S26 (Short Option) – Stars & Galaxies – Dr Andrew Bunker

Syllabus for Section 1 (Lectures 1-4)

Measurement of physical properties of stars and galaxies. Parallax and the distance ladder. Magnitude systems and their relationship to quantitative measurements of luminosity and effective temperature. Observational properties of stars and galaxies: the H-R diagram, stellar clusters, basic description of the structure of the Milky Way; the Hubble classification of galaxies; galaxy luminosity functions.

Lecture 1

Distance in Astrophysics

In astronomy, we deal with vast distances. As light travels at a finite (but admittedly fast) speed, as we look through space we look back through time. We see the Sun as it was 8 minutes ago and the nearest star as it was 4 years ago. The public often use the "light year" as a unit of distance (but astronomers prefer the "parsec", pc, which is 3.26light years, see below).

- **Light-year : Distance travelled by light in 1 year = 9.46×10^{15} m**

The centre of our galaxy is 28,000 light years away (8500pc = 8.5kpc), and the nearest large galaxy, Andromeda, is 2,500,000 light years (800,000pc=0.8Mpc), so we are seeing these as they were before recorded human history. The furthest know galaxies are seen about 95% of the way back in time to the Big Bang, 13.7Gyr (13,700,000,000years) ago*. So we effectively have a "time machine" to study galaxy evolution.

**[Note that there are various ways of quoting "distance" to objects in an expanding Universe – you can see it can get confusing: is it the distance the light has travelled? Or the distance between us and the object when the photon left the distant galaxy? Or the distance today when the photon arrived at Earth? Some of this complexity is covered in the Bv paper on Cosmology and General Relativity, but in this course I will be dealing mainly with galaxies which are nearby rather than "high redshift". You can read Prof Balbus's notes on luminosity distance, angular size distance and proper motion distance for more detail.]*

- **Astronomical unit (AU) : Mean distance between the Earth and the Sun 1.50×10^{11} m**

An important distance which we can directly measure is the average distance from the Sun to the Earth, called the Astronomical Unit (AU). It is actually determined from the distance of the Earth to Venus from radar reflections, then applying Kepler's laws of orbits.

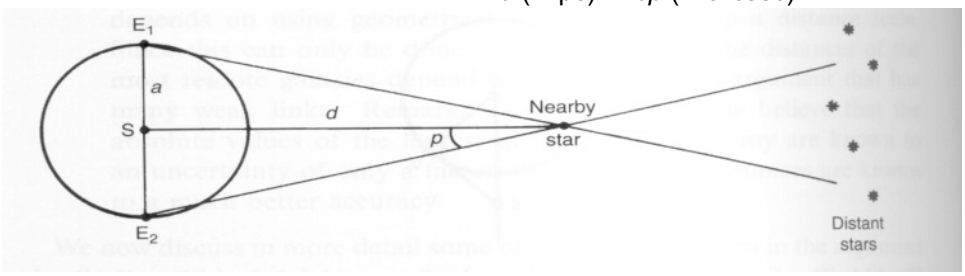
Pluto's mean distance from the Sun is about 40 AU (3.5 light hours)
The furthest any man-made objects has travelled is just beyond the edge of the Solar System, with Voyager 2. This is still only 0.03% the distance to the nearest star.

Units: astronomy often deals with quantities that are very large when quoted in SI (m k s) units, so often "physical" units are adopted. When dealing with stars, quantities are quoted relative to the Sun, which has subscript symbol ⊙. These are (on the exam Formula Sheet):
Solar mass $1M_\odot = 1.99 \times 10^{30}$kg (most stars have $0.1\ M_\odot < \text{mass} < 100\ M_\odot$)
Solar luminosity $1L_\odot = 3.85 \times 10^{26}$W (most stars have $10^{-4}\ L_\odot < \text{luminosity} < 10^{6}\ L_\odot$)
Solar Radius $1R_\odot = 6.96 \times 10^{8}$m (most stars have $0.01\ R_\odot < \text{radius} < 1000\ R_\odot$)

Trigonometric Parallax – Measuring the Distance of the Nearest Stars

The motion of the Earth around the Sun means that the apparent location of nearby stars will shift against a background of more distant objects from opposite sides of the Earth's orbit, and its distance can be obtained from trigonometry. The direct measurement of distance using *trigonometric parallax* can be done only for relatively nearby stars. The angle subtended at the star by the semi-major axis of the Earth's orbit around the Sun is called the *parallax*, and it is 0.76 arcseconds for the nearest star. The distance corresponding to a *par*allax of 1 arc*sec* is called a *parsec*; 1 pc = 3.09×10^{16} m. If *d* is the distance of a star of parallax *p*:

d (in pc) = 1/*p* (in arcsec) .



Parsec (pc) : Distance at which 1 AU subtends an angle of 1 arc second (1")

- **60 arcsec in 1 arcmin, 60 arcmin in 1 degree**
- **1 parsec = 360 x 60 x 60 / 2π AU = 206265 AU = 3.26 light-years = $3.09\text{x}10^{16}$ m**

The HIPPARCOS satellite 20 years ago gave parallax determinations accurate to 0.001arcsec, and distances out to a few hundred parsecs measured by this direct method, and GAIA is now doing even better. Note that even the centre of our own galaxy (at 8.5kpc) is beyond this, and indirect methods must be used to determine its distance, the distances to most stars in our Galaxy, and the distances to all external galaxies.

- **Distance to nearest star ~1pc; radius of Milky Way galaxy ~10kpc (10,000pc);**
- **Distance to nearest galaxy ~1Mpc (10^6pc); observable Universe ~10Gpc (10^{10}pc)***

Brightness

The energy output rate (power) in electromagnetic radiation is called ***luminosity*** (L), with the SI unit Watts. If the radiation is emitted isotropically (approximately true for stars), then at a distance D away we observe a ***flux of energy,*** f (energy crossing unit area in unit time), such that

$L = f \times 4\pi D^2$

where f has units W/m^2. Note that the flux falls off with distance as an inverse square law.
The total rate of energy output is called the ***bolometric luminosity*** (L_{bol}), but we are often concerned about the particular distribution of energy with wavelength – the shape of the spectrum. We can determine this either by dispersing the light and measuring the intensity as a function of wavelength with a spectrograph, or we can take images with different filters and measure the relative brightness of the same object in different colours.
Hence it is sometimes useful to define a ***luminosity density*** or ***intensity*** (luminosity per unit wavelength interval, L_λ in W/m, or per unit frequency interval, L_ν in W/Hz), and similarly flux density.

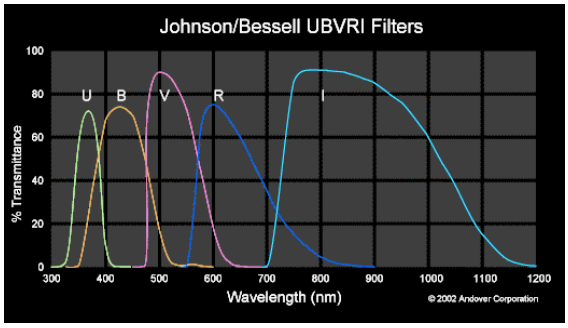
The Magnitude System

Astronomers adopt another convention to describe the brightness of objects – the magnitude system. This is a logarithmic system, with the brightest stars in the sky having magnitude about 0, and the

faintest visible with the unaided eye having magnitudes around +6 (note – the bigger the magnitude the fainter the object!). The equation relating ***apparent magnitude***, *m*, to flux is

$$m = \text{constant} - 2.5 \log_{10} flux$$

where the "constant" is actually a zeropoint, which depends on the filter you use (what colour of light). The factor of 2.5 is chosen such that a 1% brightness change corresponds to about 0.01mag difference, and 10% is 0.1mag (breaks down for large changes).Traditionally, the star Vega (α Lyrae) is defined as having magnitude zero in all filters – but unfortunately this means the system is tied to observation, and not the SI system of units.



The magnitude measured in a particular filter is m_λ, (e.g. m_V is the magnitude in the V-band, the centre of the visible spectrum at λ~550nm, m_R in the red at ~650nm, and m_B in the blue at ~450nm). Many of the filters are chosen because the Earth's atmosphere transmits in those windows. We can define a ***colour*** as the ratio of fluxes in two filters, or a difference in magnitudes:
$B\text{-}V=m_B-m_V=2.5\log_{10}(f_V/f_B)$
The larger the colour, the redder the object.
The magnitudes observed in these filters can be corrected to the bolometric (total) magnitude by a bolometric correction (B.C.) appropriate for that filter, integrating flux over all wavelengths: $m_{bol} = m_\lambda + B.C.$

So far we have considered the observed fluxes from objects represented by the apparent magnitude, but what of luminosity? We would need to view objects at a fixed distance if the magnitudes were to equate to (logarithmic) luminosities. By convention, we use the term ***absolute magnitude*** (M) to be the magnitude that would be measured if the object was at a distance of 10 parsecs.

$$M = m - 5 \log_{10} (D/10 \text{ pc}) = m - 5 \log_{10} D + 5 \text{ (where D is in parsecs).}$$

The quantity $m\text{-}M=5 \log_{10} D - 5$ is know as the ***distance modulus***

But what if some light is absorbed along the way? This phenomenon is common in astrophysics, as dust is produced in stars and dispersed into the interstellar medium in late stages of stellar evolution. This dust can scatter light, and has the effect of absorbing more blue light than redder wavelengths. We call this dust reddening – with the effect that the observed spectrum looks redder than the intrinsic spectral shape emerging from the source (not to be confused with redshift!). We can quote the absorption at a particular wavelength in magnitudes of extinction (A):

$$m_\lambda - M_\lambda = 5 \log_{10} D - 5 + A_\lambda$$

We can also express this as a colour excess *E* (in magnitudes), that is how much redder the colour (e.g. B-V) is than its intrinsic value (without dust), $B_0\text{-}V_0$:
$B\text{-}V= B_0+A_B-(V_0+A_V) = B_0\text{-}V_0 + E(B\text{-}V)$ where $E(B\text{-}V) = A_B - A_V$. Observationally, $A_V \approx 3.2E(B\text{-}V)$

<u>The Distance Ladder</u>

In order to measure greater distances than the few stars with parallaxes, we can use "standard candles" – that is, objects for which we know the luminosity, and measure their flux (apparent brightness) and use $L= f \times 4\pi D^2$. But what objects have known luminosity? Certain stars are variable (those on the "instability strip" of the HR diagram that we will study later, where the opacity of the stellar atmosphere changes with temperature), and we can recognise the change in their brightness with time over days and weeks. RRLyrae stars have a characteristic light curve with a period of 13 hours, and all have about the same absolute magnitude ($M_V \approx +0.6$ mag). We can use these to determine distances within our own galaxy, but at about 50 $L_\odot$ they are too faint to see in external galaxies.

Another more luminous class of variable star are Cepheids, which have a tight correlation between their pulsation periods and their luminosity: $M_V=-1.4-2.8\log_{10}(\text{period/days})$. Cepheids can be seen in external galaxies with periods of a few days, and measuring the

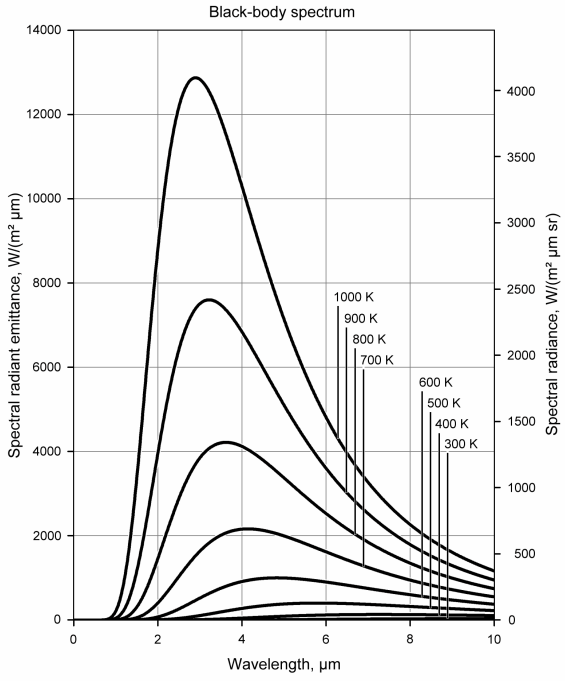
period gives us the luminosity (absolute magnitude). We measure their apparent magnitude, and hence derive the distance modulus. The Hubble Space Telescope Key Project measured distances to several galaxies, and determined the value of the Hubble constant H_0 (not really a constant, but the relation between the distance and the velocity of recession of galaxies in an expanding Universe at the current epoch):
$v=H_0 D$ (for small D) where H_0=71 km/s/Mpc (units are actually 1/time, approximately the reciprocal of the age of Universe).

The expansion history of the Universe depends on its matter content and any dark energy, and this dictates the relation between recession velocity (quoted as redshift) and luminosity distance – see the third year general relativity & cosmology course notes! But to get to great enough distances to test different cosmological models, you need even brighter "standard candles", and in recent years Supernovae Type 1a have been used to show that there may be enough "dark energy" that the expansion of the universe is accelerating.

Radius of a Star and Surface Temperature

Measuring stellar radii is very challenging (the disks of stars generally cannot be resolved), but a few measurements exist from interferometry, or binary star systems with their orbital plane in our line of sight which eclipse (by studying their light curves).
However, if we know the luminosity and spectrum of a star we can estimate its radius R_s if we assume it radiates as a black body with effective temperature T_{eff}:

$$L = 4\pi R_s^2 \sigma T_{eff}^4 ,$$



where σ is the Stefan-Boltzmann constant (the total power output is the surface area of the star multiplied by the surface emissivity of a black body). T_{eff} is determined from fitting the shape of the spectrum, although stars often deviate from a perfect black body due to emission and absorption features in the stellar atmospheres. A black body radiates a Planck spectrum:

intensity $I_\nu = B_\nu(T) \equiv \frac{2\eta\nu^3}{\chi^2} \frac{1}{\varepsilon^{\eta\nu/\kappa T} - 1}$

The wavelength with the peak energy output (λ_{max}) is inversely proportional to the temperature, so red-hot stars are actually cooler than blue stars. The Wien displacement law is a useful approximation: $\lambda_{max} \approx 0.0029/T$. Stars typically have temperature in the range 3000K to 100,000K.

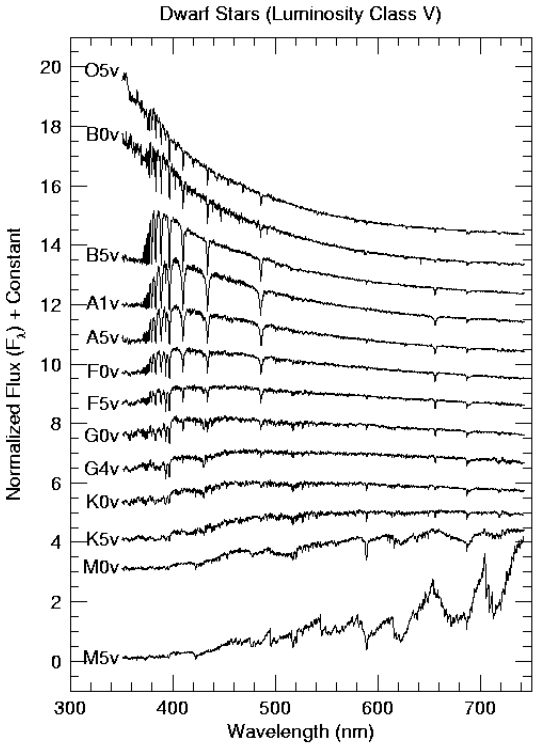
Stellar Mass.

Again very few direct determinations of masses of stars exist. We can apply Kepler's third law to binary star systems: $a^3/P^2=G(M_1+M_2)/4\pi^2$ where "a" is the average separation and P is the period to orbit each other. [Or $(M_1+M_2)=a^3/P^2$ where masses are in solar masses, a is the separation in AU and P is in years.] In visual binary cases where we know the absolute orbits around the centre of mass, we have $M_1/M_2=a_2/a_1=v_2/v_1$. More usually we have spectroscopic binaries, where we can only obtain M_1/M_2 and $M_1\sin^3 i$ (where i is the inclination of the orbits to our line-of-sight). For rare spectroscopic eclipsing binaries we know $i \approx 90°$ and hence we can derive actual stellar masses rather than lower limits.

S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 2: Spectral Classification of Stars, Star Clusters and the Herzsprung-Russell Diagram

The Surface Temperature of Stars: Stars are predominantly classified by their surface temperature - which can be estimated from their colours (blue=hot, red=cool) or by fitting a black body curve to the spectrum. It can also be determined from the strength of spectral lines (telling us the ionization & excitation state of various atom species in the stellar atmospheres), as can be seen in the figure below right.



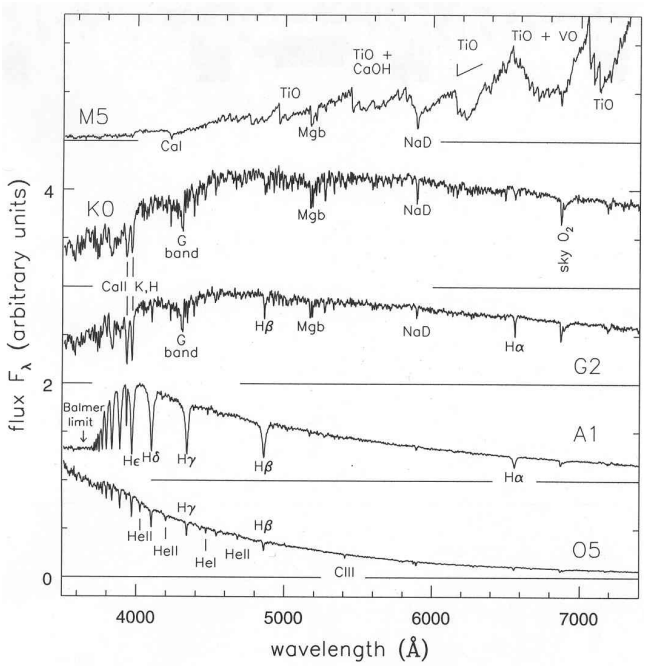


Figure 1.1 Optical spectra of main-sequence stars with roughly the solar chemical composition. From the top in order of increasing surface temperature, the stars have spectral classes M5, K0, G2, A1, and O5 – G. Jacoby *et al.*, spectral library.

The figure above left shows stellar spectra, with the hottest at the top. Note that the hottest are the most blue (relatively more flux at short wavelength) and are well fit by a smooth black-body curve, whereas the cooler stars have complicated spectra with deviations from this fit.

The fundamental independent properties of stars are mass, luminosity and radius. Luminosity is derived from apparent magnitude and distance. Effective temperature of the surface depends on luminosity and radius. The "effective temperature" $T_{\rm eff}$ of the stellar photosphere (where the light emerges) is defined by luminosity $L = 4\pi R^2 \sigma T_{\rm eff}^4$ (where R is the stellar radius).

Astronomers adopt the Harvard classification sequence for stars, denoting each spectral type by a letter. From hot to cold, these are: O B A F G K M (and L and T for some very cool dwarf stars). Each of these seven surface temperature bins split into 10 sub-units:

...A8,A9,F0,F1,... $T_{\rm eff}$ ranges between 2000K and 100,000K.

