IR regime, suffers considerably less extinction.

11.2 Brief Overview

It is useful to distinguish six stages in the star formation process:

1. The initial free-fall collapse of the parent interstellar cloud.
2. Cloud fragmentation, leading to a range of stellar masses.
3. Formation of a protostellar core. The star appears on the H-R diagram.
4. Accretion of the surrounding gas, generally through an accretion disk.
5. Dissociation of molecules and ionisation of H and He.
6. Pre-main sequence phase.
7. Star formation is considered to be completed once the star appears on the "Zero Age Main Sequence" (ZAMS).

In the following sections we shall consider these different stages in turn.

11.3 The Jeans Criterion for Gravitational Collapse

The condition for a cloud of interstellar gas to collapse and form stars was first formulated by James Jean in 1902 while working at Trinity College, Cambridge. His treatment considers only gravitation and thermodynamics and ignores other important effects such as rotation, turbulence and magnetic fields. Nevertheless, it provides a useful insight into the development of protostars.

As we saw in Lecture 7.1, the condition for equilibrium of a stable, gravitationally bound system is given by the virial theorem:

$$2K + U = 0 \tag{11.1}$$

where K is the kinetic energy and U is the potential energy. Consider now what happens if such a system is perturbed from equilibrium. If $2K > |U|$, the force due to gas pressure will dominate the force of gravity

and the cloud will expand. On the other hand, if $2K < |U|$, the reverse will be true and the cloud will collapse under the force of gravity.

We showed in Lecture 7.1 (eqs. 7.1–7.8) that the gravitational potential energy can be written as:

$$U \simeq -\frac{3}{5}\frac{GM_c^2}{R_c}\,, \tag{11.2}$$

where M_c and R_c are, respectively, the mass and the radius of the cloud under consideration.

We also saw (eqs. 7.9–7.11) that the average kinetic energy per particle is

$$K = \frac{3}{2}kT \tag{11.3}$$

where k is Boltmann constant. Thus, the total internal kinetic energy of the cloud is just:

$$K = \frac{3}{2}\mathcal{N}kT \tag{11.4}$$

where $\mathcal{N}$ is the total number of particles. We can write $\mathcal{N}$ in terms of the mass and the mean molecular weight (Lecture 9.2.1):

$$\mathcal{N} = \frac{M_c}{\mu m_H} \tag{11.5}$$

We can therefore rewrite the condition for gravitational collapse ($2K < |U|$) as:

$$\frac{3M_c kT}{\mu m_H} < \frac{3}{5}\frac{GM_c^2}{R_c} \tag{11.6}$$

We can eliminate the radius R_c from the above equation using:

$$R_c = \left(\frac{3}{4}\frac{M_c}{\pi\rho_0}\right)^{1/3} \tag{11.7}$$

where ρ_0 is the initial density of the cloud prior to collapse, with the assumption that the cloud is a sphere of constant density.

Substituting the above expression into eq. 11.6, we obtain the important concept of the *Jeans mass*:

$$\boxed{M_J \simeq \left(\frac{5kT}{G\mu m_H}\right)^{3/2}\left(\frac{3}{4\pi\rho_0}\right)^{1/2}\,.} \tag{11.8}$$

If the mass of a cloud exceeds the Jeans mass, the cloud will be unstable against gravitational collapse.

Note that for a given chemical composition (i.e. for a given μ), the Jeans mass depends only on temperature and density: the lower the temperature and the higher the density, the smaller the cloud mass which is unstable against gravitational collapse. Note also that this is an oversemplification of the real situation, not only because we have ignored rotation, internal macroscopic velocity gradients and magnetic fields, but also because we have neglected any external pressure on the cloud (such as the external pressure of an encompassing GMC on an embedded dense core).

The same criterion for gravitational collapse can also be expressed in terms of the Jeans length:

$$\boxed{R_J \simeq \left(\frac{15kT}{4\pi G\mu m_H \rho_0}\right)^{1/2}}, \tag{11.9}$$

where the condition for gravitational collapse is $R_c > R_J$, or the Jeans density:

$$\boxed{\rho_J \simeq \frac{3}{4\pi M_c^2}\left(\frac{5kT}{G\mu m_H}\right)^3}, \tag{11.10}$$

where the condition for gravitational collapse is $\rho_c > \rho_J$.

Let us consider the Jeans mass of some interstellar structures. In diffuse hydrogen clouds, typical values of temperature and density are $T \sim 100\,\mathrm{K}$ and $n_{\mathrm{HI}} \sim 10^7\,\mathrm{m}^{-3}$. Thus, $\rho_0 = n_{\mathrm{HI}}\, m_H \sim 1 \times 10^7 \cdot 1.7 \times 10^{-27} \sim 1.7 \times 10^{-20}\,\mathrm{kg\ m}^{-3}$. With the gas fully neutral, $\mu = 1$. Entering these values in eq. 11.8, we have (using S.I. units throughout):

$$M_J \simeq \left(\frac{5 \cdot 1.4 \times 10^{-23} \cdot 100}{6.7 \times 10^{-11} \cdot 1 \cdot 1.7 \times 10^{-27}}\right)^{3/2} \left(\frac{3}{4 \cdot 3.14 \cdot 1.7 \times 10^{-20}}\right)^{1/2} \tag{11.11}$$

or

$$M_J \simeq \left(6.1 \times 10^{16}\right)^{3/2} \left(1.4 \times 10^{19}\right)^{1/2} \tag{11.12}$$

$$M_J \simeq 5.7 \times 10^{34}\,\mathrm{kg} \sim 30\,000 M_\odot \tag{11.13}$$

This value of the Jeans mass is two orders of magnitude higher than the typical mass of diffuse interstellar clouds, which are therefore very stable against gravitational collapse.

On the other hand, in the dense core of a giant molecular cloud, typical values are $T \sim 10\,\mathrm{K}$, and $n_{H_2} \sim 10^{10}\,\mathrm{m}^{-3}$. The density is thus $\rho_0 \sim 2n_{H_2}m_H \sim 2 \times 10^{10} \cdot 1.7 \times 10^{-27} \sim 3.4 \times 10^{-17}\,\mathrm{kg\ m}^{-3}$, a factor of 2000 greater than in a diffuse hydrogen cloud. With $\mu = 2$, the Jeans mass will therefore be lower by a factor of $(10 \cdot 2)^{3/2} \cdot \sqrt{2000} = 4000$, i.e. $M_J \simeq 7.5 M_\odot$.

Now we see that, with typical masses $M \sim 10 M_\odot$, the dense cores of GMCs are indeed unstable to gravitational collapse, consistent with being the sites of star formation. While a GMC may originally be in pressure equilibrium with the surrounding interstellar medium, a small perturbation can initiate its contraction and trigger gravitational collapse. Such a perturbation can be provided by cloud-cloud collisions, or by a passing interstellar shock wave originating from a nearby region of star formation, where some of the most massive stars have already exploded as supernovae. This can lead to a 'forest fire' type of star formation, as has been suggested for the Carina nebula (see Figure 11.2), whereby star formation propagates from one side of a giant molecular cloud complex. Galaxy interactions are another type of process that can trigger star formation by tipping GMCs past the Jeans mass limit.

11.4 Free-fall Timescale

Although the Jeans criterion provides the necessary condition for the onset of collapse of a gas cloud, such collapse involves the release of gravitational energy. If this energy were converted directly to thermal energy, the temperature would rise and, according to eq. 11.8, the Jeans mass would increase, halting the collapse. However, in the early stages of collapse, the cloud is transparent to far-infrared radiation and can cool efficiently by converting kinetic energy of its molecules and atoms into infrared photons which can escape the cloud. Thus, the early stages of the collapse are isothermal and the cloud is essentially in free-fall collapse.

We can obtain an estimate of the free-fall timescale as follows. We begin

by writing the equation of gravitational acceleration:

$$\frac{d^2r}{dt^2} = -G\frac{M_r}{r^2} \tag{11.14}$$

where as usual M_r denotes the mass enclosed within radius r (assuming spherical symmetry). Writing the mass in terms of the initial density and radius, and multiplying both sides by the velocity at the surface of the spherical cloud, we have:

$$\frac{dr}{dt}\,\frac{d^2r}{dt^2} = -\left(\frac{4\pi}{3}G\rho_0 r_0^3\right)\,\frac{1}{r^2}\,\frac{dr}{dt} \tag{11.15}$$

which can be integrated with respect to time to give:

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 = \left(\frac{4\pi}{3}G\rho_0 r_0^3\right)\,\frac{1}{r} + C_1\,. \tag{11.16}$$

The integration constant C_1 can be evaluated with the boundary condition that the infall velocity be zero at the onset of collapse, that is $dr/dt = 0$ when $r = r_0$. This gives:

$$C_1 = -\frac{4\pi}{3}G\rho_0 r_0^2 \quad . \tag{11.17}$$

Substituting 11.17 into 11.16 and solving for the velocity at the surface, we have:

$$\frac{dr}{dt} = -\left[\frac{8\pi}{3}G\rho_0 r_0^2\left(\frac{r_0}{r} - 1\right)\right]^{1/2}\,, \tag{11.18}$$

where we have chosen the negative root because the cloud is collapsing. In order to obtain an expression for the position of the spherical cloud surface as a function of time we need to integrate 11.18. With the substitutions:

$$\theta \equiv \frac{r}{r_0}$$

and

$$\chi \equiv \left(\frac{8\pi}{3}G\rho_0\right)^{1/2}$$

eq. 11.18 can be re-written as:

$$\frac{d\theta}{dt} = -\chi\left(\frac{1}{\theta} - 1\right)^{1/2}\,. \tag{11.19}$$

With the further substitution:

$$\theta \equiv \cos^2 \xi \,,$$

eq. 11.19 can in turn be re-written as

$$\cos^2 \xi \frac{d\xi}{dt} = \frac{\chi}{2} \tag{11.20}$$

which can now be integrated with respect to time to give:

$$\frac{\xi}{2} + \frac{1}{4} \sin 2\xi = \frac{\chi}{2} t + C_2 \tag{11.21}$$

Again, we can evaluate the constant of integration C_2 by considering that at $t = 0$, $r = r_0$. Hence, $\theta = 1$ and $\xi = 0$. Therefore, $C_2 = 0$. We thus arrive at the equation of motion for the gravitational collapse of the cloud:

$$\xi + \frac{1}{2} \sin 2\xi = \chi t \tag{11.22}$$

We define the free-fall timescale as the time taken by a cloud in free-fall to collapse from $r = r_0$ to $r = 0$. (In reality, of course, this final condition is never reached, but the concept is still valid so long as $r_{\rm final} \ll r_0$). When $r = 0$, $\theta = 0$ and $\xi = \pi/2$. Then, from 11.22 we have:

$$t_{\rm ff} = \frac{\pi}{2\chi} \tag{11.23}$$

or

$$t_{\rm ff} = \left(\frac{3\pi}{32} \frac{1}{G\rho_0} \right)^{1/2} \,. \tag{11.24}$$

Note that the free-fall timescale is independent of the initial radius of the sphere, and depends only on the initial density ρ_0. Thus, in a spherical molecular cloud of uniform density (admittedly a rather implausible simplification!), all parts of the cloud will take the same length of time to collapse and the density will increase at the same rate everywhere within the cloud. This behaviour is known as *homologous collapse*.

11.5 Cloud Fragmentation

From the above treatment, one would be justified in concluding that the entire mass of GMC exceeding the Jeans limit would collapse to form a

single star. This is clearly not the case: most early-type (i.e. young) stars are found in clusters, and maybe as many as three quarters of all stars are in binary or multiple systems.

During the free-fall collapse phase, the density within the cloud increases by many orders of magnitude. If the temperature remains approximately constant, then the Jeans mass criterion (eq.11.8) implies that the mass limit for instability decreases dramatically. Any initial density inhomogeneities which may have been present within the cloud will cause individual regions within the GMC to cross the instability threshold independently and collapse locally. This could lead to the formation of a large number of smaller objects.

This is still only part of the story, however. The fragmentation process must stop at some point because: (i) the process of star formation is not 100% efficient—it is actually rather inefficient, with only about 1% of the gas mass of the parent cloud being turned into stars; and (ii) the most common type of star in the Galaxy has a mass of $\sim 1 M_\odot$, and lower mass stars are actually *less* common. Something must intervene to limit the fragmentation process.