Originally the Harvard Spectral Sequence was ordered *alphabetically* according to the strength of the hydrogen Balmer absorption lines (transitions from $n = 2$), A B F G K M O, where these lines are strongest in A-type stars and weakest in M and O. However, it was soon realised that the fundamental property governing the stellar spectra was temperature, and the series was re-ordered from hottest (O-stars) to coolest (M-stars).

If we look in detail at stellar spectra, we see that they are not smooth black-body "Planck Curve" spectra. There are absorption lines (and sometimes emission lines) from particular ions, or even neutral atoms or molecules in the coolest stars. Which lines and of which elements depends on the surface temperature of the star. So these lines provide a surface temperature diagnostic. This does NOT mean that there is no sodium in O-stars, and no hydrogen in K-stars, for example. In reality, the chemical abundances might be similar, but it is the surface temperature which mainly drives which absorption lines form.

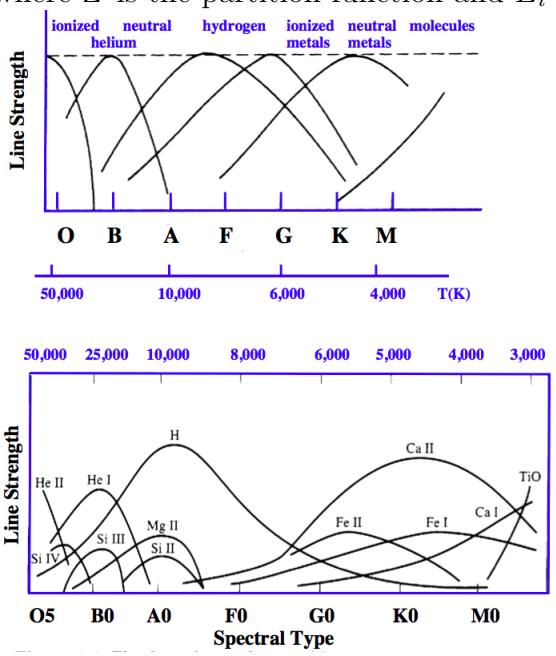
Detailed treatment of stellar atmospheres is beyond this course, but the Boltzmann equation gives the ratio of atoms in different energy levels:

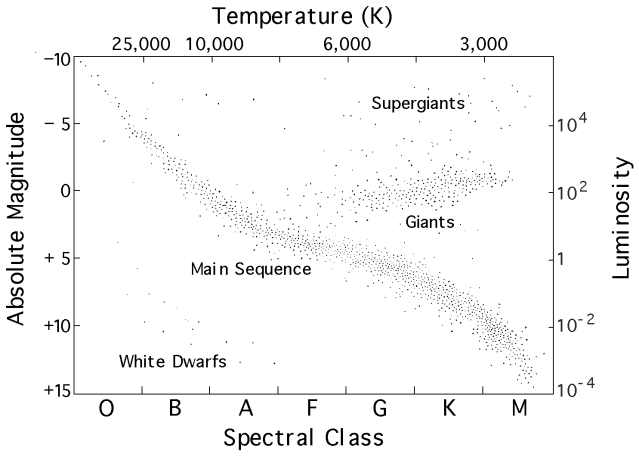
$$N_B/N_A = (g_B/g_A)e^{-(E_B-E_A)/k_BT}$$

where g is the statistical weight However, must also consider how many atoms are in that $i^{\rm th}$ ionization states, which is given by the Saha equation:

$$N_{i+1}/N_i = 2Z_{i+1}/(n_eZ_i)(2\pi m_e k_B T/h^2)^{3/2}e^{-E_i/k_BT}$$

where Z is the partition function and E_i the ionization energy.



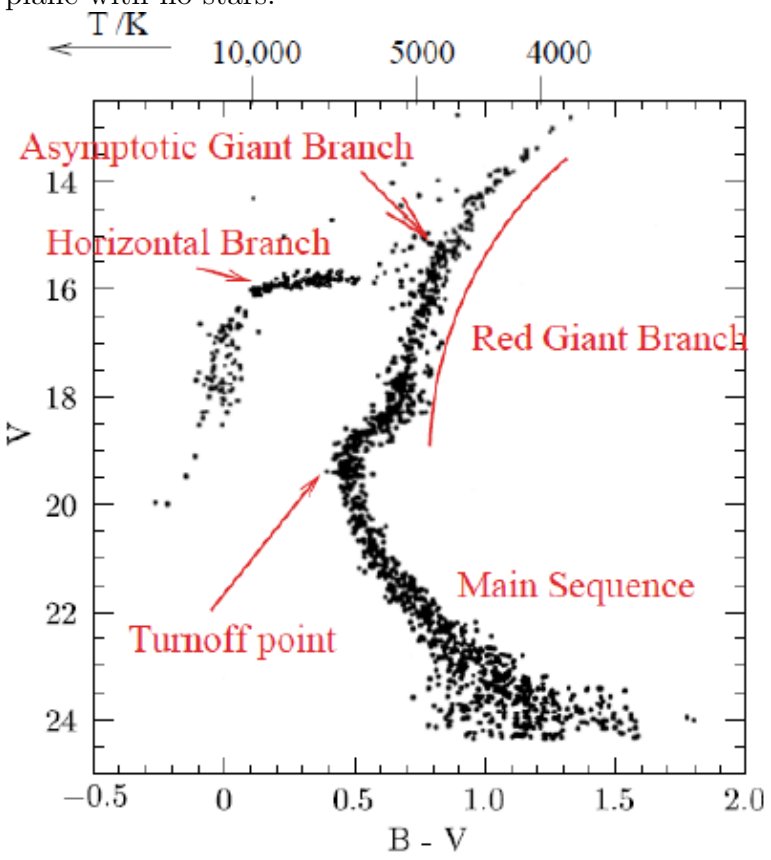


The hottest O-stars (>20000K) have high ionisation species (He^+ which is called HeII, CIII, etc.) and most of the hydrogen is fully ionised. B-stars show neutral helium. A-stars

at T=10,000K have the highest fraction of atomic hydrogen in the $n = 2$ state (hotter would imply more ionized and so no absorption lines, cooler implies more in $n = 1$ ground state or in molecules), and the Balmer absorption lines of hydrogen dominate. F and G stars have lines of neutral metallic atoms and once-ionized (e.g. Ca^+) – our own Sun is a G2 star with T=6000K. K & M stars are cooler (down to ∼3000K) and molecules can form in stellar atmospheres, creating strong absorption bands.

If we plot the luminosity of stars against their surface temperature (usually these are plotted logarithmically), this is the **Hertzsprung-Russell (HR) diagram**. We see that stars do not fall at random, and many stars show a correlation between these quantities. We could also plot the related properties of absolute magnitude (e.g. M_V instead of luminosity) against colour index (e.g. $B - V$ or spectral class rather than effective surface temperature). Note that increasing luminosity is plotted upwards, and increasing temperature to the left (backwards...).

The *main sequence* is a band running from hot, high luminosity stars (top left) to cool, faint ones (bottom right). About 90% of all stars fall on the main sequence – stars spend 90% of their time on the main sequence, which is the phase when they fuse hydrogen to helium in their cores. The remaining 10% cluster in three main regions: giants, supergiants and white dwarfs. Comparing stars at a given temperature, $L \propto R^2$, and so the most luminous stars are also the largest. There are many regions of the L, $T_{\rm eff}$ plane with no stars.



Left: Hertzsprung-Russell Diagram of a globular cluster.

Features of the Hertzsprung-Russell Diagram

MS Main sequence (90% of stars) – core fusion H → He

TO Turn-Off from main sequence (for a cluster)

AGB Asymptotic Giant Branch

RGB Red Giant Branch (cool stars, but very luminous due to large size, $100 - 1000\,R_\odot$)

HB Horizontal Branch;
note Gap - due to instability strip (often variable stars)

WD White Dwarfs (low luminosity $10^{-2} - 10^{-3}\,L_\odot$)

The *luminosity class* is given by a Roman numeral suffix according to the Morgan-Keenan (MK) Luminosity Classification: I – Supergiants; II – Bright giants; III – Giants; IV – Subgiants; V – Main sequences (dwarfs); VI – Sub dwarfs.
The Sun is a G2V main sequence star. We can estimate the luminosity class of a star from the surface gravity $g = GM/R^2$ as determined by the velocity spread of spectral lines.

Mass-Luminosity relation: For the few main-sequence stars for which masses are known, there is a *mass-luminosity relation*: $L \propto M^n$ where $n = 3 - 5$. The slope changes at extremes, less steep for low and high mass stars, and around $n = 4$ for solar mass stars. This implies that the main-sequence (MS) on the H-R Diagram is a function of mass i.e. from bottom (cool, faint) to top (hot, luminous) of main- sequence, stars increase in mass:

[lifetime on main sequence] $\propto$ [mass available] / [energy output]

Hence stellar lifetimes vary as M^{-3} or M^{-2} if $L \propto M^{3-4}$. Luminous stars are more massive and short lived. The lifetime of the Sun is $\sim 10^{10}$ years, but the lifetime of a $10M_\odot$ star is $\sim 10^7$ years (1000 times shorter)! Redder (cooler) main sequence stars are less luminous, less massive and live much longer. If we know a star is class V (on the main sequence), its spectral type ($T_{\rm eff}$) specifies its luminosity (absolute magnitude), so we can estimate the distance.

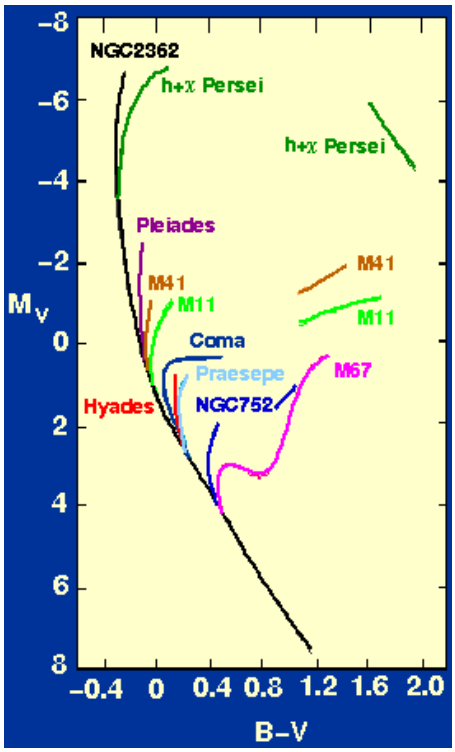
Stellar spectra: the strengths of absorption lines depend on three factors: chemical composition, temperature and luminosity. Many stars have similar compositions, so that the spectrum yields $T_{\rm eff}$ and M (absolute magnitude).

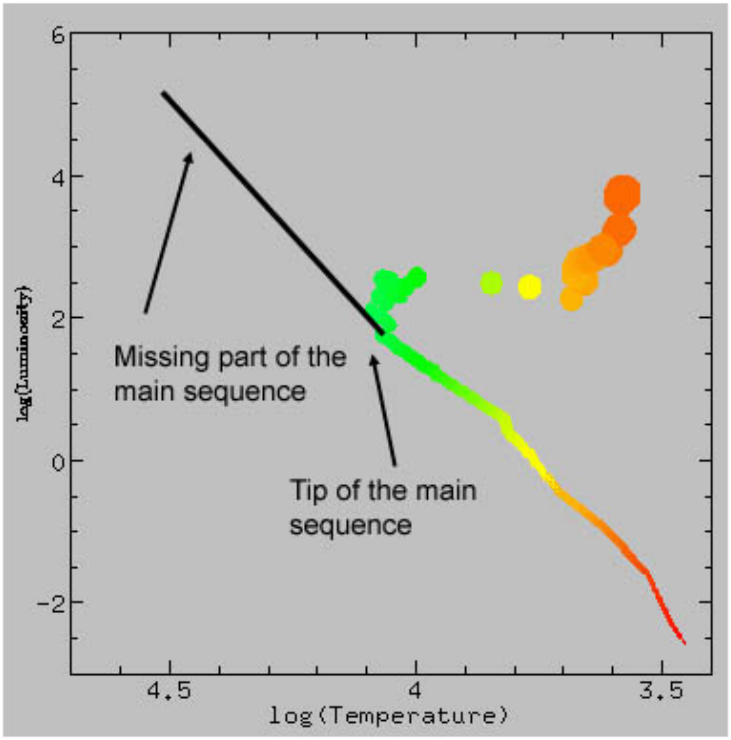
There are two other fundamental properties of stars that we can measure: age and chemical composition. All elements heavier than beryllium are made in stars by stellar nucleosynthesis (or during supernovae). We might expect successive generations of stars to enrich the gas with metals. The chemical composition is parameterised with: X, Y, Z, which are the mass fractions of hydrogen, helium and all other elements, e.g. for the Sun $X_\odot = 0.747$; $Y_\odot = 0.236$; $Z_\odot = 0.017$ (where $X + Y + Z = 1$). Note that Z (the mass fraction of all elements heavier than helium) is often referred to as *metallicity* in astrophysics. We can measure metallicity of stars from the spectral lines of elements, accounting for the surface temperature effects.

We would like to studies stars of same age and chemical composition – to keep these parameters constant and determine how models reproduce the other observables. *Star clusters* are gravitationally bound groups of stars, moving together through space, and these are useful laboratories for studying stars and stellar evolution: all the stars lie at the same distance, and presumably they all have similar ages and initial chemical composition. They provide large samples of stars at the same distance, and plotting the colours and magnitudes of these stars produces an *isochrone* on the HR diagram (all the same age).

*Globular cluster*s are very old, compact and roughly spherical, containing $10^5 - 10^6$ stars, and distributed in a spherical halo around the centre of our Galaxy. The HR diagrams of globular clusters are all very similar in shape. The main sequence is short, with a prominent giant branch, and a "horizontal branch" (on which RR Lyrae variable stars reside).

Galactic or *open clusters* are looser, smaller and more irregular groupings, and the stars formed together recently. Open clusters contain $10^2 - 10^3$ stars, and are concentrated in the plane of the Galaxy (where star formation occurs). Their HR diagrams are diverse (see figure below left). We often find very massive, luminous stars in Open Clusters.





We can estimate the distance to the clusters by comparing the apparent magnitudes of the main sequence stars, the red supergiant stars and the variable stars with those of similar nearby stars whose distances (and hence absolute magnitudes) are known, essentially moving the apparent magnitude y-axis up and down until it overlays the theoretical HR diagram plotted in absolute magnitude, and measuring the distance modulus $m - M$. This technique is sometimes called *Main Sequence Fitting*.

Evolution of Star Cluster HR Diagrams: As time progresses, massive stars (which are short lived) exhaust their core hydrogen fuel and leave the main sequence. The older the cluster, the further this "main sequence turn off" point has moved down the main sequence (see figure above right) – i.e. in an old cluster, the stars which are just about to leave the main sequence are cooler (redder), fainter (less luminous) and less massive than in a younger cluster.

If a star of luminosity L moves off the MS when it has transformed a fraction, f, of its mass from H $\rightarrow$ He (Where $f \sim 0.1$) then its MS lifetime can be estimated by:

$$\tau_{MS} = \eta f M\, c^2 / L$$

where η is the efficiency of conversion of H $\rightarrow$ He (which is about 0.7% of the mass in protons – the mass deficit of Helium-4 compared with H). For the Sun, putting in solar

luminosity and mass gets an estimated lifetime of 10^{10} years (which the Sun is half-way through).

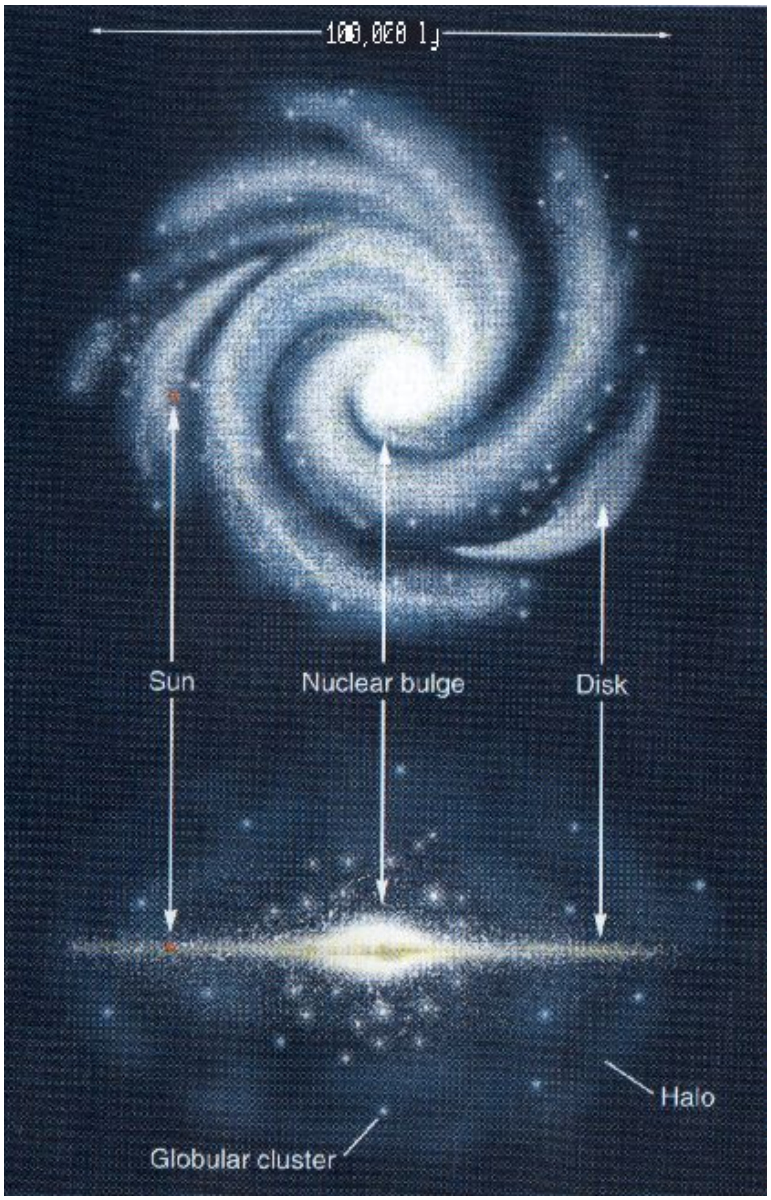
Since for MS stars $L \propto M^4$, then $\tau_{MS} \propto M^{-3}$. We can estimate the ages of star clusters from the turn-off point of their Main Sequence.

References

Carroll & Ostlie Chapter 8 ("The Classification of Stellar Spectra")
Choudhuri Section 3.6.2
Sparke & Gallagher; Sections 1.1.2, 1.1.3, 1.1.5, 2.2.3

S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 3: Our Galaxy, the Milky Way



Our Milky Way is a spiral galaxy of diameter about 30 kpc (100,000 light years). It comprises several distinct components:

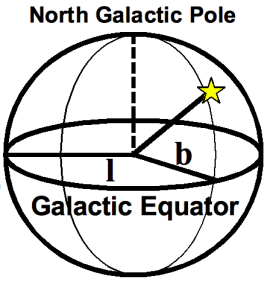
The Disc contains *Population I* stars (younger and more metal-rich), and gas and dust in the *Interstellar medium* (ISM). There is evidence for a thick disc and thin disc component to the stellar populations. The Sun is in the disk, and its orbital radius is 8.5 kpc.

The Halo contains *Population II* stars (older and metal-poor). Embedded in the halo are *Globular Clusters* (ensembles of $10^4 - 10^6$ stars held together by mutual gravity, ∼150 in Galaxy, most within 3 kpc of centre, some out to ∼25 kpc)

At the centre of the disk is **the Bulge**, which has a size of about 2kpc, and at the **Galactic Centre** (or *Nucleus*) there is evidence for a *supermassive Black hole*

As well as stars, the galaxy is made up of dark matter; gas and dust *(the interstellar medium).*

We can define a co-ordinate NGP system with the Galactic Plane as the equator and Galactic latitude (b), and the angle in this plane relative to the direction of the Galactic centre defining Galactic longitude (l).



The Galaxy hosts two distinct populations of stars which differ spatially, kinematically, and chemically:
Population I in the Galactic disc on nearly circular orbits in Galactic plane around Galactic centre. Ages of $10^6 - 10^9$ years. Range of metallicities (elements except H and He termed metals) Solar: $Z \sim 0.02$ (mass fraction of "metals"). Population I stars have $Z = 0.001 - 0.02$.
Population II in the Galactic halo. Orbits randomly inclined to Galactic plane. All old: ages $> 10^{10}$ years. Metal poor: $Z < 0.001$.

The **Dust** grains in the interstellar medium are $0.1 - 1$ microns in size. Dust distribution is non-uniform and clumpy. Effects on star light: scattered, reddened, polarized and converts optical and UV radiation to infrared and radio re-radiation. The dust is mainly confined to a layer (disk) 300 pc above and below Galactic plane (the Sun is near mid-plane). Obscuration depends on angle to plane, strongest within plane of disk.

The **Gas in the Interstellar Medium** has 4 different temperature/density regimes (molecular, atomic, ionized, coronal):
Molecular hydrogen (H_2) clouds, of which >4000 known, each of up to 10^5 solar masses. These are often the sites of star formation. The temperatures are cold ($T = 5 - 30$ K), and the density is high ($> 10^9$ molecules/m^3). There are associated dark dust clouds;
Atomic hydrogen (HI) clouds are well distributed through the Galactic plane, < 200 pc thick with cool temperatures ($T \sim 100$ K), and relatively high density ($10^6 - 10^9$ particles/m^3). HI emits a hyperfine structure line at 21cm (1420 MHz) due to spin transition of electron from spin parallel to proton to spin anti-parallel;
Ionised hydrogen (HII) arises from hydrogen ionised by UV radiation from nearby hot, massive, young stars. Hence HII regions at edges of molecular clouds (newly formed stars). HII captures free electrons which cascade down through various energy levels emitting UV and longer wavelength radiation as line emission. HII regions can be used to identify Galactic spiral arms;
Supernova remnants can result in very hot coronal gas, temperatures up to 10^6 K; Density $< 10^6$ particles/m^3. Supernova remnants form interconnecting tunnels around cooler clouds – ISM has the morphology of "Swiss cheese". Detected by: X-ray emission; Radio emission from synchrotron radiation; Absorption lines from ground states of highly-ionised atoms (e.g. OVI, NV) superimposed on spectra of background hot stars.

The Size and Structure of the Milky Way was originally derived from star counting (first done by Herschel, later in 1920 by Kapetyn – although not correcting for dust obscuration gave wrong size).

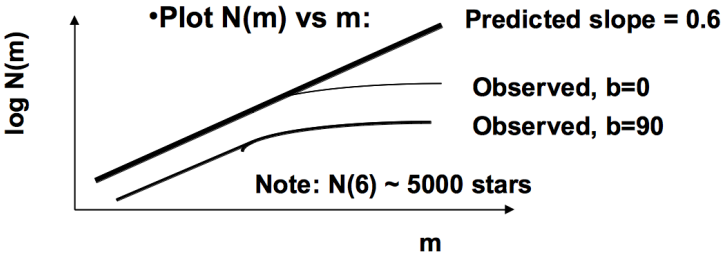
Star counts: imaging a Universe with a constant density of stars all of the same luminosity (absolute magnitude, M). If we take a series of images, each one progressively more sensitive, how will the number of stars rise with the limiting apparent magnitude? Remember that flux falls as $f = L/4\pi d^2$ and $m - M = 5 \log_{10} d - 5$, where d is the distance out to which we can see the stars at the limiting apparent magnitude m.
Also the volume out to which we can observe stars is $V = 4\pi d^3/3$.

So ratio of stars brighter than apparent magnitude (m+1) to those brighter than m is:

$$\frac{N(m+1)}{N(m)} = 10^{0.4\times 3/2} = 10^{0.6} \approx 4$$

(seeing fainter stars sees a larger volume).

These number counts diverge at large m (faint magnitudes) - brightness of night sky would be infinite! (this is *Olber's paradox*). It isn't, so the Galaxy is finite, and has dust obscuration, and there are redshift effects (and a horizon size of the Universe) for external galaxies.



Curves do not follow prediction as you go to fainter stars
$b = 90$ degrees: Galactic poles – Star density is lower – not uniform
$b = 0$ degrees: Galactic disc – Dust blocks starlight, so stars appear fainter than they should (points move to the right). Some stars not visible (points lower than they should be).

The difference between $b = 0°$ and $b = 90°$ is due to the vertical structure of stellar Galactic disc. Model this as an exponential fall in star counts: $D(z) = D_0 e^{-z/H}$ where D_0 is the density at $z = 0$ (stars per pc^3) and H is the scaleheight (about 100 pc).

The surface density of stars (per pc^2): $\Sigma = \int_{-\infty}^{\infty} D(z)dz \simeq 2D_0H$

Generally the structure is modelled as two exponentials (a *Thin Thick Disc*).

We can estimate the Star formation rate (SFR) in the disk of our Galaxy.
The local main-sequence stellar density in disc is $\approx 30 M_\odot\,\mathrm{pc}^{-2}$. If we assume this is the value over all Galactic disc (area πR^2), and the disc radius $R \approx 10\,\mathrm{kpc}$, then the total mass in stars is $M_{tot} \approx 1 \times 10^{10} M_\odot$. If we take the timescale to be the age of the Galaxy ($\approx 1.5 \times 10^{10}$ years, the age of oldest stars), then SFR$=$ M_{tot}/age $\approx 0.6 M_\odot$/yr. The typical mass of a star is $\sim 0.3 M_\odot$, therefore, SFR ~ 2 stars per year in our Galaxy. The above is a bad approximation – the disk is not uniform with radius, instead the stellar density is exponentially declining with radius.

From radial light profiles of other spiral galaxies $\Sigma(R) = \Sigma_0 e^{-(R-R_0)/L}$, Where L is the disc scale length ($2 - 5\,\mathrm{kpc}$) and R_0 the radius at which the surface density is Σ_0.

$$M = \int \Sigma(R) 2\pi R\, dR = \int_0^\infty \Sigma_0 e^{-(R-R_0)/L} 2\pi R\, dR = 2\pi\Sigma_0 L^2 e^{R_0/L} \int_0^\infty e^{-x} x\, dx = 2\pi\Sigma_0 L^2 e^{R_0/L}$$

(where $x = R/L$).

For our Galaxy, the disc scale length $L \approx 3.5\,\mathrm{kpc}$ so the mass estimate for stars in disk increases by a factor of ≈ 2.4 to $2.5 \times 10^{10} M_\odot$. There will also be additional mass in the form of dead stars (stellar remnants), since the age of Galaxy ($t_{\mathrm{gal}} \sim 1.5 \times 10^{10}$ years) is less than the Main Sequence lifetime for stars with $M_* < 0.8 M_\odot$. Main sequence stars with

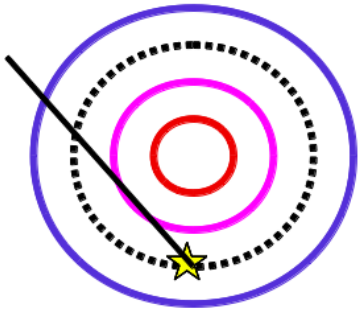
$M_* > 0.8M_\odot$ evolve into giants, white dwarfs, neutron stars and black holes within the age of the Milky Way, and these *stellar remnants* further increase the mass estimate by factor of ~ 1.7 to $\sim 50 M_\odot\,\mathrm{pc}^{-2}$, so $M_{\rm stars} = 4 \times 10^{10} M_\odot$. However, we can directly measure the total mass from the dynamics of the Sun's orbit: $M \approx V_0^2 R_0/G$ where $V_0 \approx 220\,\mathrm{km/s}$, $R_0 \approx 8\,\mathrm{kpc}$ for the Sun, so $M_{\rm dyn} \approx 1 \times 10^{11} M_\odot$. We are still only accounting for 40% of the dynamical mass in the form of stars and stellar remnants. Most of the mass missing! It may be in the form of *Dark matter*.

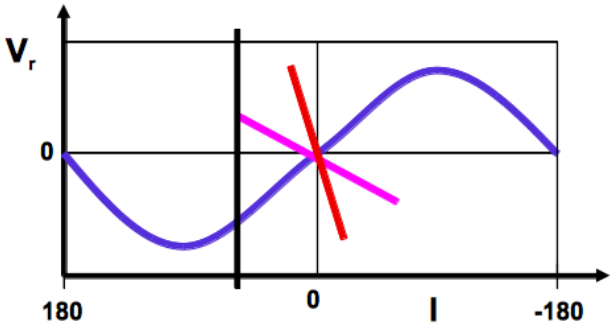
Mapping the Rotation of our Galaxy can be done using HI (neutral atomic hydrogen) which is distributed throughout the Galactic disc. We can map HI using 21cm emission and trace molecules (e.g. CO at 2.6mm). The Doppler shift gives radial velocity and hence distance.

Distance from radial velocities – assume gas on roughly circular orbits. Assume angular velocity ω decreases outwards (this is true for any reasonable model of the Galaxy). Radial velocities depend on angular velocity and Galactic longitude (l) of emitting gas

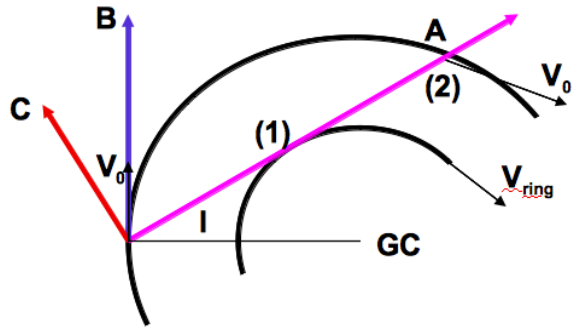
$$V_r = R_0(\omega - \omega_0)\sin l = (\frac{R_0}{R}V - V_0)\sin l$$

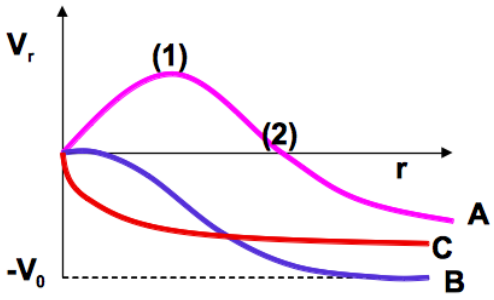
(R_0 is the radius of the Solar orbit, and V_0 is the speed of the Solar orbit).
Measure emission as function of longitude and radial velocity.
Determine Galactic rotation velocity as a function of radius.
Consider 3 rings of emitting gas with: $R \ll R_0$; with $R < R_0$; with $R > R_0$.





Inner rings: highest velocity at tangent point.
Positive velocities for $l > 0$ (ring rotates faster than Sun).
Outer rings: full sine curve. Negative velocities for $l > 0$ (ring rotates slower than Sun).
Max velocity at $R \to \infty$ is rotation speed of local standard of rest (LSR). The maximum velocity is $V_0 = R_0\omega_0$.





A: (1) Tangent point: max radial velocity in line of sight $V_{\max} = V_{\text{ring}} - V_0 \sin l$
(2) Points on solar circle: constant separation: $V_r = 0$
B: at $r \to \infty$, $\omega \to 0$ and $V_r = 0 - V_0 \sin l = -V_0$

Determining the rotational profile (inwards from the Sun's orbit, $0 < l < 90°$):
Tangent point maximum velocity along line of sight $V_{\max}$ measures $V_{\text{ring}} - V_0 \sin l$;
Determine R from $R = R_0 \sin l$;
Determine V_{ring} (rotation velocity) as function of R;
From tangent points, derive Galactic rotation velocity inside Sun's orbit.

We determine that the Milky Way has a flat rotation curve from $\sim 0.2 - 0.95 R_0$ ($1.5 - 8$ kpc). The tangential velocity roughly equal to $V_0 = 220$ km/s. The mass enclosed within an orbital radius R around the Galactic centre is $M(R) \approx V^2 R/G$. This tells us mass distribution must vary as $M(R) \propto R$ to produce the observed *Flat Rotation Curve* (V is constant with R). A flat rotation curve is *not expected* from Galactic disc if mass traces light.

Mass within R_0: Dynamical mass $\sim 10^{11} M_\odot$; Stellar + stellar remnant mass $\sim 4 \times 10^{10} M_\odot$; Interstellar medium (gas+dust) $\sim 2 \times 10^9 M_\odot$; Visible halo $\sim 10^8 M_\odot$; Central black hole $\sim 2.6 \times 10^6 M_\odot$. Hence we infer most of the mass of the Milky Way is dark.

Also, the flat rotation curve ($M(R) \propto R$) allows us to determine the dark matter distribution of the Galactic halo:

$$M(R) = 4\pi \int \rho R^2 \, dR \Longrightarrow \rho(R) = \rho_0 \left(\frac{R}{R_0} \right)^{-2} \text{where } \rho_0 = \frac{(V_0/R_0)^2}{4\pi G} \approx 0.01 M_\odot \text{pc}^{-3}$$

Density of dark matter falls off roughly as R^{-2} – known as a "singular isothermal sphere".

Possible forms of dark matter:
MACHOs (Massive Compact Halo Objects) or Black holes – Problem is that these would have to be massive and would disrupt Globular clusters;
Neutron stars $M \approx 1.4 M_\odot$ – Problem: formed from stars and too many metals would be produced;
Brown dwarfs $M < 0.08 M_\odot$ – Problem: similar numbers of stars and brown dwarfs formed today;
WIMPs (Weakly Interacting Massive Particles) or Neutrinos or Super-symmetric particles?

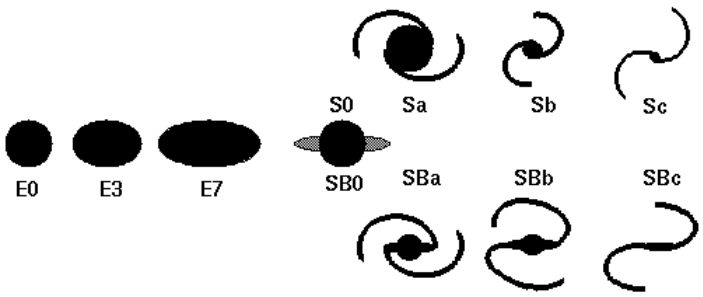
S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 4: External Galaxies

References: Carroll & Ostlie chapter 25 "The Nature of Galaxies"
Choudhuri sections 9.1 & 9.2;
Sparke & Gallagher chapters 5 & 6

Less than 100 years ago it was still unknown if our own Galaxy was the entire Universe, of if there are galaxies beyond our own. We now know that there are many galaxies seen out to great distances (over a large fraction of the age of the Universe). In 1929 Edwin Hubble used the shapes (morphologies) to suggest a classification scheme we still use today: most galaxies seem to belong to one of three distinct classes: smooth ellipticals; spirals with arms; and chaotic irregulars. The spirals are sub-divided into two branches (forks) depending on if they have a central bar of stars. Galaxy appearance can also depend on inclination, dust, observed wavelength and redshift (due to galaxy evolution).

Morphological Classification of Galaxies – the Hubble Tuning Fork



The amount of gas and the disk component become more important from left to right – ellipticals (E) have no disk, whereas spirals (S) and barred spirals (SB) have disks with spiral arms. Lenticular galaxies (S0) have a disk and prominent bulge but no spiral arms, and form a link from spirals to ellipticals (which are essentially "all bulge" with no disk component).

Elliptical (E) galaxies: the image on the sky is ellipitical, with no disc component. Given a number to indicate amount of elongation. Elliptical shape – projected axial ratio b/a; ellipticity $= 10 \times (1 - b/a) = 0 -$ 7, E0 (spherical) to E7 (cigar-shaped). Axis ratios from 1 to 3, i.e. major axis, a, $1 - 3$ times minor axis, b.

Spirals (S and SB) have subtypes a–c (Sa has tightly-wound arms and significant bulge, Sc loosely and disk dominates light). From a to c: Disc component becomes more important (as do gas & young stars); Bulge becomes smaller relative to disc; Arms become less tightly wound; Total luminosity decreases. Spirals account for about 2/3rds of the galaxies locally with luminosities similar to the Milky Way, with Ellipticals comprising 10% and Lenticulars 20%. However this depends on environment – in clusters of galaxies, ellipticals are much more common. Some ellipticals may form from the mergers of disk galaxies – this process would be more common when galaxies are closer (e.g. in clusters).

Irregular galaxies do not fit the "Hubble Tuning Fork" classification, and have no nucleus, no disk and no spiral arms. They are often gas rich, and can be actively star forming and often have blue young stellar populations. They typically have low mass and hence low luminosity, and examples include the Magellanic Clouds (satellite galaxies of the Milky Way). Luminous irregulars are rare today, but were more common early in the history of the Universe when galaxy collisions and mergers were frequent.

Many more luminous galaxies do exhibit peculiarities (tidal tails, low-surface-brightness shells, rings of star formation) which are evidence of galaxy interactions, such as mergers or dwarf galaxies being tidally-stripped as they fall in to larger galaxies.

The Colours and Masses of Galaxies

Each star has a colour measured, e.g., using $(U - B)$ or $(B - V)$. Stellar populations can be studied using their colours. Old populations only contain low-mass stars (high mass stars are short-lived and died a long time ago), and hence old stellar populations have the red colours of the low-mass, low-surface-temperature stars. Young populations contain some massive stars (which are hot, luminous and short-lived), and hence these young populations typically have bluer colours, although this can be reddened by dust.

Ellipticals are referred to as "early-type galaxies" and have old populations (red).

Spirals and irregulars are "late-type galaxies" and have young populations (blue).

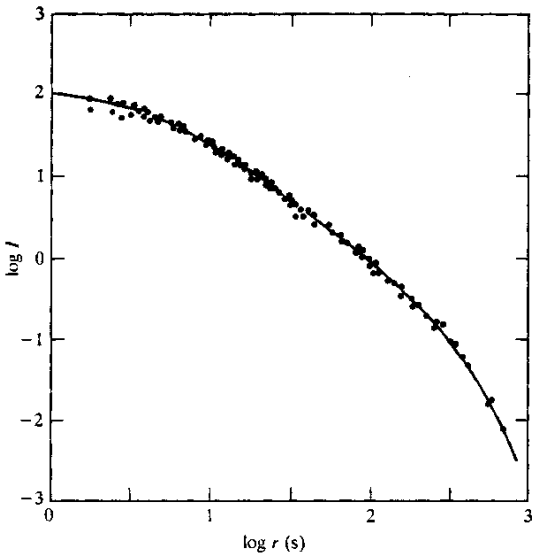
Examining the distribution of luminosity of a galaxy is the first step to identifying the components of the galaxy and their masses. We can make the simplistic assumption that 'mass follows light' to determine mass of each component. This is not the full picture (as seen with our Galaxy, most of the mass is dark), so we need to use kinematics along with the distribution of luminous matter.