Fragmentation stops when the assumption of isothermal contraction breaks down. The increasing density of the collapsing cloud fragment eventually renders the gas opaque even to infrared photons. As a result, radiation is trapped within the central part of the collapsing cloud, leading to heating and an increase in gas pressure. When radiation can no longer escape the cloud, the collapse turns from isothermal to adiabatic. (Of course, in reality the collapse is never totally isothermal nor adiabatic, but somewhere between these two limits. As is often the case, we can gain insight into the relevant physical processes by considering the two limiting cases.)

We can appreciate the transition from isothermal to adiabatic collapse by recalling (Lecture 8) the adiabatic relationship between pressure and density:

$$P = K\rho^\gamma$$

where γ is the ratio of specific heats and K is a constant. Combined with the ideal gas law

$$P = \frac{\rho k T}{\mu m_H} ,$$

it implies an adiabatic relationship between temperature and density:

$$T = K'\rho^{\gamma-1}\,. \tag{11.25}$$

Substituting 11.25 into the expression for the Jeans mass (eq. 11.8), we have:

$$M_J \propto \rho^{(3\gamma-4)/2}\,. \tag{11.26}$$

For atomic hydrogen[2] $\gamma = 5/3$, giving $M_J \propto \rho^{1/2}$; in other words, the Jeans mass *increases* as the density increases (for a perfectly adiabatic collapse). This behaviour results in a minimum fragment mass determined by the transition from a predominantly isothermal to a predominantly adiabatic collapse. In turn, the transition is driven by the opacity of the interstellar mix of molecules, gas and dust to infrared radiation.

We can have a stab at estimating the lower mass limit of the fragmentation process as follows. As we have already seen in lecture 7.11, the energy released during the collapse of a protostellar cloud is half its potential energy:

$$\Delta E_g \simeq \frac{3}{10}\frac{GM_J^2}{R_J} \tag{11.27}$$

for a spherical cloud satisfying the Jeans instability criterion. Averaged over the free-fall time, the luminosity produced in the collapse is:

$$L_{\rm ff} = \frac{\Delta E_g}{t_{\rm ff}} \simeq \frac{3}{10}\frac{GM_J^2}{R_J}\cdot\left(\frac{3\pi}{32}\frac{1}{G\rho_0}\right)^{-1/2} \tag{11.28}$$

Using $\rho_0 = M_J/(\frac{4}{3}\pi R_J^3)$, we have:

$$L_{\rm ff} \sim G^{3/2}\left(\frac{M_J}{R_J}\right)^{5/2}\,. \tag{11.29}$$

Now, if the cloud were optically thick and in thermodynamic equilibrium, this energy would be emitted as blackbody radiation, with its luminosity given by the familiar expression:

$$L_{\rm bb} = 4\pi R^2\sigma T^4\,.$$

[2]At the low temperatures of interstellar clouds, molecular hydrogen behaves like a monoatomic gas. This observation was extremely puzzling in the early part of the twentieth century and engaged the minds of some of the most prominent physicists of the time. It took the development of the quantum theory to show that diatomic hydrogen, with its tiny rotational inertia, requires a large amount of energy to excite its first excited molecular rotation quantum state. Since it can not get that amount of energy at low temperatures, it acts like a monoatomic gas.

However, the collapsing cloud is not in thermodynamic equilibrium, so we introduce an efficiency factor e ($0 < e < 1$) into the above equation to give:

$$L_{\rm rad} = 4\pi R^2 e\sigma T^4 \,. \tag{11.30}$$

Equating the two expressions for the cloud luminosity, we have:

$$M_J^{5/2} = \frac{4\pi}{G^{3/2}} R_J^{9/2} e\sigma T^4 \,. \tag{11.31}$$

Expressing the radius in terms of the mass and density (eq. 11.7), and then the density in terms of the Jeans mass (eq. 11.8), we finally arrive at an expression for the minimum obtainable Jeans mass corresponding to when adiabatic effects become important:

$$M_{J_{\rm min}} = 0.03 \left(\frac{T^{1/4}}{e^{1/2}\mu^{9/4}} \right) M_\odot \tag{11.32}$$

where T is in kelvins. With $\mu \sim 1$, $e \sim 0.1$, and $T \sim 1000\,{\rm K}$ at the time when adiabatic effects may start to become significant, we have $M_{J_{\rm min}} \sim 0.5 M_\odot$; that is, fragmentation ceases when individual fragments are approximately solar mass objects. Note that the parameters which determine $M_{J_{\rm min}}$ in eq. 11.32 are all to relatively low powers. Thus the conclusion that fragmentation does not continue much beyond masses comparable to a solar mass is not highly dependent on the values of e and T. For example, if $e = 1$ (fully adiabatic), then $M_{J_{\rm min}} \sim 0.2 M_\odot$.

11.6 Protostars

Once the density of a collapsing fragment has increased sufficiently for the gas to become opaque to infrared photons, radiation is trapped within the central part of the cloud, leading to heating and an increase in gas pressure. As a result the cloud core is nearly in hydrostatic equilibrium and the dynamical collapse is slowed to a quasistatic contraction. At this stage we may start to speak of a protostar.

The mass of such a 'protostar' is still only a small fraction of the mass it will have once it reaches the Main Sequence. The surrounding gas continues to free-fall onto the protostellar core, so that the next phase is dominated by

accretion. In general, the infalling gas will form an accretion disk around the protostar, reflecting the fact that the contracting gas cloud has a net angular momentum. Such accretion disks are commonly seen around very young stars.

The accretion of gas generates gravitational energy, part of which goes into further heating of the core and part of which is radiated away, providing the luminosity of the protostar:

$$L \sim L_{\rm acc} = \frac{1}{2}\frac{GM\dot{M}}{R} \tag{11.33}$$

where $\dot{M}$ is the mass accretion rate. The core heats up almost adiabatically since the accretion timescale $t_{\rm ff}$ is much smaller than the thermal timescale, which is just the Kelvin-Helmholtz timescale, $t_{\rm KH}$, we have already encountered in Lecture 7.1.