Elliptical Galaxies have the simplest structure. They are smooth in shape and lack the structure of spirals: there is no current star formation (H II regions) and little gas & dust. They have primarily old stars (K-giants dominate today). Ellipticals are comparitively simple systems – one component; contours of equal surface brightness (isophotes) are concentric ellipses.There is an intrinsic distribution of ellipticities (the statistics are inconsistent with a single shape viewed at random angles), and ellipticals are triaxial (i.e. different axis length along 3 dimensions).

It used to be thought that these were axisymmetric ensembles of stars, flattened due to rotation – but we now know this is wrong: observations show Ellipticals essentially do not rotate globally (not enough to cause flattening). The otation velocity is less than or similar to the overall velocity dispersion in ellipticals, $\sigma > v_{\rm rot}$. This is unlike spirals, which have a flat stellar disk which is essentially all rotation with little velocity dispersion.

Ellipticals are found preferentially in denser galaxy environments (this is the "morphology–density relation"). They can be very luminous, but there are also dwarf Ellipticals (dE).

Surface brightness profiles (how the isophotes drop with distance from the centre).



Ellipticals display a pronounced maximum brightness at the centre of the galaxy, rapid and uniform decrease with distance as quasi-power law involving $R^{1/4}$ (De Vaucouleur):

$$I(R) = I_e e^{-7.67[(R/R_e)^{1/4}-1]}$$

This approximation to the profile becomes invalid close to and far from the centre. R_e Is the half-light radius (or effective radius), which encloses half the light of the galaxy (R_e can be up to $\sim$30 kpc). I_e is the surface brightness at R_e, sometimes quoted as $\langle I \rangle_e$ – the average surface brightness within R_e ; for an $R^{1/4}$ law, $\langle I \rangle_e = 3.6 I_e$
Total light intensity: $I_{\rm tot} = 2\pi R_e^2 \langle I \rangle_e$

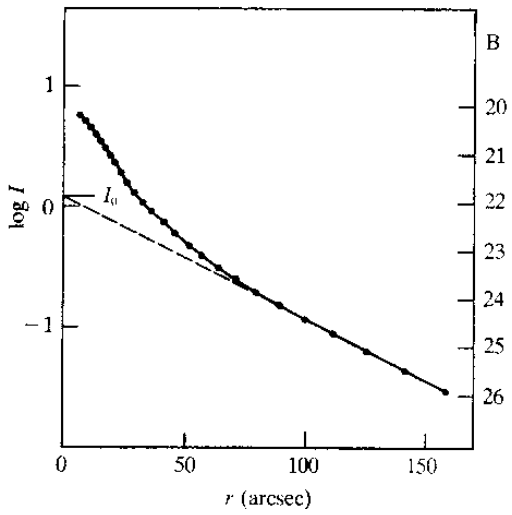
A more generic Sérsic $r^{1/n}$ profile has been developed, where $n = 4$ is the de Vaucouleur law typical of ellipticals and $n = 1$ is an exponential disk typical of spirals.

Scaling relations in Ellipticals: Kormendy (1977) found a link between the surface brightness of ellipticals and their radius. The de Vaucouleur's radius R_e is related to the surface brightness at this radius B_e (or μ_e) – that larger ellipiticals have lower densities.

The Faber-Jackson law links luminosity to velocity dispersion: $L \propto \sigma^4$. These four quantities (L, σ, R_e, B_e) form a two-dimensional family: Elliptical galaxies lie on a "fundamental plane" in this space. The origin of the fundamental plane not fully understood. It can, however, be used to measure the distances of galaxies (we know the apparent brightness and velocity dispersion, and hence infer the absolute luminosity and distance).

Spiral Galaxies contain a central bulge and a highly-flattened (thin) disc. Stars and gas are on near-circular orbits about centre in the plane of the disk (random motions are tiny) – disks are dynamically "cold". Spirals have multiple components (bulge – old; also halo and perhaps a thick disk too).

Central bulges of spirals resemble elliptical galaxies in morphology, luminosity profile, and stellar content (they have old stars), but rotation is more important in bulges than in ellipticals. Bulges follow the $R^{1/4}$ law along major axis, but not minor axis, and are more flattened than typical ellipticals, and are not as dense or luminous as elliptical galaxies.



Spiral Disc surface brightness: Freeman's Law (1970), exponential light profile $I = I_0^{-R/R_0}$ where I_0 is the central luminosity (extrapolated) and R_0 (sometimes called R_d or L) is the characteristic disk scale length, $R_0 \sim 1 - 5\,\mathrm{kpc}$. Not as clear a law as ellipticals due to contribution of the bulge, spiral structure, rings, etc. Stellar disks – present in Spiral (S,SB), Lenticular (S0) and some Irregular galaxies (Irr).

Dynamics of Spiral Galaxies

As we have seen from our Galaxy, the mass of a spiral Galaxy can be determined from its rotation curve.

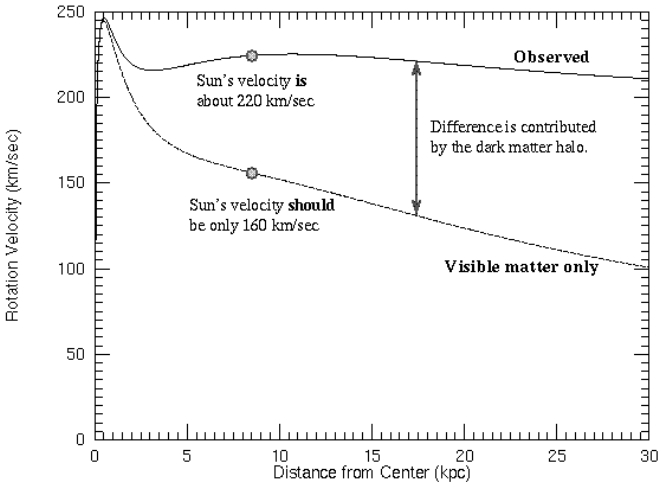
Stellar velocities: assume that the stellar velocity dispersion is much less than the rotation speed (true if the stars are in a disc).

Gas velocities: assume thermal and turbulent velocities are much less than the rotation speed (true if gas in a disc).

$M_{\mathrm{tot}} = V_{\mathrm{rot}}^2 R/G$ for a sphere, where the disc is cut off at $r = R$, V_{rot} is the rotation speed at radius R, and the mass is that within R.

Some aspects of spiral galaxies are easier to study in other galaxies than in our Milky Way (Spiral structure, rotation curve, bulge dynamics, etc.). In particular, global properties such as size, mass, rotation curve, etc. are easier to determine from the outside than from the inside.

Rotation curves can be determined by the Doppler shift of absorption lines in the spectra of stars, or emission lines from gas (such as the HI 21-cm line in the radio). Generally, the rotation speed of gas and stars are the same, within measurement errors, except for the bulge. Gas essentially has only rotational motions, centrifugally supported, and traces the mass distribution. Bulge stars have large velocity dispersion and lower net rotation, similar to elliptical galaxies. We use the simpler gas rotation curves to measure the mass distribution.



The gravity of the visible matter in the Galaxy is not enough to explain the high orbital speeds of stars in the Galaxy. For example, the Sun is moving about 60 km/sec too fast. The part of the rotation curve contributed by the visible matter only is the bottom curve. The discrepancy between the two curves is evidence for a **dark matter halo.**

Generally in spirals, rotation curves (plotting rotation speed against radius) rise rapidly (in less than 5 kpc) towards a maximum $V_{\max}$, and then are flat out to outer limits of galaxy, implies mass increasing with radius. The largest galaxies made only $\sim$10 outer rotations since formation.
As with our galaxy, other spiral galaxy rotation curves can be decomposed into bulge, disc, and dark matter halo components. Two components: visible matter & dark matter. Surprising that these combine to give flat rotation curves.

As luminosity increases, so does $V_{\max}$ and central velocity gradient. $V_{\max}$ is largest for *Sa* and decreases toward *Sc*, probably due to the bulge/disc ratio increasing towards early types (matter is more centrally concentrated, producing stronger rotation for given total mass). When normalising the rotation curves (to a characteristic radius & velocity), rotation curves and mass distribution are almost independent of type & luminosity for spiral galaxies.

Mass-to-light ratio in Spiral galaxies

Luminosity density (visible mass) decreases exponentially with radius, but the dynamical mass enclosed increases as $M(R) \propto R$.

Mass-to-luminosity (M/L) ratio $> 1 M_\odot/L_\odot$ (in solar units) and increases with radius – this cannot continue indefinitely. M/L is roughly constant with morphological type, with $M/L_{\rm vis}$ varying from ≈ 1 for *Sc* spirals to ≈ 3 for *Sa* spiral galaxies. Earlier types have older stars (fewer high-mass stars). Most mass is in low-mass stars, but most luminosity comes from high-mass stars. So the total mass M increases from late to early-type, and hence, $M/L_{\rm vis}$ increases with total mass M. Dwarf galaxies have proportionally more dark matter than massive galaxies.

Mass determination in Ellipiticals is harder – they do not contain much gas and do not rotate significantly. We assume in virial equilibrium, and use virial theorem (see later in course) to estimate the gravitational potential energy. Assume luminosity traces mass and using a de Vaucouleur's $R^{1/4}$ profile, the potential energy is then:

$W = -\dfrac{G\,M^2}{3R_e}$ where radius R_e contains half the luminosity (approximately half the mass)

Relations between the total luminosity of a galaxy, L, and its velocity dispersion σ or

rotation velocity, V: Tully-Fisher relation for spiral galaxies – $L \propto V^4$; Faber-Jackson relation for ellipticals – $L \propto \sigma^\beta$ with $\beta \approx 4$.

Best when we use infrared luminosity in the near-infrared H-band (i.e. L_H) since this traces the old population of stars which exist in all galaxies. Can use the Tully-Fisher law for distance measurement: velocity dispersion gives absolute luminosity, and can compare absolute to apparent luminosity to obtain distance modulus. Virial Theorem + a Universal Mass Distribution implies $M \propto V^4$. This gives the Tully-Fisher law if M/L is a constant (true for L_H, $M/L_H \approx 2M_\odot/L_\odot$). The origin of this universal law is not yet clear, but could be due to galactic systems being formed by collisions with sub-systems.

Spiral Arms and the Winding Problem

Most stars, especially massive ones, are formed in spiral arms when gas falls down potential well of spiral potential. The gas is compressed in shocks, molecular clouds undergo gravitational collapse, and form stars. New stars, HI and HII downstream of the shock. Spiral arms lead to angular momentum being transported outward in the disc (mass losing angular momentum and falling toward the centre), suggesting an evolution from Sc to Sa classes (growth of bulge) and the arms become more tightly-wound.

The Winding Problem: The rotation curve of spiral galaxies is flat and $V = R\omega = \text{constant}$, so $\omega = V/R$ and stars & gas further out in radius take longer to perform an orbit.

The outer edge of the galaxy may only have rotated a few times since formation, but there have been thousands of rotations near centre. Thus, spiral arms cannot be material arms because they would wind up too much compared with observation. Lindblad (1963) proposed spiral arms rotate as a solid body, with the same angular speed Ω_p (pattern speed) from centre to edge. Material rotates faster than the spiral near centre, but trails behind the spiral pattern at large radius. The pattern is maintained by the self-gravity of the stars.

Density wave theory explains existence of "grand design" spirals (long-lived, global). The spiral arms are regarded as waves that propagate like waves in the sea, and this pattern needs to be maintained for many mean rotations of galaxy. Observations of old & young stars show spiral arms present in both. Young (massive, blue) stars form along arms. Old (low-mass, red) stars show that the spiral exists for long times. A different short-lived spiral phenomenon is seen in 'stochastic' (or 'flocculent') spirals with filamentary arms.

Barred Spirals 2/3rds of spiral galaxies show some form of bar, and 1/3 show a strong bar. The bars are vertically thin (like disk) not round (like bulge). Bars are a few kpc across. Flow of gas in a barred potential naturally leads to trailing spiral arms in the gas. Bars supported by disk self-gravity, velocity dispersion σ (from randomly-ordered orbits) must be low. Bars composed of old (low-mass, red) stars. Active star formation primarily at ends of bar or in outer spiral arms. Our galaxy is thought to have a bar. Gas on elliptical orbits aligned along bar. Because the barred potential is not axisymmetric gas is subjected to tangential forces. Gas streams collide with each other, orbits changed and gas spirals in towards galactic centre.

The Galaxy Luminosity Function

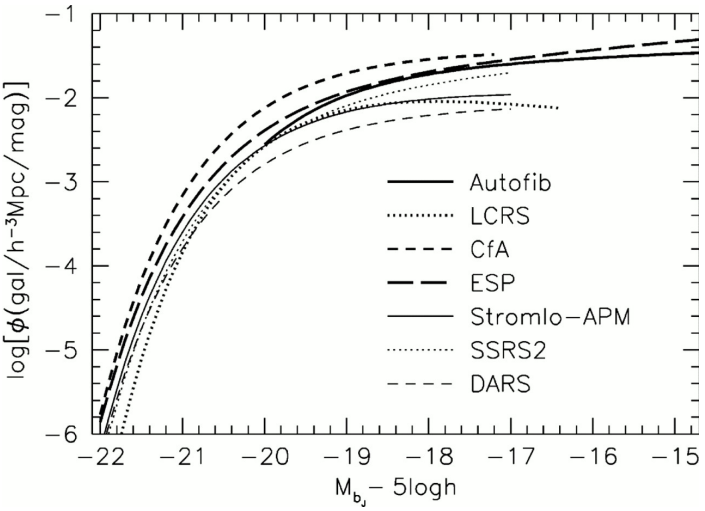
The distribution of galaxy luminosities is quantified by the luminosity function, $\phi(L)$. The number of galaxies δN located in a volume δV with luminosity in the range $(L, L+\delta L)$ is:

$$\delta N = \phi(L)\,\delta L\,\delta V$$

This assumes (unrealistically) a universal luminosity function, independent of location! In principle we should have a luminosity function that depends on position, but we are making two approximations. Firstly, we are assuming the luminosity function is separable from the density function $\phi(L, r) \propto \phi(L)D(r)$, and secondly we are assuming $D(r) \sim$ constant. In reality the luminosity function depends on morphology, environment, bandpass etc. The luminosity function also evolves with cosmic epoch... We need to know the luminosity function to interpret galaxy number counts as a function of apparent magnitude and redshift. The most common modern parameterization of the luminosity function is the Schechter function (1978). This is a power law truncated by an exponential (when expressed using linear flux units):

$$n(x)\delta x = \phi^* x^\alpha e^{-x}\delta x \quad \text{where } x = L/L^*$$

$L^* \sim 10^{10} L_\odot$ and $\phi^* \sim 0.01\,\mathrm{Mpc}^{-3}$ are the characteristic luminosity and number density at the "knee" or break of the luminosity function, and $\alpha \approx -1.5$ is the faint end slope.



S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 5: Stellar Structure

Syllabus for lectures 5-9: *The equations of stellar structure: hydrostatic equilibrium, virial theorem, convection and energy transport. Structure of main sequence stars; use of scaling relations to derive relationships between stellar masses, luminosities, radii and lifetimes. The Chandrasekhar limit and degenerate stellar cores; introduction to post-main sequence evolution.*

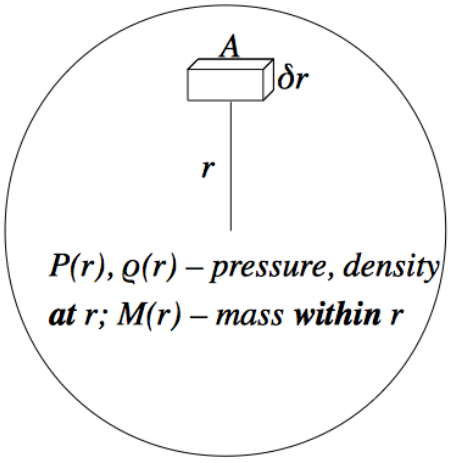
In the next 5 lectures we will study the physics of interiors of stars, and how this relates to stellar evolution. Stars usually evolve on long timescales, meaning that properties change slowly with time, and hence we can neglect rate of change of many properties for an approximate treatment.

Another approximation is that stars are spherical, and symmetric about the centre – in this approximation all physical quantities depend just on the distance from the centre, r. In reality stars may be slightly flattened through rotation, and have magnetic poles, but the spherical approximation is good enough for many purposes.

We also assume that stars can be approximated as being self-gravitating bodies in dynamic equilibrium. What does this mean? A condition called *hydrostatic equilibrium* is the balance between the inward pull of gravity and the outward push of internal pressure due to a pressure gradient across a shell of gas. We know from geological timescales that the Sun's output cannot have changed very greatly over millions or even billions of years, so its luminosity (which depends on radius and surface temperature) is almost constant, and hence it must be in near-equilibrium.

Derivation of hydrostatic equilibrium: Consider a slab at radius r with area A and thickness δr (so mass $= \rho\, A\, \delta r$).

If there is a gradient in the pressure with radius, the pressure: $P(r+\delta r) = P(r) + \delta P$
Difference in force on bottom and top: $F = P\,A - (P+\delta P)A = -\delta P\, A$ outward.



Force due to gravity:

$$F = \frac{G\,M\,m}{r^2} = \frac{G\,M(r)\,A\,\rho\,\delta r}{r^2}$$

where $M(r)$ is the mass within radius r. Horizontal pressure forces cancel. Balance net outwards pressure force against inward gravitational force.
Spherical symmetry: Newton's theorem (like Gauss's law) allows replacement of mass distribution by equivalent mass at centre just mass within r, $M(r)$.

$$\frac{\partial P}{\partial r} = \frac{G\,M(r)\,\rho(r)}{r^2} = \rho(r)\,g$$

There is a second important *equation of mass continuity*: Consider a spherical shell at radius r with thickness δr and density $\rho(r)$. Volume of shell $\delta V = 4\pi r^2 \delta r$ and mass of shell $M(r+\delta r) - M(r) = \delta M = 4\pi r^2 \rho(r) \delta r$. Rate of change of mass enclosed within radius r with r:

$$\frac{\partial M}{\partial r} = 4\pi r^2 \rho(r)$$

Estimate dynamical time ("free fall") from rest: use equation of motion $s = ut + \frac{1}{2} g\, t^2$:

$$t_{\rm dyn} \approx \sqrt{\frac{2R_S^3}{G\, M_S}} \approx \frac{1}{\sqrt{2\, G\, \bar{\rho}}} \approx 40\,{\rm min}$$

for the Sun (of course, g is not constant – only approximate).

Clearly, the Sun does not collapse on a dynamical timescale, but we can ask how finely matched the balance of pressure and gravitational force is. We start by considering the acceleration from net force, where ϵ is the fractional imbalance of pressure force and gravity: $\ddot{r} = \epsilon g$. Take $\epsilon \ll 1$ (i.e. forces nearly balance and acceleration small) so:

$$\rho\, \ddot{r} = \frac{\partial P}{\partial r} - \frac{G\, M\, \rho}{r^2} = \frac{\partial P}{\partial r} - \rho\, g = -\epsilon\, \rho\, g$$

Even if $\epsilon \sim 10^{-10}$ then the Sun's size changes by 10% in a few years, which is clearly not the case! (We know this from the eclipse record, where the angular sizes of the Sun and the Moon are similar for all recorded history; also climate and tree rings related to solar luminosity. Not much change for a billion years). Hence the approximation of hydrostatic equilibrium is extremely good, and forces almost exactly balance (to 1 part in 10^{18} or so).

Stars adjust very quickly to maintain an equilibrium between pressure and gravity.

Estimate of Central Pressure and Temperature

The Sun is slightly denser than water, but is at high temperature so the material is mainly ionized. The mean density of the Sun is:

$$\bar{\rho} \approx \rho_r = \frac{3M_S}{4\pi R_S^3} \approx 1400\,{\rm kg\,m^{-3}}$$

We can make a crude estimate of pressure P and temperature T in a stellar core assuming a constant gradient of pressure $dP/dr \sim -P_C/R_S$ (where P_C is central pressure, $P = 0$ at surface and R_S is stellar radius). Take a point at $r = R_S/2$, and assume this encloses about half the total mass M_S, $M_r \sim M_S/2$ (a star is centrally-condensed so this is a better approximation than uniform density)

$$\text{which gives } P_C = \frac{3\, G\, M_S^2}{2\pi R_s^4} \approx 5 \times 10^{14}\,{\rm N\,m^{-2}} = 5 \times 10^9\,{\rm atm}$$

We can also get a *lower limit* on the central pressure, using $r < R_S$ within the star:

$$\frac{dP(r)}{dr} \div \frac{dM(r)}{dr} = \frac{dP}{dM} = -\frac{G\,M}{4\pi r^4} \text{ hence } \int_0^{M_S} \frac{G\,M}{4\pi r^4} dM > \int_0^{M_S} \frac{G\,M}{4\pi R_S^4} dM = \frac{G\,M_S^4}{8\pi R_S^4}$$

$$\text{so } (P_C - P_S) \approx P_C > \frac{G\,M_S^2}{8\pi R_s^4}$$

To estimate the central temperature, we assume the ideal gas law: $P_r = \rho_r\, k_B\, T/(\mu\, m_H)$.

If the gas is ionized, the "mean molecular weight" is 1/2 the mass of a proton ($m_H/2$), so $\mu \approx 1/2$ ("mean molecular weight" is actually the average particle mass in atomic mass units – there are no molecules).

$$\frac{\partial P}{\partial r} = -\frac{G\,M_r\,\rho_r}{r^2} \approx -\frac{P_C}{R_S}$$

We can make the approximation that the pressure gradient is constant. Hence for Sun:

$$T_C \approx \frac{2G\,M_S\,\mu m_H}{k_B\,R_S} \approx 2\times 10^7\,\text{K and } k_B T_C \sim 2000\,\text{eV, so H is ionized } (\gg 13.6\,\text{eV}).$$

The above estimate involved some major approximations about the pressure gradient and density of a star, but we can get a lower limit on the core temp from <u>*the Virial theorem*</u>:

$$T_C > \frac{G\,M_S\,\mu m_H}{6k_B\,R_S},\ T_C > 2{\times}10^6\,K \text{ where } \int_0^{M_S} \frac{G\,M}{r}\,dM > \int_0^{M_S} \frac{G\,M}{R_S}\,dM = \frac{GM_S^2}{2R_S} \text{ see below}$$

The virial theorem states:

$$E_{\text{TOT}} = E_{\text{KIN}} + E_{\text{GRAV}} = U + \Omega \text{ and } 2E_{\text{KIN}} = -E_{\text{GRAV}} \Longrightarrow E_{\text{TOT}} = -E_{\text{KIN}} = \frac{1}{2}E_{\text{GRAV}}$$

Proof for stars: multiply hydrostatic equilibrium equation by $4\pi r^3$ and recognise integration by parts:

$$\frac{\partial P}{\partial r} = -\frac{G\,M_r\,\rho_r}{r^2} \rightarrow 4\pi r^3\,dP = -\frac{G\,M_r}{r} 4\pi r^2 \rho_r\,dr \text{ gravitational energy}$$

$$4\pi[r^3 P_r]_{r=0,P=P_C}^{r=R_s,P=0} - 3\int_0^{R_S} P_r 4\pi r^2 dr = -\int_0^{R_S} \frac{G\,M_r}{r} 4\pi r^2\,\rho_r\,dr$$

$$\int_0^{R_S} P_r\,4\pi r^2 dr = \int_0^V P\,dV \Longrightarrow -3\int_0^V P\,dV = -\int_0^{R_S} \frac{G\,M_r}{r} 4\pi r^2 \rho_r\,dr = \int_0^{M_S} \frac{G\,M}{r}\,dM$$

Now consider thermal (kinetic) energy per unit volume, u, with f degrees of freedom and ratio of heat capacities $\gamma = C_P/C_V$:

$$u = \frac{1}{2} f\, n\, k_B\, T = \frac{f\, \rho_r}{2\mu\, m_H} k_B\, T = \frac{\rho_r\, k_B\, T}{\mu\, m_H\, (\gamma - 1)} = \frac{P}{(\gamma - 1)}$$

so $3(\gamma - 1)U + \Omega = 0$ and for fully-ionized ideal gas ($f = 3$ and $\gamma = 5/3$), $2U + \Omega = 0$ and $E_{\rm TOT} = -U = \Omega/2$

Implications of Virial Theorem: $E_{\rm TOT} = -U$ stars have negative heat capacity (heat up when energy drops). If no energy source (e.g. no nuclear fusion) then stars contract and heat up as they radiate energy from the surface.

Has the effect of self-regulating nuclear burning through negative feedback (in non-degenerate stellar cores) – increase in nuclear burning would cause expansion and cooling, suppressing reactions.

If no nuclear processes, how long would a star take to radiate away its thermal energy?

$t_{\rm th} \approx E_{\rm TOT} = \dfrac{\Omega}{2L_S} \approx \dfrac{G\, M_S^2}{2R_S L_S} \approx 1.5 \times 10^7$ years (for the Sun), Kelvin-Helmholtz timescale

So gravitational energy insufficient to sustain output for $> 10^9$ yr. Stars lose energy by radiation (luminosity from the surface), need an energy source to prevent collapse if supported by thermal pressure. Need nuclear fusion to account for luminosity and lifetime.

Energy Sources

ϵ_r is energy generation rate (nuclear) per unit mass per second, for a shell at radius r (depends on T_r , ρ_r and composition)

$$L_S = \int_0^{R_S} 4\pi r^2\, \rho_r\, \epsilon_r\, dr \Longrightarrow \frac{dL}{dr} = 4\pi r^2 \rho_r \epsilon_r$$

We now have 3 key equations, for dL/dr, dP/dr and dM/dr

$$\frac{\partial P}{\partial r} = -\frac{G\, M_r\, \rho_r}{r^2} \ , \ \frac{\partial M}{\partial r} = 4\pi r^2 \rho_r \text{ and } \frac{\partial L}{\partial r} = 4\pi r^2 \rho_r \epsilon_r$$

There is a relation for $\partial T/\partial r$ which we will encounter later when we in study energy transport. Together these are the 4 equations of stellar structure.

References: Carroll & Ostlie chapter 10 "The Interior of Stars"
Choudhuri section 3.2

S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 6: Energy Transport in Stars

References: Carroll & Ostlie chapter 10 "The Interior of Stars"

Summary of Lecture 6: Energy transport in stars by radiation and convection. Opacity and optical depth. Stellar models (Choudhuri 3.3). Ideal gas, and mean particle mass (concept of X, Y, Z abundance). Scaling relations (Choudhuri 3.4) e.g. H–R (Hertzsprung–Russell) main sequence, mass–luminosity relation.

Methods of energy transport – There are three energy transport mechanisms:
1) Convection – bulk motion of gas
2) Conduction – exchange of energy during collisions (usually electrons)
3) Radiation – emission and absorption of photons transports energy
Within stars, the radiation field is mainly X-ray photons which are in thermal equilibrium with the (ionized) gas particles. A typical particle has thermal energy $\approx 3k_BT/2$ which is similar to the X-ray photon energy hc/λ

Bound-free absorption (ionization of inner electrons) means inside of stars are opaque to X-rays (i.e. mean free path length is short, $\sim 1\,$cm, but longer than that of the actual particles at high gas density):
mean free path of photon $\sim 10^{-2}\,$m $\gg$ mean free path of particle $\sim 10^{-10}\,$m (size of atom).
Photons can move across temperature gradients more easily, hence larger transport of energy. Conduction is negligible, EXCEPT in the unusual case of degenerate matter, such as white dwarf stars, where conduction by electrons is very effective. Radiation transport usually dominates, where photons "random walk" and diffuse out from centre of star; energy is degraded from X-ray to optical frequencies through repeated scatters.

Time for a photon to random walk to surface: $t \sim 3R_S^2/(c\,d) \approx 3 \times 10^4\,$years for the Sun ($d$ is the mean free path, and c the speed of light).

Energy Transport by Radiative Transfer

Radiation pressure is: $P_{\rm rad} = \frac{1}{3} a\,T^4$ where $a = 4\sigma/c$ is the radiation density constant. Where does this expression come from? It is derived from the flux of momentum, where photon momentum $p = E/c$ and flux component normal to a surface dA is $dE\,\cos\theta\,/c$. Divide this by $dAdt$ to get momentum flux per unit area (I_ν is the intensity per solid angle, where $d\Omega = sin\theta\,d\theta\,d\phi$):

$$\frac{dE\,\cos\theta}{c\,dA\,dt} = \frac{I_\nu}{c}\cos^2\theta\,d\Omega \;\text{ integrate for } P_{\rm rad}:\; P_{\rm rad} = \frac{1}{c}\int I_\nu \cos^2\theta\,d\Omega = \frac{4\pi I_\nu}{3c}$$

$$\text{And total energy } U_\nu = \int \frac{I_\nu}{c} d\Omega = 4\pi\frac{I_\nu}{c} \text{ so } P_{\rm rad} = \frac{U_\nu}{3} = \frac{a\,T^4}{3}$$

Consider rate at which momentum is deposited in a shell of radius r, width δr:

$$-\frac{dP_{\rm rad}}{dr}4\pi r^2 dr$$

The rate of momentum absorption of radiation in shell $= L(r)\,\kappa\,\rho\,dr/c$ where the opacity κ (sometimes called χ) is the attenuation of the light intensity (units are m^2/kg).

$\frac{dI}{I} = -\kappa\,\rho\,dx$ and $I = I_0 e^{-\tau}$ where optical depth $\tau = \int \kappa\,\rho\,dx$, $\frac{1}{\kappa\,\rho} =$ mean free path.

Equating the rate of momentum absorption with momentum deposition:

$$\frac{L_r}{c}\kappa\,\rho\,dr = -\frac{P_{\rm rad}}{dr}4\pi r^2 dr = 4\pi r^2 dr\frac{d}{dr}\left(\frac{a\,T^4}{3}\right)$$

This gives us a $4^{\rm th}$ equation of stellar structure:

$$L_r = -T_r^3\frac{dT_r}{dr}4\pi r^2\frac{4\,a\,c}{3\,\kappa_r\,\rho_r} \text{ or } \frac{dT_r}{dr} = -\frac{3\,\kappa_r\,L_r\,\rho_r}{16\,\pi\,a\,c\,r^2\,T_r^3}$$

which is appropriate for regions within stars where energy transport is mainly by radiative transfer. We need a different expression for convection, to be derived later in this lecture.

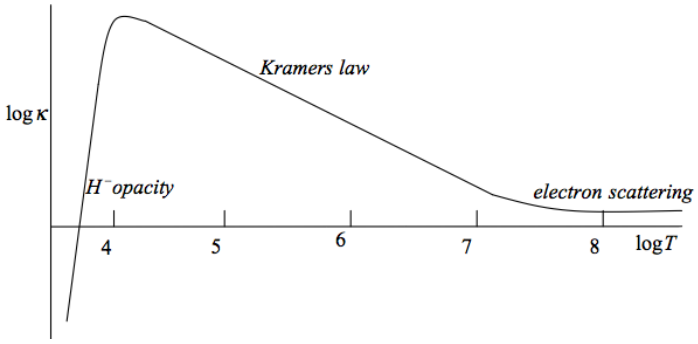
We now have 4 differential equations for the internal structure of stars (in the absence of convection). Where we have 7 unknowns as a function of radius, r : P_r = pressure at r; M_r = mass of material within r; ρ_r = density at r; T_r = temperature at r; L_r = luminosity at r (rate of energy flow across sphere of radius r); κ_r = Rosseland mean opacity at r; ϵ_r = energy release per unit mass per unit time.

So 7 unknowns but only 4 equations cannot be solved immediately. But we have some other relations:
$P = P(\rho,\ T,$ chemical composition) – this is the Equation of State (e.g., ideal gas law); $\kappa_R = \kappa_R(\rho,\ T,$ chemical composition); and $\epsilon = \epsilon(\rho,\ T,$ chemical composition).

We can impose some Boundary Conditions: At the centre of the star $M_r = 0$, $L_r = 0$ at $r = 0$. Stars do not have a sharp edge (no solid surface). Usually take "surface" to be photosphere – that is where the optical depth rises above ≈ 1 and where most of the photons we see come from. For the Sun, the density of the photosphere is $10^{-4}\,\mathrm{kg\,m^{-3}}$, which is much smaller than mean density of the Sun ($\approx 1.4\times10^3\,\mathrm{kg\,m^{-3}}$) so we can take the density and pressure at the surface to be negligible. Know the surface temperature ($T_{\rm eff} = 5780\,\mathrm{K}$) is much smaller than its minimum mean temperature ($2\times10^6\,\mathrm{K}$). Surface boundary conditions: $P = \rho = T = 0$ at $r = R$ (or more precisely $T = T_{\rm eff}$), i.e. that the star does have a sharp boundary with the surrounding vacuum

Opacity: The opacity κ (called χ in Choudhuri) governs rate of energy flow by radiative transfer. It is the interaction cross-section per unit mass. Opacity has several sources: bound-free, free-free absorption and scattering by electrons. Sometimes use opacity averaged over frequency (Rosseland mean opacity).

At high temperature, opacity is constant $\kappa_1 = 0.02(1+X)\,\mathrm{m}^2/\mathrm{kg}$
At intermediate temperatures, Kramer's law $\kappa = \kappa_2 \rho T^{-3.5}$
At low temperature $\kappa = \kappa_3 \rho^{\frac{1}{2}} T^4$
The κ_1 , κ_2 , κ_3 are constants which depend on chemical composition. X is mass fraction of hydrogen.



Equation of State – For an ideal gas, the equation of state is:

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

where N is the number density of particles and μ is the mean particle mass (in terms of m_H, protons). [IF THE GAS IS DEGENERATE, THE EQUATION OF STATE IS VERY DIFFERENT – see later in the course, e.g. white dwarf stars, lecture 9].
X=mass fraction of hydrogen; provides 1 electron per atom
Y=mass fraction of helium; provides 2 electrons per atom
Z=mass fraction of metals (heavier elements); provide on average $A/2$ electrons per atom, where A is the average atomic weight (in atomic mass units, a.m.u.)

If fully ionized (good approximation inside stars except near surface):
Number density of particles $N = (2X + \frac{3}{4}Y + \frac{Z}{2})\rho/m_H = (6X + Y + 2)\rho/(4\,m_H)$
Note: $X + Y + Z = 1$

$$\text{So: } \frac{1}{\mu} = \frac{6X+Y+2}{4} \text{ and if } Z \text{ is negligible, } Y = 1-X, \quad \mu = \frac{4}{3+5X}$$

Mass dependency of radiation to gas pressure

Previously we assumed that the radiation pressure was negligible compared to gas pressure. Is this valid for all stars? The pressure exerted by photons on the particles in a gas is:

$$P_{\mathrm{rad}} = \frac{a\,T^4}{3} \text{ and the gas pressure is } P_{\mathrm{gas}} = \frac{k_B\,T\,\rho}{\mu\,m_H}$$

Consider the ratio of these components of the pressure at a point within a star:

$$\frac{P_{\mathrm{rad}}}{P_{\mathrm{gas}}} = \frac{a\,T^4}{3} \Big/ \frac{k_B\,T\,\rho}{\mu\,m_H} = \frac{\mu\,m_H\,a\,T^3}{3\,k_B\,\rho}$$

In the Sun, gas pressure dominates over radiation pressure: If $\bar{T} = 2\times10^6\,\mathrm{K}$, $\bar{\rho} = 1.4\times10^3\,\mathrm{kg\,m^{-3}}$ and $\mu\,m_H = \frac{1}{2}\times1.67\times10^{-27}\,\mathrm{kg}$ gives $P_{\mathrm{rad}}/P_{\mathrm{gas}} \sim 10^{-4}$. However P_{rad} does become significant in higher mass stars. Substituting in for ρ:

$$\frac{P_{\mathrm{rad}}}{P_{\mathrm{gas}}} = \frac{\mu\,m_H\,a\,T^4}{3k_B\left(\frac{3M_S}{4\pi R_S^3}\right)} = \frac{4\pi\mu m_H a\,R_S^3 T^3}{9k_B M_S} \text{ and from the Virial Theorem: } \bar{T} \propto \frac{M_S}{R_S} \Rightarrow \frac{P_{\mathrm{rad}}}{P_{\mathrm{gas}}} \propto M_S^2$$

So the most massive stars ($>\sim 100 M_\odot$) are supported against collapse by radiation pressure. Relations among stellar quantities (see Choudhuri section 3.4): similar to dimensional analysis ("homology relations" offer a slightly more sophisticated approach, but beyond this course).

Starting with hydrostatic equilibrium

$$\frac{dP(r)}{dr} = -\frac{G\,M(r)\,\rho(r)}{r^2} \quad \text{Replace} \frac{dP}{dr} \approx -\frac{P}{R}\ (P \text{ is typical pressure within star}) \text{ and } \frac{M_r}{r^2} \approx \frac{M_S}{R_S^2}$$

$$\frac{P}{R} \propto \frac{M}{R^2}\rho \Longrightarrow P \propto \frac{M^2}{R^4} \text{ since } \rho \propto \frac{M}{R^3}. \text{ From the ideal gas law } P \propto \rho T \Longrightarrow P \propto \frac{M}{R^3}T$$

(neglect variation in μ). And for both of these to be valid: $T \propto M/R$ (temperatures inside stars scale with M/R).