11.6.1 Dissociation and Ionisation

When the temperature of the core reaches $T \sim 2000\,{\rm K}$, the average particle energy is comparable to the dissociation energy of molecular hydrogen. The energy produced by contraction is now absorbed by this process, rather than providing the pressure gradient necessary to maintain hydrostatic equilibrium. As a result, the core becomes dynamically unstable, and a second collapse occurs during which the gravitational energy released is absorbed by the dissociating molecules without a significant rise in temperature. When H_2 is completely dissociated into atomic hydrogen, hydrostatic equilibrium is restored and the temperature rises again. Somewhat later, further dynamical collapse phases follow when first H and then He are ionised at temperatures $T \sim 10^4\,{\rm K}$.

When ionization of the protostar is complete, it settles back into hydrostatic equilibrium at a much reduced radius. We can estimate the radius R_p of a protostar after the dynamical collapse phase is over by equating the change in potential energy to the sum of the energies of: (i) dissociation of molecular hydrogen ($\chi_{\rm H_2} = 4.48\,{\rm eV}$ per H_2 molecule), (ii) ionisation of atomic hydrogen ($\chi_{\rm HI} = 13.6\,{\rm eV}$ per H atom), and (iii) ionisation of helium ($\chi_{\rm He} = 79\,{\rm eV}$ per He atom: $24.6\,{\rm eV}$ for $\rm He^0 \rightarrow He^+$ and $54.4\,{\rm eV}$ for $\rm He^+ \rightarrow$

He^{++}). Thus we have:

$$\Delta E_g \simeq \frac{3}{10}\frac{GM^2}{R_p} \approx \frac{M}{m_H}\left(\frac{X}{2}\chi_{H_2} + X\chi_H + \frac{Y}{4}\chi_{He}\right) \equiv \frac{M}{m_H}\chi\,, \qquad (11.34)$$

where we have taken the collapse to start from infinity because $R_p \ll R_{\rm init}$. With $X = 0.74$ and $Y = 1 - X$ (ignoring the metals), $\chi = 16.9\,{\rm eV}$ per baryon. Therefore we have:

$$R_p \approx \frac{3}{10}\frac{GMm_H}{\chi} \approx 35R_\odot\left(\frac{M}{M_\odot}\right)\,. \qquad (11.35)$$

We can use the virial theorem to estimate the average temperature of the protostar, $\langle T_p\rangle$. Following the same steps as in Lecture 7.1.1 (eqs. 7.9–7.13), we find:

$$\langle T_p\rangle = \frac{1}{5}\frac{\mu m_H}{k}\frac{GM}{R_p} = \frac{2}{3}\frac{\mu}{k}\chi \approx 8\times10^4\,{\rm K}\,, \qquad (11.36)$$

for $\mu \simeq 0.6$ appropriate to a fully ionised plasma (Lecture 9.2.1). There are two things of note about eq. 11.36. First, the temperature is independent of the mass of the protostar. Second, the temperature reached once the protostar settles back into hydrostatic equilibrium (following H_2 dissociation and He and He ionisation) is still much lower than the temperature necessary to ignite nuclear reactions. At these 'low' temperatures, the opacity is very high and is dominated by the H^- ion (Lecture 5.4.1). Under these circumstances, radiative energy transport is very inefficient, making the protostar *convective* throughout, except for the outermost layer (photosphere). Thus a new star is chemically homogeneous. This will change as the star ages.

11.7 The Hayashi Track and the Final Approach to the Main Sequence

Fully convective stars of a given mass occupy an almost vertical line in the H-R diagram. The line is called the *Hayashi line* from the Japanese astrophysicist who worked on stellar evolution during the pre-main sequence stages in the early 1960s. A protostar of a given mass moves along its Hayashi line as it approaches the main sequence (see Figure 11.3).

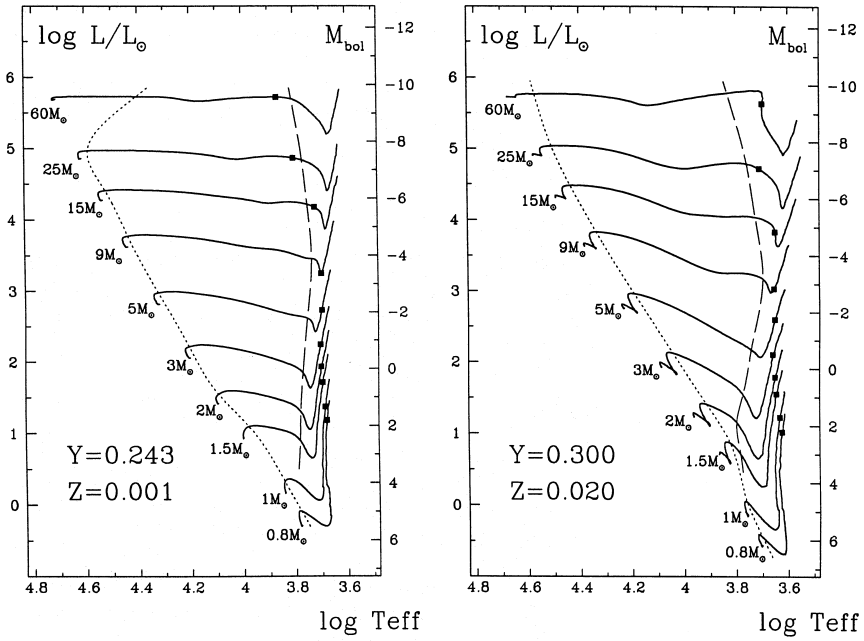

Figure 11.3: Theoretical pre-main sequence evolutionary tracks computed for stars of various masses for two different metallicities, as indicated. A pre-main sequence star evolves along its track from right to left. The black square on each track indicates the onset of D burning. (Reproduced from Bernasconi & Maeder 1996, A&A, 307, 829).

The Hayashi track is actually a boundary. It represents the minimum effective temperature for a star in hydrostatic equilibrium. The right of the Hayashi track, there is no mechanism that can adequately transport the luminosity out of the star at such low effective temperatures; hence no stable star can exist there and this region is sometimes referred to as the *forbidden region* of the H-R diagram. On the other hand, stars to the left of the Hayashi line (at higher $T_{\rm eff}$) cannot be fully convective but must have some portion of their interior in radiative equilibrium.