Justifying the Mass-Luminosity Relation

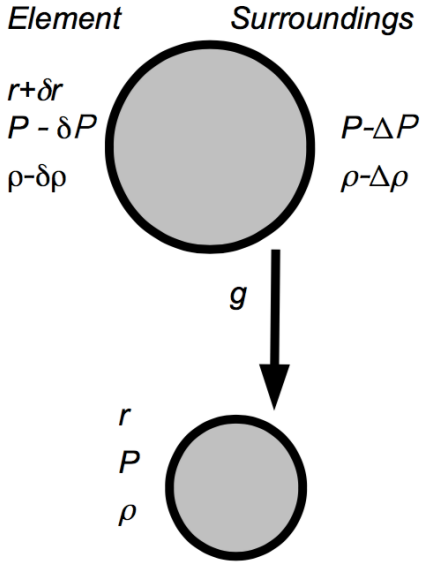
For the equation of radiative transfer, if the opacity, κ, is constant (as it is at high temperatures, so in massive stars)

$$\frac{dT_r}{dr} = -\frac{3L_r\kappa_r\rho_r}{16\pi r^2 a\,c\,T_r^3} \Longrightarrow \frac{T}{R} \propto \frac{L\,\rho}{R^2T^3} \propto \frac{L\,M}{T^3R^5}$$

$$\text{so, using } T \propto \frac{M}{R}: \ L \propto \frac{(T\,R)^4}{M} \Longrightarrow L \propto M^3$$

This echoes our observations of the Main Sequence on the H–R diagram! (Lecture 2). Also, using effective temperature and luminosity: $L = 4\pi R_S^2 \sigma T_{\rm eff}^4 \propto R^2T^4$ we can obtain $M \propto T^2$ and $L \propto T_{\rm eff}^6$ for the Main Sequence (MS) on the H–R diagram, and MS lifetime $t_{\rm MS} \propto M/L \propto M^{-2}$, appropriate for high mass stars on the Main Sequence.

Energy Transport by Convection



We previously considered radiative transfer, where energy is transported without movement of the matter. In contrast, convection is the mass motion of gas elements, and in some stars this can be the dominant mechanism for energy transport – when the temperature gradient exceeds some critical value.

Consider an element of gas (constant mass) at distance r from the centre of star. We assume this element is in equilibrium with its surroundings. If, as it rises to $r + \delta r$, it expands then the pressure and density fall to $P - \delta P$ and $\rho - \delta\rho$. If gas element is denser than surroundings at $r + \delta r$ (which have density $\rho - \Delta\rho$ and pressure $P - \Delta P$) then it will sink (i.e. stable). If it is less dense then it will keep on rising – it is *convectively unstable*.

Condition for convective instability: $(\rho - \delta\rho) < (\rho - \Delta\rho)$. Governed by:
(a) The rate at which the element expands due to decreasing pressure;
(b) The rate at which the density of the surroundings decreases with height.
If the speed that a convective element rises is much less than the sound speed, then the pressure will be that of the surroundings at that radius (sound waves smooth out pressure differences), i.e. pressure equilibrium, so $(P - \delta P) = (P - \Delta P)$. We will assume the element rises adiabatically: $PV^\gamma = $constant, where $\gamma = C_P/C_V$ (ratio of heat capacities at constant pressure, to constant volume)

$$\text{and } \rho \propto \frac{1}{V} \text{ so } \frac{P}{\rho^\gamma} = \text{constant. So } \frac{P}{\rho^\gamma} \text{ is the same at } r \text{ and } r + \delta r : \quad \frac{P - \delta P}{(\rho - \delta\rho)^\gamma} = \frac{P}{\rho^\gamma}$$

Doing a binomial expansion: $(\rho - \delta\rho)^\gamma = \rho^\gamma(1 - \frac{\delta\rho}{\rho})^\gamma \approx \rho^\gamma - \gamma\,\rho^{\gamma-1}\delta\rho$.
So $\delta\rho = \frac{\rho}{\gamma P}\delta P$ for the element. For the surroundings, change in density: $\Delta\rho = \frac{d\rho}{dr}\delta r$.

$$\text{Hence } \frac{\rho}{\gamma P}\delta P < \frac{d\rho}{dr}\delta r \text{ and since } \delta r \to 0, \ \frac{\delta P}{\delta r} = \frac{dP}{dr} \text{ and so } \frac{\rho}{\gamma P}\frac{dP}{dr} < \frac{d\rho}{dr}$$

for convection to occur (assumes pressure of element same as surroundings). A gas is convectively unstable if the actual temperature gradient is steeper than the adiabatic gradient. Note that dP/dr is negative, hence the inequality sign must change (or use modulus convention).

Schwarzschild Stability Condition for Convection

$$\text{Dividing previous equation by } \frac{dP}{dr} : \ \left(\frac{P}{\rho}\right)\frac{d\rho}{dP} < \frac{1}{\gamma} \text{ and for } P = \frac{\rho\, k_B\, T}{\mu\, m_H} \Rightarrow \frac{dP}{p} = \frac{d\rho}{\rho} + \frac{dT}{T}$$

(by differentiating logs and assuming ideal gas & negligible radiation pressure).
Hence: $\ln P = \ln\rho + \ln T + $constant.

$$\text{For convection to occur: } \frac{P}{T}\frac{dT}{T} > \left(\frac{\gamma - 1}{\gamma}\right) \text{ Alternatively, } \left|\frac{dT}{dr}\right| > \left(\frac{\gamma - 1}{\gamma}\right)\frac{T}{P}\left|\frac{dP}{dr}\right|$$

$$\text{So a layer of a star is stable AGAINST convection if: } \left|\frac{dT}{dr}\right| < \left(\frac{\gamma - 1}{\gamma}\right)\frac{T}{P}\left|\frac{dP}{dr}\right|$$

(this is the *Schwarzschild stability condition* from 1906). If convective, large scale rising and falling motions transport energy upwards. Convection possible if temperature gradient is large (e.g. nuclear reactions in some cores), or if γ is close to unity (e.g. cooler outer layers of some stars, partially ionized). Convection is complicated to model – needs "mixing length theory". Poorly understood and a major obstacle to understanding stellar evolution.

Implications of Convection:

- Turbulent motions (slow – do not upset hydrostatic equilibrium).
- Convection is a very efficient energy transport mechanism, as the particles have high thermal energy.
- Convection mixes material so the chemical composition of a convective zone is very homogeneous.
- The difference between the actual temperature gradient and the adiabatic value is very small.

$$\text{Close to neutral buoyancy: } \left|\frac{dT}{dr}\right| \approx \left(\frac{\gamma - 1}{\gamma}\right) \frac{T}{P} \left|\frac{dP}{dr}\right|$$

- Convection important in some regions of some stars.
- Where there are partially-ionized zones (such as cooler outer layers of some stars), much of the energy goes into ionization rather than heating up the gas, so $C_P \approx C_V$ and $\gamma \approx 1$ and so these are often convective zones.
- Where there are steep temperature gradients (e.g. the CNO cycle of massive stars on the Main Sequence, or helium-burning by the triple-α process post-main-sequence, have a very strong temperature dependence and will produce convective mixing).

S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 7: Energy Generation in Stars

References: Carroll & Ostlie chapter 10.3 "Stellar Energy Sources"
Summary of Lecture 7: Binding energy, nuclear reaction rates (Choudhuri 4.1). PP-I chain, CNO cycle, Helium triple-alpha (Choudhuri 4.2). Chemical enrichment, alpha-elements.

Summary of Stellar Interiors *(the story so far...)*

Stars are self gravitating bodies in hydrostatic equilibrium (internal pressure balanced by gravity for each radial shell), $dP_r/dr = -GM_r\rho_r/r^2$

Stars radiate energy from their surfaces at $L = 4\pi R_S^2 \sigma T_{\rm eff}^4$. If supported by thermal pressure (not degeneracy pressure) need an energy source to avoid collapse. Nuclear energy production rate per unit mass, ϵ: $dL_r/dr = 4\pi r^2 \rho_r \epsilon_r$.

Normal stars have "negative heat capacity" from the virial theorem, so if total energy decreases (e.g. nuclear energy stops) they contract and heat up.

Nuclear burning is self-regulating (in non-degenerate cores): increase in nuclear burning rate would cause core to expand and cool (virial theorem) reducing nuclear reaction rate, a negative feedback loop and so stable nuclear burning.

Internal temperature structure, T_r, mainly depends on type of energy transport (how energy gets from the nuclear-burning centre to the photosphere "surface").

Three energy transport mechanisms:
1) conduction (important in degenerate matter such as white dwarfs);

$$\text{2) radiative transport where: } \frac{dT_r}{dr} = -\frac{3L_r\kappa_r\rho_r}{16\pi r^2 a\, c\, T_r^3}$$

$$\text{3) or convection where } \frac{P}{T}\frac{dT}{dP} > \left(\frac{\gamma-1}{\gamma}\right)$$

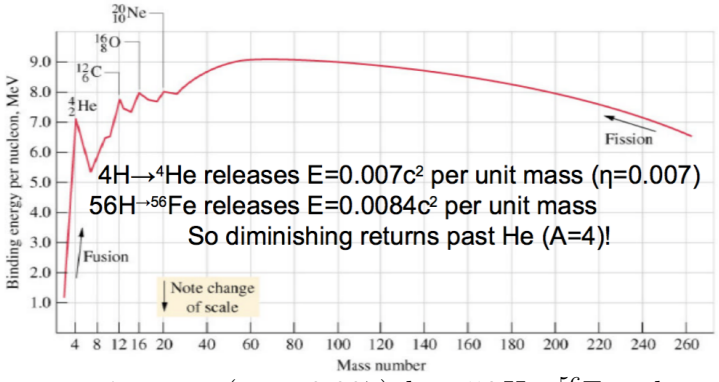
Global stellar structure determined by satisfying the above equations at each radius (i.e. the observed properties L, $T_{\rm eff}$, M, R are set by internal structure). In this lecture we will consider the energy generation mechanism within stars, nuclear fusion.

Nuclear Binding Energy

Binding energy of a nucleus with atomic number Z (= number of protons) and mass number A (= number of nucleons = protons + neutrons):
$Q(Z,N) = [Zm_p + Nm_n - m(Z,N)]c^2$

Energy less than constituent nucleons - a mass defect (ΔM), equivalent to net energy released in building nucleus ($E = mc^2$). Useful to specify the binding energy per nucleon: $Q(Z,N)/A$

Binding energy peaks at iron, $A = 56$. Lighter elements can potentially release energy by fusion, heavier elements by fission (splitting). Big energy release per unit mass from H to He, then progressively less as other light elements are formed.



$4\,\mathrm{H}\rightarrow{}^{4}\mathrm{He}$ releases $E = 0.007c^2$ per unit mass ($\eta = 0.007$) but $56\,\mathrm{H}\rightarrow{}^{56}\mathrm{Fe}$ releases $E = 0.0084\,c^2$ per unit mass. So diminishing returns past He ($A = 4$)!

Nuclear Reaction Rates

The rate of reaction between two types of particle is proportional to the number density of particles n_1 AND the number density of particles n_2. The frequency of collisions $= v\sigma(v)$, where $\sigma(v)$ is the cross section for interaction, and v is the relative velocity of the particles. Reaction rate: $R_{1+2} = n_1 n_2 \langle v\sigma(v)\rangle$. Velocity v is given by Maxwellian distribution of speeds,

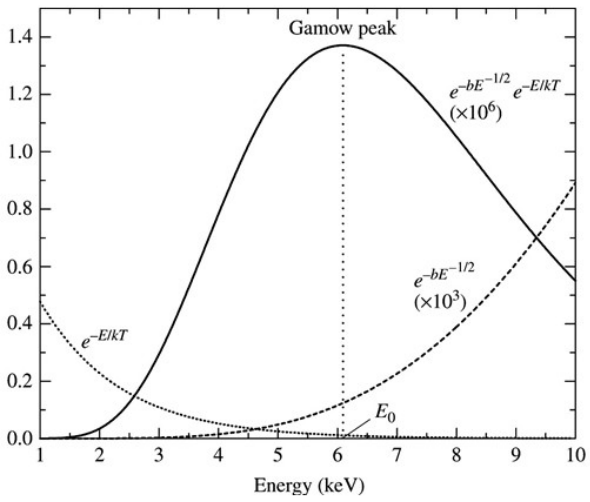
$$f(v)dv = 4\pi v^2 dv \left(\frac{m}{2\pi k_B T}\right)^{\frac{3}{2}} e^{\frac{-mv^2}{2k_B T}}$$

To achieve fusion we have to get nuclei sufficiently close ($\sim 10^{-15}$ m) that the short-range strong nuclear force dominates over electrostatic repulsion. The cross section σ depends on probability of penetrating Coulomb barrier, $P_p(v)$ or $P_p(E)$. Define cross-section factor S(E): $\sigma = P_p(E)\,S(E)/E$, where $S(E)$ depends on details of nuclear reactions but is relatively insensitive to particle energy or velocity (expect for nuclear resonances). $S(E)$ is determined from laboratory measurements.

There is a problem that the electrostatic repulsion for two nuclei is $F = Z_1 Z_2/(4\pi\epsilon_0 r^2)$ must be overcome, with the potential energy $E = Z_1 Z_2/(4\pi\epsilon_0 r) = 2.3\times10^{-13}$ J =1.4 MeV for two protons at $r = 10^{-15}$ m ($Z_1 = Z_2 = 1$ for hydrogen fusion).

Taking the average energy per proton to be $3k_BT/2$ this implies $T \approx 10^{10}$ K for H fusion – much hotter than estimate of Sun's core temp! The solution is *Quantum Tunnelling*. Nuclear fusion proceeds through quantum mechanical tunnelling through the Coulomb barrier. The tunnelling probability is approximated by: $P_{\text{tunnel}} \propto e^{-\sqrt{E_G/E}}$, where E_G is the Gamow Energy: $E_G = (\pi\alpha Z_A Z_B)^2 2m_r c^2$ and $m_r = m_A m_B/(m_A + m_B)$ is the reduced mass and $\alpha = e^2/(4\pi\epsilon_0\hbar c) \approx 1/137$ (fine structure constant).

Figure credit: Figure 10.6, Introduction to Modern Astrophysics, 2nd ed., Carroll & Ostlie, 2006.



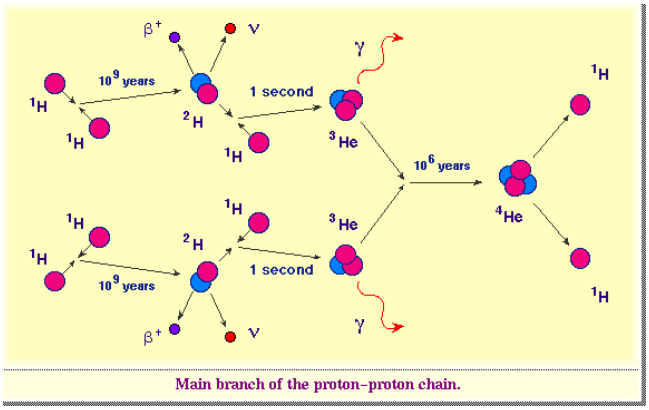
Dashed lines: competition between Tunnelling factor, $P_{\rm tunnel} \propto exp(-\sqrt{E_G/E})$ and Maxwell-Boltzmann factor, $P_{\rm Boltz} \propto exp(-E/k_BT)$. Solid line: Product gives the reaction rate, peaks at about $E_0 = (k_BT/2)^{2/3}E_G^{1/3}$ – *the Gamov Peak*. For the Sun: $E_G = 493\,$keV; $T = 1.6 \times 10^7\,$K $\Rightarrow k_BT = 1.4\,$keV. Hence $E_0 \approx 6.2\,$keV ($\approx 4.4k_BT$). Actually $E_0 = 15 - 30\,$keV for $k_BT = 1 - 3\,$keV, and width of Gamov peak $\Delta = 4 - 10\,$keV.

Energy generation per unit volume: $\epsilon\,\rho = R_{12}\Delta E = n_1 n_2 \langle v\sigma\rangle \Delta E$ and so $\epsilon = C\rho X_1 X_2 T^{-2/3} \exp[-3(m\pi^2 2e^4 Z_1^2 Z_2^2/8\epsilon_0^2 k_B h^2 T)^{1/3}]$ (see Choudhuri section 4.3).

The PP chain: the main branch (PP-I) – primary fusion mechanism in the Sun

Protons (ionized hydrogen) fuse to form helium-4 in stages from close encounters of pairs of nucleons. The most common in low mass stars is called the PP chain (proton-proton chain reaction).

1) $p + p \longrightarrow {}^2D + e^+ + \nu_e + 1.44\,$MeV; typical reaction time $1.4 \times 10^{10}\,$yr BOTTLENECK
2) ${}^2D + p \longrightarrow {}^3He + \gamma + 5.49\,$MeV; $t \approx 6\,$seconds
3) ${}^3He + {}^3He \longrightarrow {}^4He + 2p + 12.85\,$MeV; $t \sim 10^6\,$years



$\epsilon_{PP} \propto X_H^2 \rho T^4$ and $Q_{pp} = [4\Delta M(^1H) - \Delta M(^4He)]c^2 = 26.7\,$MeV.

Note that stages 1) and 2) have to occur twice to make a single ^{4}He. Deuterium is rapidly consumed. The neutrino ν_e has an energy 0.26 MeV (1% of the total energy released). The total energy in photon luminosity produced in fusing 4 protons is $\approx$ 27 MeV. The PP-I Chain accounts for $\approx$ 85% of the energy production in the Sun.

Branches of the PP chain: for step 3. there are two other (less frequent) branches, PP-II & PP-III PP-II Chain ($\approx$ 15% of energy production in Sun)
3′) ${}^3He + {}^4He \longrightarrow {}^7Be + \gamma + 1.6\,$MeV
4′) ${}^7Be + e^- \longrightarrow {}^7Li + \nu_e + 0.9\,$MeV
5′) ${}^7Li + p \longrightarrow {}^4He + {}^4He + 17.4\,$MeV

PP-III Chain ($\sim 0.02\%$ of energy production in Sun)

4″) $^7Be + p \rightarrow\ ^8B + \gamma + 0.1\,\mathrm{MeV}$

5″) $^8B \rightarrow\ ^8Be + e^+ + \nu_e$

6″) $^8Be \rightarrow\ 2\,^4He + 18.1\,\mathrm{MeV}$

For the PP-II and -III chains, ^{4}He acts as a catalyst for the reaction $^3He + p \rightarrow\ ^4He + \nu_e$. These chains are favoured by hotter temperatures and abundant ^{4}He.

The CNO Cycle

In higher mass stars ($> 1.5M_\odot$), the core temperatures are higher (on the main sequence) and hydrogen fusion to form ^{4}He proceeds by another path, catalysed by the elements carbon, nitrogen and oxygen (the CNO cycle).

Observational evidence: some red giants with $^{13}C\,/\,^{12}C \approx 1/5$ (compare to 1/90 on Earth), and some stars with very high nitrogen abundance.

$^{12}C + p \longrightarrow\ ^{13}N + \gamma$

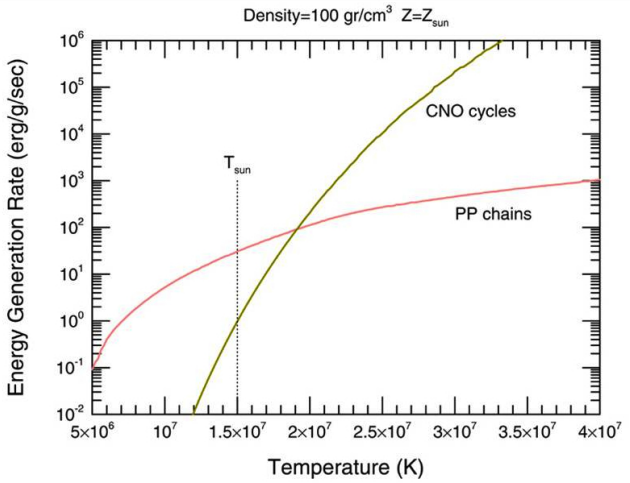
$^{13}N \longrightarrow\ ^{13}C + e^+ + \nu_e$

$^{13}C + p \longrightarrow\ ^{14}N + \gamma$

$^{14}N + p \longrightarrow\ ^{15}O + \gamma$

$^{15}O \longrightarrow\ ^{15}N + e^+ + \nu_e$

$^{15}N + p \longrightarrow\ ^{12}C +^4 He$



$\epsilon_{\mathrm{CNO}} \propto X_{\mathrm{H}} X_{\mathrm{CNO}} \rho T^{17}$.

Slowest reaction in this case: $^{14}N + p \longrightarrow\ ^{15}O + \gamma$ (sets cycle timescale).

^{14}N lives for $\sim 5 \times 10^8\,$yr. Approach to equilibrium set by second slowest reaction ^{12}C+p. In equilibrium:

$$\lambda_{12C} n(^{12}C) = \lambda_{13C} n(^{13}C) = \lambda_{14N} n(^{14}N) = \lambda_{15N} n(^{15}N)$$

where $n()$ is number density and λ is reaction rate. Isotopes which react more slowly have higher abundances: $^{12}C \approx 4\%$, ^{13}C $\approx 1\%$, ^{14}N $\approx 95\%$, ^{15}N $\approx 0.004\%$

Life after the Main Sequence – Helium Burning by the triple-α process

After hydrogen is exhausted in core, contraction and heating up of the core (virial theorem). Simplest reaction in a helium gas should be the fusion of two helium nuclei. BUT there is no stable nucleus configuration with $A = 8$ (beryllium isotope ^{8}Be has a lifetime of only $2.6 \times 10^{-16}\,$s before decaying into two ^{4}He). If a third helium nucleus can be added to ^{8}Be before decay, then the very stable ^{12}C nucleus can be formed. This is the the "triple-alpha" reaction (three Helium-4 nuclei, or "α particles", fuse).

^{4}He + ^{4}He $\longrightarrow\ ^8$Be

^{8}Be + ^{4}He $\longrightarrow\ ^{12}$C+γ; $Q_{3\alpha} = [3\Delta M(^4He) - \Delta M(^{12}C)]c^2 = 7.275\,\mathrm{MeV}$

The reaction has to occur fast enough since two ^{4}He scatter off each other in only $t \sim 2d/v = 2 \times 10^{-15}/2 \times 10^5 \sim 10^{-20}$ seconds.

At $T \sim 10^8$ K and $\rho \sim 10^8\,\mathrm{kg\,m^{-3}}$, the fraction of ^{8}Be is $n(^8Be)/n(^4He) \sim 10^{-9}$.

Fred Hoyle suggested this small probability of α-capture by short lived ^{8}Be would be greatly enhanced if the C nucleus had an energy level close to the combined energies of the reacting ^{8}Be and ^{4}He nuclei (it does!). Reaction is a faster "resonant" reaction. Thus helium burning proceeds in a 2-stage reaction, and energy released per unit mass is 10% that of H-burning (and the duration of the core-burning helium-burning phase is about 10% of the time spent in core-burning hydrogen defining the Main Sequence). Exceptionally strong temperature dependence: $\epsilon_{3\alpha} \propto \rho^2 T^{40}$

Summary of energy generation: the equation of energy generation is $dL/dr = 4\pi r^2 \rho\epsilon$ where $\epsilon = \epsilon\mathrm{pp} + \epsilon_{\mathrm{CNO}}$ is energy generation rate per kg from PP & CNO cycles on main sequence. $\epsilon \propto T^\alpha$ where $\alpha \approx 4$ for p-p chain and $\alpha \approx 17 - 20$ for CNO cycle, and for helium burning $\alpha \approx 40$!

Abundances of nuclei

All nuclei heavier than boron must be made in stars (stellar nucleosythesis) rather than in Big Bang nucleosynthesis. Carbon fusion (C+C) occurs at stellar core temperatures $> 5 \times 10^8$ K, and oxygen burning (O+O) above 10^9 K. Note that low mass stars may never get this hot in their cores, and finish nuclear burning before these elements are created. For the more massive stars, the duration of burning these heavier elements is very short indeed.

$^{12}C + ^{12}C \longrightarrow {}^{24}Mg + \gamma \longrightarrow {}^{23}Mg + n \longrightarrow {}^{23}Na + p \longrightarrow {}^{20}Ne + \alpha \longrightarrow {}^{16}O + 2\alpha \approx 13\,\mathrm{MeV}$ ($\approx 5.2 \times 10^{13}\,\mathrm{J\,kg^{-1}}$).

$^{16}O + ^{16}O \longrightarrow {}^{32}S + \gamma \longrightarrow {}^{31}S + n \longrightarrow {}^{31}P + p \longrightarrow {}^{28}Si + \alpha \longrightarrow {}^{24}Mg + 2\alpha \approx 16\,\mathrm{MeV}$ ($\approx 4.8 \times 10^{13}\,\mathrm{J\,kg^{-1}}$).

Beyond this, complicated by photodistintegration at of oxygen at $T = 1 - 1.5 \times 10^9$ K, and Si at $T > 3 \times 10^9$ K: $^{16}O + \alpha \longleftrightarrow {}^{20}Ne + \gamma$

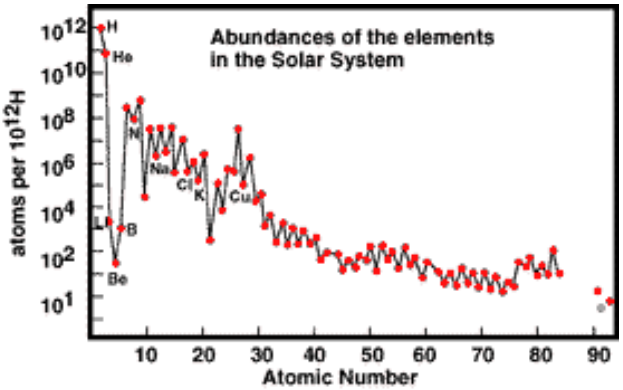
Nuclear reactions reach an equilibrium including two ^{28}Si fusing to ^{56}Fe – end of fusion (endothermic). As iron-56 is the most stable element (highest binding energy per nucleon), no energy released in further fusion. Only the most massive stars will achieve the central temperatures required to reach this stage, and afterwards will collapse rapidly to form a black hole or neutron star. The re-bound of the outer layers will produce a Supernova Type II (an exploding star caused by core collapse).

The s-process and r-process – creation of heavy elements

There are neutrons, protons and α-particles (helium-4 nuclei) produced in nuclear reactions, which can be absorbed by other nuclei in further nuclear reactions. In particular, the interaction between nuclei and free neutrons (neutron capture) is important since neutron capture by heavy nuclei is not limited by the Coulomb barrier (neutrons have no charge) – so could proceed at relatively low temperatures. However, free neutrons are typically scarce, but in late-stage evolution of stars there are enough for chain reactions

to be possible.

$I(A,Z) + n \longrightarrow\ I_1(A+1,Z)$
$I_1(A+1,Z) + n \longrightarrow\ I_2(A+2,Z)$
$I_2(A+2,Z) + n \longrightarrow\ I_3(A+3,Z)$... etc.
If a radioactive isotope is formed it may undergo β-decay if it is unstable, creating a new element.
$I_N(A+N,Z) \longrightarrow\ J(A+N,Z+1) + e^- + \bar{\nu}_e$ (anti-neutrino).



If the new element is stable, it will resume neutron capture, otherwise it may undergo a series of β-decays. Neutron capture reactions may proceed more *slowly* or more *rapidly* than the competing decays. The different chains of reactions and products are hence called the *s-process* and *r-process*.

S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 8: The Life and Death of Stars

Scope of Lecture 8: Stellar evolution – main sequence, turn-off, red giant branch (AGB), helium flash, horizontal branch, asymptotic–GB, planetary nebula (Choudhuri 4.5).

A Brief Life History of Stars

The birth of stars New-born stars are embedded in dense molecular gas clouds. They burn deuterium (at lower temperatures than the PP chain) and accrete from a birth-cloud. The proto-stars contract to join the Main Sequence along the *Hayashi track* (almost vertical in H-R diagram, constant $T_{\rm eff}$), until the centre becomes hot enough for hydrogen to fuse to helium (as the proto-star contracts it heats up through the Virial Theorem). At this point, the nuclear energy source is enough to balance radiation heat loss, and contraction ceases and hydrostatic equilibrium is established.

The main sequence Very little change for a nuclear timescale – i.e. until nuclear fuel exhausted, star spends its time at approximately the same place on the Main Sequence (same luminosity and effective temperature), slight changes due to chemical composition of core.

Post main sequence After hydrogen in core exhausted, a series of phases of alternating contraction (releasing gravitational energy until centre hot enough) and further nuclear reactions (helium to carbon, etc., possibly up to iron), causing outer layers ("the envelope") to expand up into a giant star.

The Death of stars After all possible nuclear fuels exhausted, star contracts to a dead compact object: white dwarf, neutron star or black hole. Outer layers have been "puffed off" before this – mass-loss due to stellar winds or explosion.

Main Sequence Lifetimes (revisited) Following on from lecture 6:

Temperatures inside stars scale with M/R: $T \propto M/R$. So from radiative transfer: $L \propto (T\,R)^2/(\kappa M)$. We showed that for massive stars where ideal gas law holds: $P = \rho\, k_B T/(\mu\, m_H)$. And opacity κ is approximately constant for hot massive stars (mainly due to electron scattering), so:

$$L \propto \frac{(\mu\, R\, T)^4}{\kappa\, M} \propto \frac{\mu^4 M^3}{\kappa} \sim M^3 \Longrightarrow t_{MS} \propto \frac{M}{L} \sim M^{-2}$$

For lower-mass stars , Kramer's opacity law: $\kappa \propto \rho\, T^{-3.5}$ leads to

$$L \propto \frac{\mu^{7.5} M^{5.5}}{R^{0.5}} \sim M^5 \text{ if } M \propto R \Longrightarrow t_{MS} \propto M/L \sim M^{-4}$$

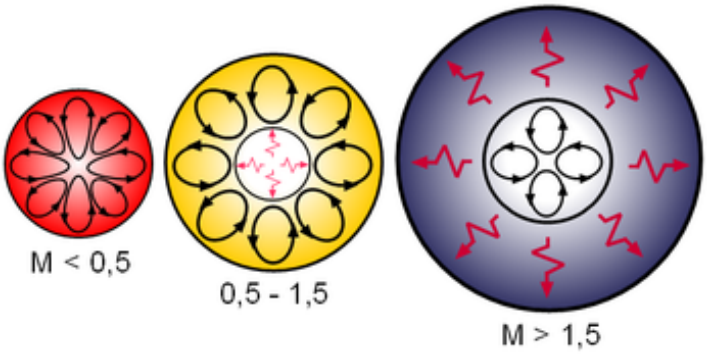
The Sun is in the intermediate range, so $L \propto M^4$, $\tau_{MS} \propto M^{-3}$.

The Most massive stars ($> 100\, M_\odot$) are supported by radiation pressure, and $\kappa \sim$ constant

$$P_{\rm rad} = \frac{a\, T^4}{3} \text{ and } P \propto \frac{GM^2}{R^4} \text{ so } T \propto \frac{M^{0.5}}{R};\; L \propto \frac{(\mu\, R\, T)^4}{\kappa\, M} \Longrightarrow L \propto M$$

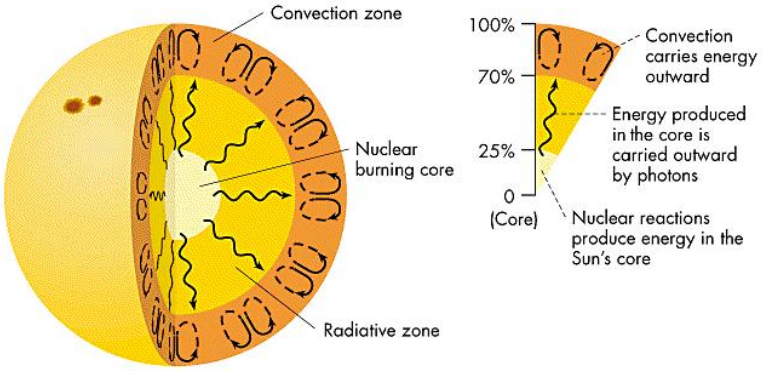
In What Regions is Convection Important in Stars of Different Masses?

Convection important in the cores of high-mass stars because of large temperature gradients (due to strong dependence of CNO cycle and triple-α process on T) mixing cores. Convection is also important in outer envelope of solar-type stars and throughout low-mass stars, partially ionized regions with small $\gamma = C_P/C_V$.



Hydrogen burning at $> 0.08\,M_\odot$. Below this, brown dwarfs supported against collapse by electron degeneracy pressure. Maximum stellar mass $\approx 150\,M_\odot$ – above this stars blow themselves apart by radiation pressure.

Structure of the Sun

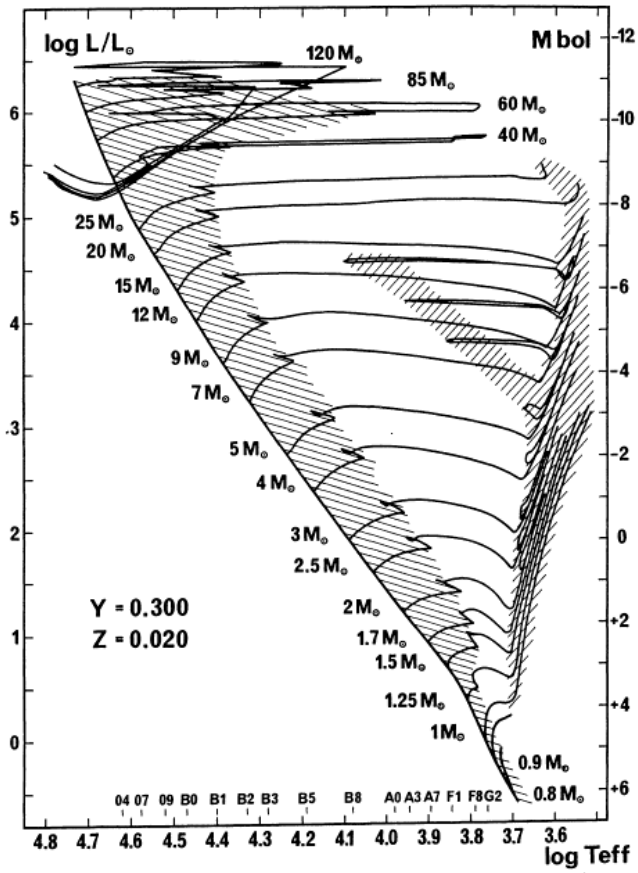


Core : $R < 0.2 R_\odot$ energy generated by nuclear fusion.
Radiative Zone: $0.2\,R_\odot < R < 0.7\,R_\odot$. Energy transport by radiation, X-ray mean free path ~ 1 cm, photons take ≈30,000 years to escape.
Convective Zone : $0.7\,R_\odot < R < 0.99\,R_\odot$. Energy transport by mass motion, seen as solar granulation, material is not fully ionised.
Photosphere : ≈100 km thick, cool surface layer imprints absorption lines on spectrum.

Lower Main Sequence: Hydrogen Exhaustion – Leaving the Main Sequence

The cores of $\approx 1\,M_\odot$ ("lower main sequence", LMS) stars become Helium rich. No convective processes in core, so material not fully mixed. Fusion is most efficient in the centre, where T is highest.

As He content increases, the core shrinks (can no longer support its own weight through thermal pressure as nuclear reaction rate drops) and the core heats up and the He rich core grows.



The T is not high enough for the triple-α process at this evolutionary stage, but extra energy produced in the shrinking core triggers *SHELL BURNING PHASE:* H-burning in a shell around the core, and as T rises, the CNO process can occur in the shell (note: originally p-p process dominated in the core). As $\epsilon_{\rm CNO} \propto T^{17}$ energy generation concentrated in regions of highest T & $X_{\rm HI}$ (hydrogen mass fraction), shell $T \approx 20 \times 10^6$ K. This high temperature causes high pressure outside the core and the H envelope expands.

This expansion becomes more pronounced when $>$ 10% of the stellar mass in the He core. The star leaves MS. The luminosity initially remains approximately constant (a slight brightening), hence $T_{\rm eff}$ must decrease as the envelope expands, and the star moves right along red subgiant branch.

The Red Giant Phase – Red Giant Branch (RGB)

The shell source slowly burns, moving out through the star. The inert He core grows (until $\approx 0.5\,M_\odot$) but when the temperature of the outer layers reach < 5000 K the envelopes become fully convective. This enables greater luminosity to be carried by the outer layers and hence quickly forces the star almost vertically in the H-R diagram (becomes a Red Giant). The star approaches the Hayashi line, and a small increase in the He core mass causes a relatively large expansion of the envelope (can expand to $\sim 100\,R_\odot$). Core mass grows $\rightarrow$ shell temperature increases $\rightarrow$ luminosity increases $\rightarrow$ ascends RGB.

The helium core does not reach threshold temperature for further burning as it ascends the RGB, and as it is not producing energy it continues to contract until it becomes degenerate.

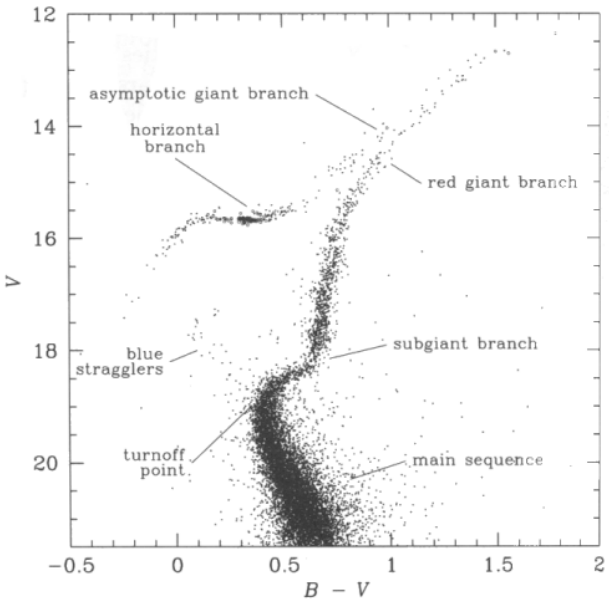
The Helium Flash & Helium Burning in the Core

Stars with $> 2\,M_\odot$ ignite helium in non-degenerate core, but for the lower main sequence ($\approx 1\,M_\odot$) core contraction stops only when electron degeneracy pressure can support the core against gravity when the core mass $= 0.48\,M_\odot$. At the tip of the RGB electrons in core are completely degenerate. Pressure is due to degenerate electron pressure, which is independent of temperature (see Lecture 9). So there is no self-regulation mechanism, and temperature is defined mainly by the energy distribution of the heavy

particles (He nuclei).

For $T \sim 10^8\,\mathrm{K}$, triple-α reactions start in the very dense core. They generate energy, heating the core, and the KE (temperature) of He nuclei increases, increasing the energy production (since $\epsilon_{3\alpha} \propto T^{40}$). Energy generation and heating under degenerate conditions leads to a runway process – *the Helium Flash*. The core temperature changes within seconds, and this rapid increase in T changes the velocity distribution of electrons, and the degeneracy is removed. The pressure increases and core expands, and the star achieves new equilibrium configuration: an expanded non-degenerate core hot enough to burn He. The H-burning shell source has also expanded, and has lower T and density (generates less energy than before).