As a newly formed star emerges from the dynamical collapse phase, it settles on the Hayashi line appropriate for its mass, with a radius roughly given by eq. 11.35. From this moment on we speak of the pre-main sequence phase of evolution. The pre-main sequence (PMS) star radiates at a luminosity determined by its radius on the Hayashi line. Since it is still too cool for nuclear burning, the energy source for its luminosity is gravitational contraction. As dictated by the virial theorem, this leads to an increase of its internal temperature. As long as the opacity remains high and the PMS star remains fully convective, it contracts along its Hayashi line and thus its luminosity decreases. One such evolutionary track is given approximately by:

$$\log L = 10 \log M - 7.24 \log T_{\rm eff} + C \qquad (11.37)$$

This track is steeply descending, and shifts upwards with increasing initial mass.

As the central temperature rises, we move beyond the peak in the opacity curves shown in Figure 5.5, into a region where the opacity follows a Kramers law, $\langle\kappa\rangle \propto T^{-3.5}$ (Lecture 5.5). A radiative core develops, growing with time to encompass more and more of the star's mass. The point of minimum luminosity in the evolutionary tracks shown in Figure 11.3 corresponds to the development of a radiative core which allows energy to escape more readily into the convective envelope, causing the luminosity of the star to increase. The effective temperature continues to increase since the star is still shrinking.

Contraction continues, as dictated by the virial theorem, until the central temperature becomes sufficiently high to ignite nuclear fusion. Once the energy generated by H fusion compensates for the energy loss at the surface, the star stops contracting and settles on the *zero-age main sequence* (ZAMS), if its mass is above the hydrogen burning limit of $0.08 M_\odot$ (Lecture 10.4).

Before thermal equilibrium on the ZAMS is reached, however, several nuclear reactions have already set in. In particular, a small quantity of deuterium ($\mathrm{D} \equiv {}^2_1\mathrm{H}$) is present in the interstellar gas out of which stars form, with mass fraction of order 10^{-5}. D is a very fragile nucleus that reacts easily with atomic hydrogen. What little is left over from Big-Bang nucleosynthesis [$(\mathrm{D/H})_{\mathrm{prim}} = 2.5 \times 10^{-5}$ by number] is destroyed by the second reaction in the p-p chain (see Figure 7.5):

$$ {}^2_1\mathrm{H} + {}^1_1\mathrm{H} \rightarrow {}^3_2\mathrm{He} + \gamma \tag{11.38} $$

which destroys all D present in a star at $T \sim 1 \times 10^6\,\mathrm{K}$, while the protostar is still on the Hayashi line (see Figure 11.3).[3]

The reaction (11.38) is exothermic and produces 5.5 MeV of energy, sufficient to halt the contraction of the PMS star for $\sim 10^5\,\mathrm{yr}$. (A similar, but much smaller effect, occurs somewhat later when lithium is burned at a higher T). Furthermore, the ${}^{12}\mathrm{C}(\mathrm{p},\,\gamma){}^{13}\mathrm{N}$ reaction is already activated

[3]This process is known as the *astration* of deuterium, and makes the calculation of the abundance of deuterium as star formation progresses in a galaxy particularly simple. For the same reason, the D/H ratio measured in the present-day interstellar medium is a *lower limit* to the primordial abundance of deuterium.

at a temperature below that of the full CNO cycle, due to the relatively large initial abundance of ^{12}C compared to the equilibrium CNO abundances. Thus almost all ^{12}C is converted into ^{14}N before the ZAMS is reached. The energy produced in this way also halts the contraction temporarily and gives rise to the wiggles in the evolutionary tracks just above the ZAMS location visible in Figure 11.3. Note that this occurs even in low mass stars, with $M < 1M_\odot$, even though the p-p chain is their main energy generation mechanism.

Finally, the time taken for a protostar to reach the ZAMS depends on its mass. This time is essentially the Kelvin-Helmholtz contraction timescale which we considered in Lecture 7.1. Since contraction is slowest when both R and L are small (recall from Lecture 7.1 that $\tau_{\rm KH} = \Delta E_g/L$ and $\Delta E_g \propto 1/R$ for a given mass M), the pre-main sequence lifetime is dominated by the final stages of contraction, when the star is already close to the ZAMS. For stars with $M \geq 1M_\odot$, we find $\tau_{\rm KH} \simeq 5 \times 10^7 (M/M_\odot)^{-2.5}$ yr, while for lower mass stars the mass dependence is somewhat shallower (see Figure 11.4). Thus, massive protostars reach the ZAMS much earlier than lower-mass stars (and the term zero-age main sequence is somewhat misleading in this context, although it hardly makes a difference to the total lifetime of a star).

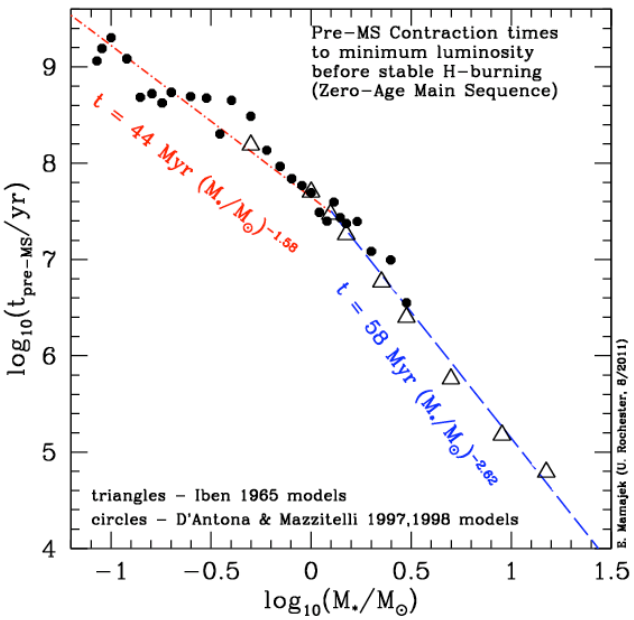

Figure 11.4: Pre-main sequence contraction times as a function of stellar mass, from the evolutionary tracks calculated by D'Antona & Mazzitelli and by Iben, as indicated.

11.8 Nomenclature of Objects Associated with Star Formation

Before concluding the topic of star formation, we briefly mention some of the objects which are associated with this process.

T Tauri stars (named after the first star of their class to be identified as an object of special interest) are an important class of low-mass, pre-main sequence objects. As can be seen from Figure 11.5, many of them lie on the Hayashi tracks. Some of their characteristics [many of them common to Young Stellar Objects (YSO) in general] are:

1. Variability in the light output with timescales of the order of days, as material falls down onto the surface of the star from a residual disk;

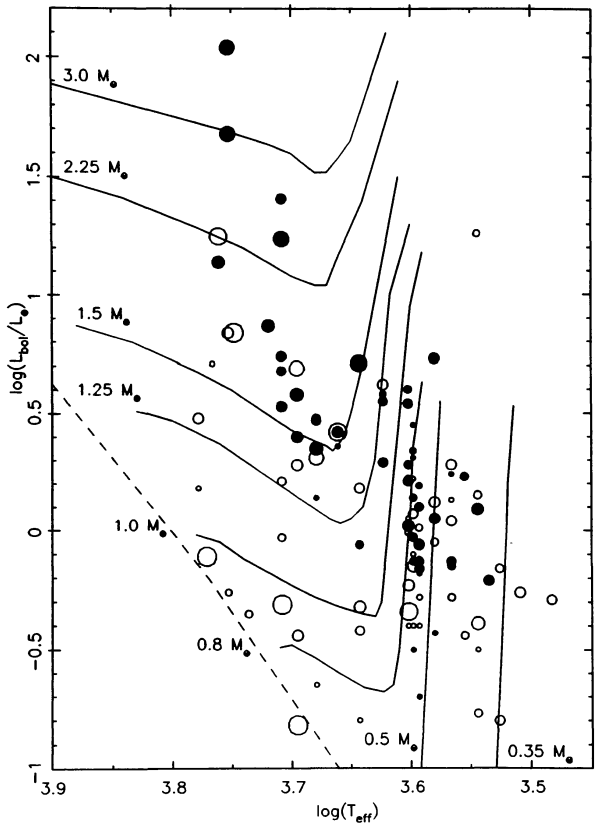

Figure 11.5: Positions of T Tauri stars on the H-R diagram. The sizes of the circles are proportional to the rate of rotation. Filled circles indicate stars with strong emission lines. Theoretical pre-main sequence evolutionary tracks are also shown. (Reproduced from Bertout 1989, ARA&A, 27, 351).

2. Emission lines from both permitted and forbidden transitions, from the disk, or the bipolar outflow, or both. Some emission lines have a characteristic P-Cygni profile, which is a combination of emission and absorption indicative of mass loss.

3. Higher infrared luminosity than main sequence stars of the same mass, because there is more dust in their immediate vicinity.

4. A high level of what is called activity is seen, meaning flares, star spots and emission from a hot corona (including X-ray emission). There seem to be two reasons for this: many T Tauri stars are rapid rotators, with rotation periods from hours to days, as opposed to a month for the Sun.[4] Second, as we have seen, the convection zone extends deeper into the star when it is on the Hayashi track than when it reaches the main sequence. The combination of rapid rotation and deep convection results in a strong dipole magnetic field which, in turn, drives the activity.

Herbig-Haro objects are apparently associated with the jets produced by T Tauri stars. As the jets expand supersonically into the ambient interstellar medium, collisions excite the gas resulting in bright nebulosities with emission line spectra. Continuous emission is also observed in some protostellar objects as the central star, which is hidden from direct view behind dust in the accretion disk, illuminates the surfaces of the disk. These accretion disks seem to be responsible for many of the characteristics of YSOs, including emission lines, mass loss, jets, variability, and ultimately planet formation.

OB associations are groups of stars dominated by O and B-type main sequence stars (lower mass stars are of course also present, but they are less conspicuous being less luminous). Since the lifetimes of O stars are $< 10^7$ yr, OB associations mark the sites of recent star formation (see Figure 11.6), and indeed they are usually found close to molecular clouds (although not necessarily the parent cloud from which they formed, which is rapidly dispersed following the formation of massive stars—see later). Studies of the kinematics of individual stars in OB associations generally

[4]Stars are born with high rotational velocities which subsequently decline with age. The braking is thought to result from the interaction of the star's magnetic field and the stellar wind emitted from the photosphere causing a steady transfer of angular momentum away from the star.

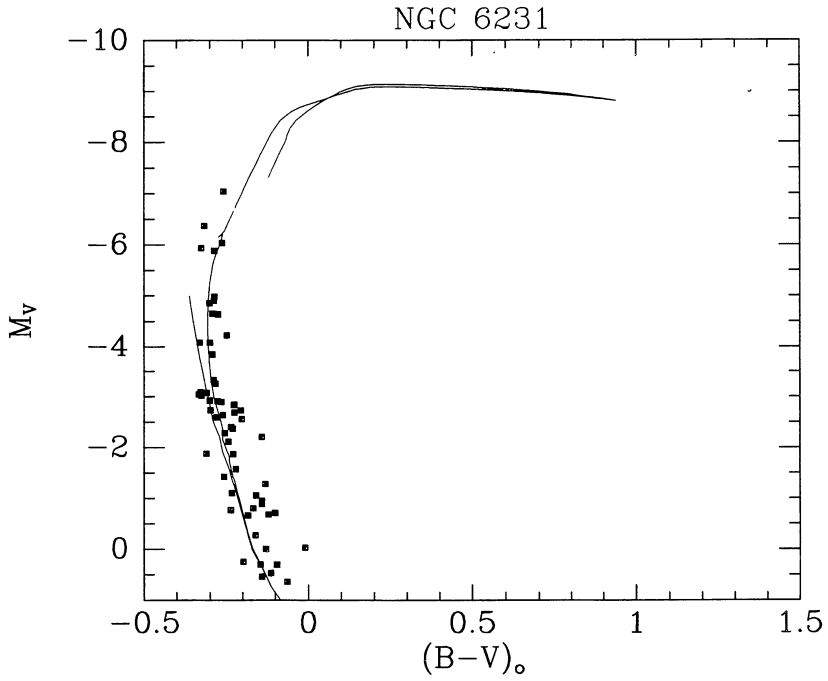

Fig. 10. Same as in Fig. 9 for NGC 6231, $m - M = 12.50$, $E(B - V) = 0.46$, log $age = 6.75$

Figure 11.6: Isochrone fitting to the upper main sequence of the open cluster NGC 6231 indicates an age of $\sim 5.6 \times 10^6$ yr. (Reproduced from Meynet et al. 1993, A&A Supp., 98, 477).

lead to the conclusion that they are not gravitationally bound and they they will eventually dissolve into the field stellar population.

On the other hand, **super star clusters**, most commonly found in regions undergoing very intense episodes of star formation—or starbursts, consist of 10^4 to 10^6 stars concentrated within a few pc. Given the high density of stars, they may remain bound even after their massive stars have exploded as Type II supernovae and may also survive other disruptive processes later on. If so, after several billion years, they would evolve into objects similar to the old globular clusters in the halo of the Milky Way. Thus, super star clusters may simply be young globular clusters.

11.8.1 H II Regions

Whether isolated or in a cluster, massive stars of spectral type O and B have a profound effect on their surroundings once they arrive on the main sequence. With effective temperatures $T_{\rm eff} > 30\,000$ K, the peak of their blackbody curve is at ultraviolet wavelengths. Photons with $\lambda < 912$ Å are sufficiently energetic to ionise ground-state hydrogen atoms (with

ionisation potential IP = 13.6 eV) in the surrounding interstellar medium, producing an H II region around the star(s).

We can calculate the size of an H II region provided we know the number of ionising photons emitted by the star and the ambient density. In a steady state, the ionisation rate (number of ionisations per unit time) must balance the recombination rate (the inverse process). If such an equilibrium did not develop, the size of the H II region would continue to grow until Lyman continuum photons are diluted enough for equilibrium to be established.

The recombination rate per unit volume is given by:

$$\mathcal{R}_{\rm rec} = \alpha(T)\, n_H\, n_e$$

where n_H and n_e are the volume densities of H^+ ions and electrons respectively, and $\alpha(T)$ is the temperature dependent radiative recombination coefficient. Since hydrogen is the most abundant element and is fully ionised, $n_H \simeq n_e$ Assuming spherical symmetry, we therefore have:

$$Q_* = \mathcal{R}_{\rm rec} \frac{4}{3} \pi r_{\rm HII}^3 \tag{11.39}$$

where Q_* is the number of ionising photons emitted by the star per unit time and r is the radius of the H II region, also called the Strömgren radius from the Danish astrophysicist who first carried out the analysis in the late 1930s. Solving for the Strömgren radius, we have:

$$r_{\rm HII} = \left(\frac{3Q_*}{4\pi\alpha} \right)^{1/3} n_H^{-2/3} \,. \tag{11.40}$$

The Strömgren sphere of a star marks the sharp transition between fully ionised circumstellar gas and mainly neutral interstellar gas. Within the H II region, hydrogen is continuously being ionised and recombining with electrons. In general, recombination takes place to a high energy level, followed by cascading of the electron through intermediate levels to the ground state. Each step in the process is accompanied by the emission of a photon with energy *lower* than the original 13.6 eV responsible for ionising H from the ground state. The dominant visible wavelength photon produced in this way results from the transition between the $n = 3$ and $n = 2$ levels, corresponding to the Hα line of the Balmer series, at a wavelength of 6563 Å, in the red region of the optical spectrum. It is this process that causes H II regions to fluoresce in red light.

11.8.2 Coeval Star Formation in Clusters?

The effect of massive stars on their environment goes far beyond ionising the gas within the Strömgren sphere. All stars more massive than $\sim 20 M_{\odot}$ experience mass loss driven by radiation pressure while still on the main sequence. For the most massive stars, the energy deposited into their surroundings via strong stellar winds is comparable to that associated with the explosion of a Type II supernova, $\sim 10^{51}$ ergs.

The 'double-whammy' of stellar winds and supernova explosions can disperse the remainder of the parent molecular cloud on a timescale of only a few million years, depending on the richness of the newly formed cluster of OB stars, effectively shutting down any further star formation. This timescale is shorter than the contraction times of lower mass stars, and yet many clusters are known to contain both early-type and late-type stars. This apparent contradiction, first noted by the American astronomer George Herbig in the 1960s, has led to the view that, when stars in a cluster form, the low- and intermediate-mass stars form first, with the process continuing gradually until the high-mass stars form and quickly halt all subsequent star formation. It may also be the case that the most massive stars form by mergers of smaller stars in dense protostellar environments.

11.9 The Initial Mass Function

The end product of an episode of star formation is a cluster of stars of different masses. Observationally, it is clear that low mass stars are much more common than high mass stars. The distribution of stellar masses immediately after a population of stars arrives on the main sequence is described by the *Initial Mass Function* (IMF). For a young cluster, the IMF can be deduced from the observed *Present-day Mass Function*, after correcting for the fact that the most massive stars have evolved off the main sequence in clusters older than a few million years. In the field, we need to combine the PDMF with a knowledge (or model) of the past history of star formation to reconstruct the IMF. In both cases, models are required to convert the measured stellar luminosities to stellar masses.

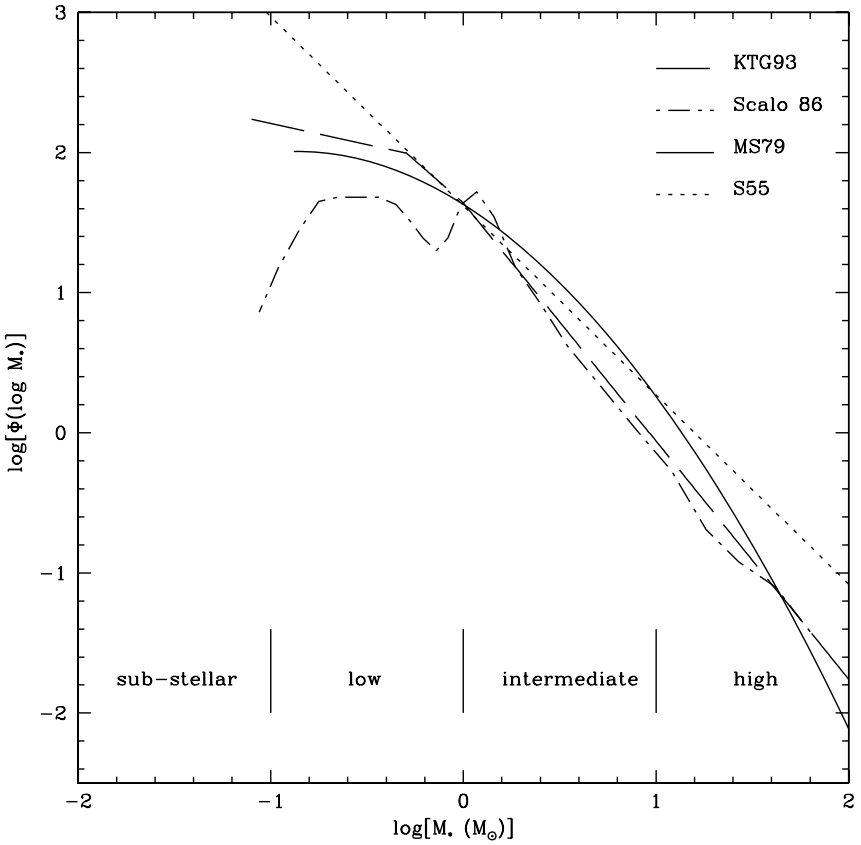

Figure 11.7: Different realisations of the stellar Initial Mass Function. S55 is for Salpeter (1955), and KTG93 is for Kroupa, Tout & Gilmore (1993). (Figure reproduced from Meyer et al. 1999, astro-ph/9902198).

The simplest form of the IMF is a single power law of the form:

$$N(M)\,dM \propto M^{-\alpha}\,dM \tag{11.41}$$

where $N(M)\,dM$ is the number of stars per unit volume with mass between M and $M + dM$.

Edwin Salpeter suggested in 1955 that, in the solar neighbourhood, the frequency distribution of stars more massive than the Sun follows eq. 11.41 with $\alpha = 2.35$. Other formulations have been proposed, usually consisting of a combination of different power laws in different mass intervals (see Figure 11.7). Thus, for example, the Kroupa, Tout & Gilmore (1993) IMF is a combination of three power laws:

$$N(M)\,dM \propto \begin{cases} M^{-2.7}\,dM & \text{if } M > 1M_\odot \\ M^{-2.2}\,dM & \text{if } 1 \geq M \geq 0.5M_\odot \\ M^{-(0.70\,\text{to}\,1.85)}\,dM & \text{if } 0.5 \geq M \geq 0.08M_\odot \end{cases} \tag{11.42}$$

The faintness of low mass stars, the rarity and short lives of high mass stars,

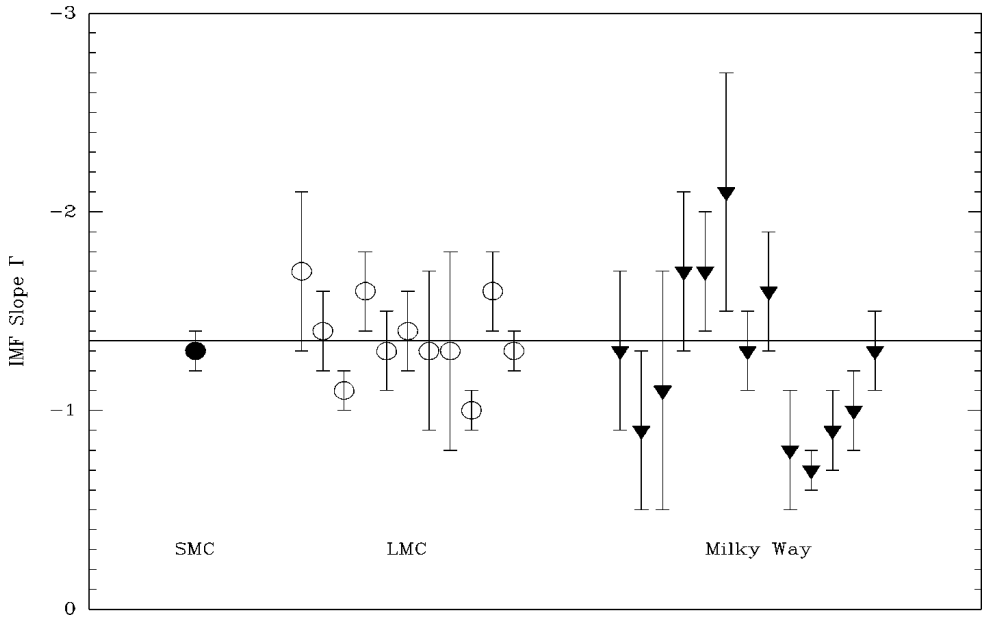

Figure 11.8: The slope of the IMF (in its integral, rather than differential, formulation) measured in OB associations and clusters in the Milky Way and our companion galaxies, the Large and Small Magellanic Clouds, is consistent with a Salpeter slope of $\alpha + 1 = -1.35$. (Figure reproduced from Massey 2003, ARA&A, 41, 15).

and the frequency of binaries are all issues affecting the determination of the IMF.

A much debated question is whether the IMF is 'universal', or whether it varies with local conditions, such as metallicity, or with redshift. Opinions are divided as to the reality of claimed IMF variations, although it seems to be well established that the same IMF applies to stars in the Milky Way and Large and Small Magellanic Clouds (our companion galaxies—see Figure 11.8). Most astronomers consider it likely that the IMF of the First Stars that formed in the Universe was 'top-heavy', i.e. lacking in low mass stars.

The stellar initial mass function is a crucial factor in many astrophysical problems. Examples are the determination of the star formation rate (SFR) in galaxies and studies of the chemical evolution of galaxies. In the former case, the SFR is usually determined from a tracer—such as the Hα emission line—which is produced by stars in a limited mass range (for Hα the most massive stars dominate), and then the result is extrapolated to all stellar masses. In the latter, the relative proportions of different chemical elements depend on the slope and mass range of the IMF because stars of different masses synthesise different elements in different proportions. Typical examples are Oxygen which is produced and ejected into

the ISM mainly by massive stars, and Iron which is due mostly to low- and intermediate-mass stars. Thus, altering the IMF would affect the resultant (O/Fe) ratio following an episode of star formation. The determination of the mass-to-light ratio in galaxies is another example of a problem where the slope and mass range of the IMF are crucial: how much mass is 'hidden' in very low mass stars and 'brown dwarfs' depends sensitively on the poorly known IMF in the subsolar regime.