There is a rapid decrease in luminosity and the star sits in the Red Clump (metal rich stars) or the Horizontal Branch (metal poor stars). At this stage, the star has core helium burning and shell hydrogen burning. RRLyrae stars are at this stage of evolution.
Globular clusters are old and metal poor – we don't see a red clump. We see a horizontal branch: H-burning shells, He burning cores.
Mass-loss drives bluewards evolution.
Lowest mass H-envelope stars are bluest. More metal rich stars appear towards red.
Clump stars — extreme red end of HB.



The Asymptotic Giant Branch (AGB) and thermal pulses

What happens when helium in core is exhausted? The triple-α reaction is very temperature dependent ($\epsilon \propto T^{40}$), hence energy generation is even more centrally condensed. Note the H-burning shell is generating energy. The core will soon consist only of C+O with $R \sim 0.01\,R_\odot$, and in a similar way to before, the inert CO-core grows in mass and contracts while a He-burning shell source develops, with H-burning at larger radii.

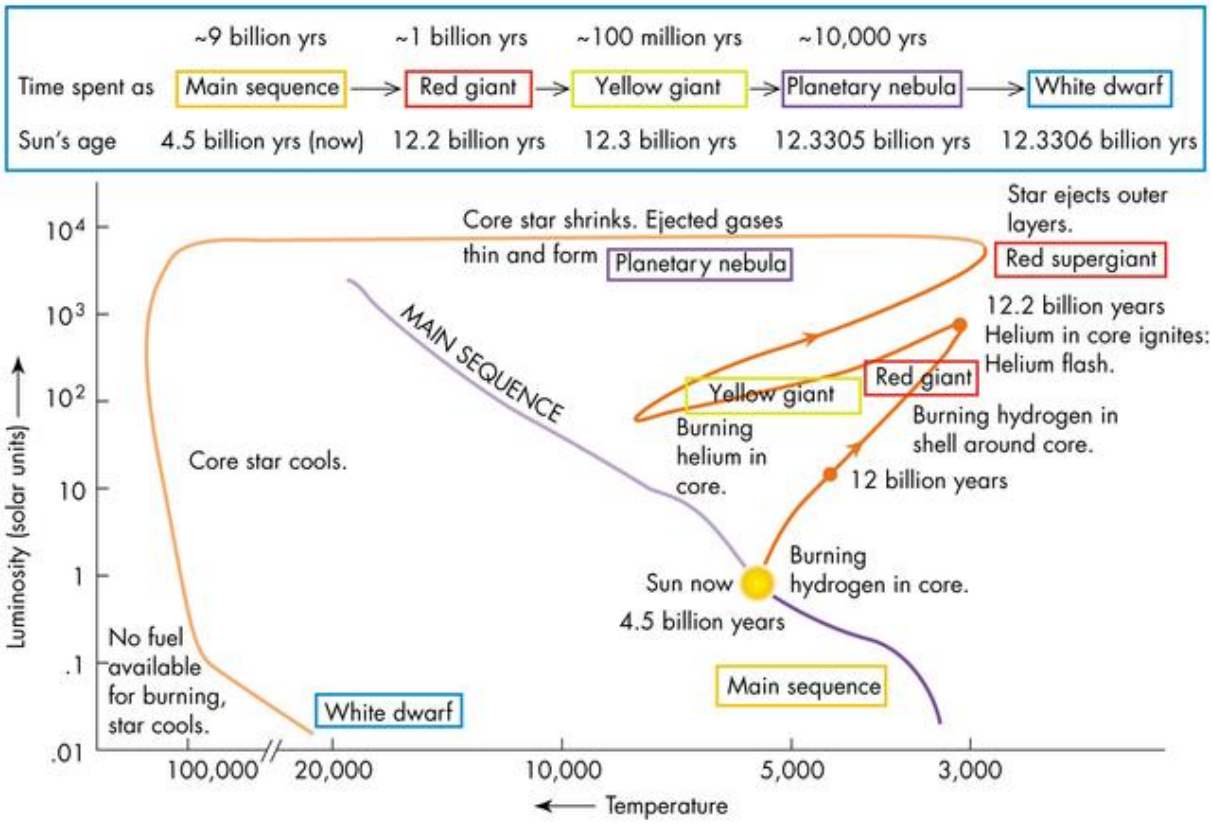
These two shell sources force expansion of the envelop and the star evolves up the red giant branch a second time – this is called the *asymptotic giant branch (AGB)*. In this phase the star can expand to $100 - 1000\,R_\odot$ (similar to Earth's orbit).

Shell-burning phases alternate, creating *Themal Pulsations* and driving off the envelope, which can form a planetary nebula. For high metallicity stars, the AGB coincides closely with the first RGB. For globular cluster stars (low metallicity of 1% solar) they appear separated.

Stellar winds, planetary nebulae and white dwarf stars

Large radiation pressure at tip of AGB probably drives mass-loss. Particles may absorb photons from radiation field and be accelerated out of the gravitational potential well. Mass loss rates are in the range 10^{-9} to $10^{-4}\,M_\odot$/yr (higher loss rates are superwinds – result in stellar ejecta in an observable shell around the central star: this is called a planetary nebula, PN). Envelope attains positive binding energy (perhaps due to energy reservoir in ionization energy) and can be dispersed away to infinity without energy. The hot CO core is exposed after ejection (ionizing the ejected shell – which is seen as planetry nebula, persisting for ∼10,000 years).

CO core cools, becomes degenerate – a white dwarf (density $\sim 10^9\,\mathrm{kg/m^3}$). The core mass at the tip of the AGB is about $0.6\,M_\odot$, and most white dwarfs have about this mass.



Lifetime and H-R evolution of the Sun

Phase	*Timescale*, τ
Main-sequence	9×10^9 years
Subgiant	3×10^9 years
Redgiant Branch	1×10^9 years
Red clump	1×10^7 years
AGB evolution	$\sim 5 \times 10^6$ years
Planetary nebula	$\sim 1 \times 10^5$ years
White dwarf cooling	$> 8 \times 10^9$ years

Evolution of Massive Stars – Main Sequence to Helium Burning

The cores of massive stars are convective, hence newly formed He is evenly mixed. As the hydrogen is consumed, the core contracts and also shrinks in mass. The convective core exhausted homogeneously, contracts and becomes hotter. There is an H-burning shell around core.

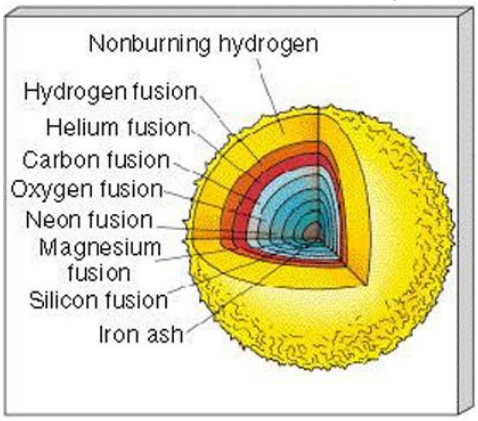
The temperature at the bottom of the hydrogen envelope is too high to sustain hydrostatic equilibrium. The envelope expands and the star becomes cooler, moving to the red region of the HR diagram. It becomes a red supergiant star. Due to the rapid drop in temperature throughout the outer atmosphere, the criterion for convection is reached in this region and a very deep convection zone develops.

It dredges up material from original convective core, which can appear at the surface.

Very Massive Stars – from Helium Burning to Core Collapse

The He burning core (∼10,000 km across) is surrounded by a H-burning shell. The triple-α process liberates less energy per unit mass than for H-burning (∼10%). Hence the lifetime of this phase is shorter (around 10% MS lifeime). There is *no* He-flash as densities in He-core not high enough for electron degeneracy. Eventually have core of ^{12}C and ^{16}O, surrounded by He and H burning shells. The core again contracts and temperature rises, allowing C & O burning to Mg & Si. This process continues (depending on mass of star), with increasing atomic number Z, building up heavier and heavier elements until the iron group elements of Ni, Fe and Co are formed. The core is surrounded by a series of shells at lower T, and lower ρ. When the core reaches iron, no more energy can be liberated and it will collapse, resulting in a Type II Supernova explosion. For a $M \approx 15\,M_\odot$ star:

Central burning Phase	*Timescale*
Hydrogen	10×10^6 years
Helium	1×10^6 years
Carbon	400 years
Nitrogen & Oxygen	1 year
Silicon	10^{-2} years ($\sim$ 1 day!)



S26 (Short Option) – Stars & Galaxies – Prof Andrew Bunker

Lecture 9: Degenerate Dead Stars — White Dwarfs & Neutron Stars

Scope of Lecture 9: White dwarfs – degenerate gas,relativistic and non-relativistic (Choudhuri 5.2,5.3). Stellar evolution of high mass stars: shell model, burning higher mass nuclei until iron core, Type II supernovae, Neutron star, Chandrasekhar mass.
Additional reference: Carroll & Ostlie chapter 16 "The Degenerate Remnants of Stars".

Equation of state of a degenerate gas

Where the density of matter is high, the Pauli exclusion principle for fermions becomes important (that is, no state of position and momentum r, p can be occupied by more than one identical particle). Electrons are fermions and have spin 1/2 so you can have two electrons per state, spin-up and spin-down. We will show that when all available states up to the Fermi energy are occupied, the gas experiences pressure due to this degeneracy. Unlike the ideal gas law ($P = N\,k_B T$), this degeneracy pressure is independent of temperature. Implications for a degenerate gas:
— Temperature regulation of nuclear burning processes no longer hold (can get thermonuclear runaway – such as the Helium flash).
— Even if there is no energy generation in the star, can be supported against collapse by degeneracy pressure (hydrostatic equilibrium maintained not by thermal gas pressure but by Pauli exclusion). There is pressure even if $T = 0$.

No more than two electrons (of opposite spin) can occupy the same quantum cell. The quantum cell of an electron is defined in phase space of position and momentum, (x, y, z, p_x, p_y, p_z). The volume of allowed phase space is given by $\Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z = h^3$. In a dense stellar core (e.g. a white dwarf), volume of phase space almost full for electrons — it is not possible for more than two electrons to occupy the same position unless their momenta are very different.

If electrons occupy a real space volume V and have momenta in the range d^3p (=a shell in 3D momentum space between p and $p+\delta p$) then the states available are $2V\,d^3p/h^3$. The volume of momentum space occupied by electrons with momenta in the range $p \rightarrow p + \delta p$ is given by the volume of a spherical shell of radius p, thickness δp: $4\pi p^2 \delta p$. The volume of phase space occupied is the product of the volume in real space and the volume in momentum space: $V_{\rm ph} = 4\pi p^2 V\,\delta p$. Number of quantum states in this volume is $V_{\rm ph}$ divided by volume of a quantum state (h^3), $4\pi p^2 V\,\delta p/h^3$ which gives the maximum number of electrons with momenta in this interval, $N_e \delta p$, with 2 spins:

$$N\,\delta p \leq 8\pi p^2 V\,\delta p$$

Define a completely degenerate gas: one in which all of momentum states up to the Fermi momentum p_F are filled, while the states with momenta greater than p_F are empty:

$$p \leq p_F :\ N = \frac{8\pi p^2\,V}{h^3};\quad p > p_F :\ N_e = 0$$

The pressure P is mean rate of transport of momentum across unit area

$$P = \frac{1}{3}\int_0^\infty \frac{N_e}{V} p\, v\, dp$$

where v = velocity of electrons with momentum p. We can use the relation between p and v from theory of special relativity

$$p = \frac{m_e v}{(1-\frac{v^2}{c^2})^{\frac{1}{2}}} \longrightarrow\ v = \frac{\frac{p}{m_e}}{(1+\frac{p^2}{m_e^2 c^2})^{\frac{1}{2}}}$$

where m_e = rest mass of electron. Combining the three expressions for N, P, and v, we obtain pressure of a completely degenerate gas

$$P = \frac{8\pi}{3h^3 m_e}\int_0^{p_F} \frac{p^4}{(1+\frac{p^2}{m_e^2 c^2})^{\frac{1}{2}}} dp$$

For a non-relativistic degenerate gas ($p_F \ll m_e\, c$)

$$P = \frac{8\pi}{3h^3 m_e}\int_0^{p_F} p^4\, dp = \frac{8\pi p_F^5}{15h^3 m_e} \longrightarrow\ P = \frac{1}{20}\left(\frac{3}{\pi}\right)^{\frac{2}{3}} \frac{h^2\, n_e^{\frac{5}{3}}}{m_e}$$

The electron degeneracy pressure for a non-relativistic degenerate gas:

$$p_F = \left(\frac{3N_e h^3}{8\pi V}\right)^{\frac{1}{3}} = \left(\frac{3n_e h^3}{8\pi}\right)^{\frac{1}{3}}$$

Where $n_e = N_e/V$ (number density of electrons). Relativistic degenerate gas ($p_F \gg m_e c$; when v approaches c and momentum $\to \infty$):

$$P = \frac{8\pi}{3h^3 m_e}\int_0^{p_F} \frac{p^4}{(1+\frac{p^2}{m_e^2 c^2})^{\frac{1}{2}}} dp \approx \frac{8\pi}{3h^3 m_e}\int_0^{p_F} p^3 c\, m_e\, dp = \frac{2\pi p_F^4 c}{3h^3} = \frac{1}{8}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} h\, c\, n_e^{\frac{4}{3}}$$

The aim is to obtain equation of state for a degenerate gas. We must convert n_e to mass density (using similar arguments to derivation of mean molecular weight). For each mass of H (m_H) there is one electron. For He and heavier elements there is approximately $\frac{1}{2}e^-$ for each m_H. Thus:

$$n_e = \frac{\rho X}{m_H} + \frac{\rho(1-X)}{2m_H} = \frac{\rho(1+X)}{2m_H}$$

Equation of state for a degenerate non-relativistic white dwarf: $P_{\rm gas} = K_1\, \rho^{5/3}$.
Equation of state for a degenerate relativistic white dwarf: $P_{\rm gas} = K_2\, \rho^{4/3}$.

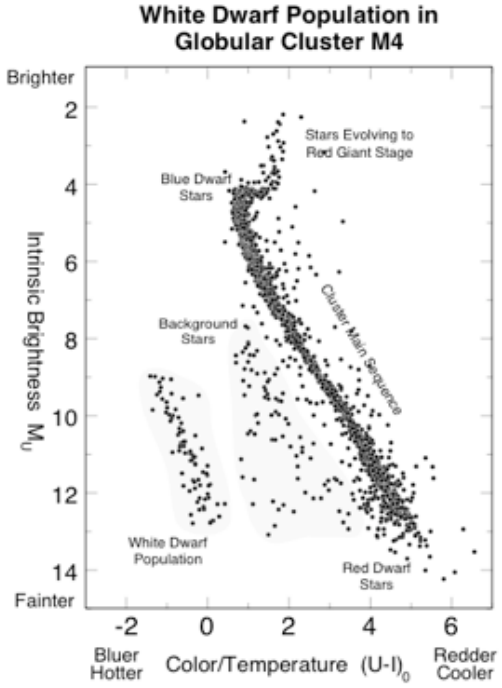
$$K_1 = \frac{h^2}{20m_e}\left(\frac{3}{\pi}\right)^{\frac{2}{3}}\left(\frac{1+X}{2m_H}\right)^{\frac{5}{3}} \qquad K_2 = \frac{h\, c}{8}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\left(\frac{1+X}{2m_H}\right)^{\frac{4}{3}}$$

In a completely degenerate gas the pressure depends only on the density and chemical composition. It is independent of temperature.

Degenerate stars

There is not a sharp transition between relativistically degenerate and non-relativistically degenerate gas. Similarly there is no sharp transition between an ideal gas and a completely degenerate one. *Partial degeneracy* situation requires much more complex solution.

White dwarfs are intrinsically faint, hot stars. Typical observed masses $0.1 - 1.4\,M_\odot$.



Left Figure: Example of WD discovered in Globular cluster M4 (cluster age $\sim$ 13 Myrs). WDs represent cooling sequence some have similar intrinsic brightness as main-sequence members, but much hotter (hence bluer).

White Dwarf Cooling

Conduction by electrons is very effective in white dwarfs, and is the dominant mode of energy transport — degenerate electrons can cover long distances before colliding with a nucleus (since the lower-energy states are fully occupied). It is so efficient that the interior of a white dwarf is nearly isothermal (temperature T_C), with the temperature declining only in the thin outer surface (where the degeneracy is lifted), with surface temperature

$T_{\rm eff}$. If this outer envelope is radiative:

$$L = 4\pi R^2 \sigma\, T_{\rm eff}^4 \propto T_C^{\frac{7}{2}} \qquad L = 6.65\times 10^{-3}\left(\frac{M}{M_\odot}\right)\frac{\mu}{Z(1+X)}T_C^{\frac{7}{2}}$$

(see Carroll & Ostelie Section 16.5). If the White Dwarf has uniform composition, with nucleon number A (e.g. $A = 12$ for Carbon), the number of nuclei:

$$N = \frac{M_S}{A\, m_H} \text{ and each has average thermal energy } KE = \frac{3k_B T}{2}$$

$$\text{so total thermal energy available: } U = \frac{M_S 3k_B T_C}{2A\, m_H}$$

and the white dwarf shines just by cooling (no nuclear processes):

$$-\frac{dU}{dt} = L_S \Longrightarrow -\frac{d}{dt}\left(\frac{M_S 3k_B T_C}{2A\, m_H}\right) = C\, T_C^{\frac{7}{2}}$$

$$\text{Cooling timescale } \tau_{\rm cool} \approx \frac{U}{L_S} = \frac{1}{CT^{\frac{5}{2}}}\frac{M_S}{A\, m_H}\frac{3}{2}k_B \ (\sim 170\,\text{million years})$$

where $C = 6.65\times 10^{-3}\left(\frac{M}{M_\odot}\right)\frac{\mu}{Z(1+X)}$ so $T_C(t) = T_0\left(1+\frac{5t}{2\tau_{\rm cool}}\right)^{-\frac{2}{5}}$, T_0 is initial temp

$$L_S(t) = L_0\left(1+\frac{5t}{2\tau_{\rm cool}}\right)^{-\frac{7}{5}} = 4\pi R_S^2\,\sigma\, T_{\rm eff}^4$$

Using equation for non-relativistic degeneracy, and hydrostatic equilibrium with $dP/dM \sim -P_C/M_S$

$$\frac{dP}{dM} = -\frac{G\,M}{4\pi r^4} \Longrightarrow P_C \approx \frac{G\,M_S^2}{R_S^4} = \left(0.77\frac{G\,M_S^2}{R_S^4}\right) \qquad P \sim P_e = \frac{h^2}{20m_e}\left(\frac{3}{\pi}\right)^{\frac{2}{3}}\left(\frac{1+X}{2m_H}\right)^{\frac{5}{3}}\rho^{\frac{5}{3}}$$

Done properly, for a self-gravitating sphere $\rho_C = 6.0\bar{\rho}$ and $\bar{\rho} = M_S/(\frac{4}{3}\pi R_S^3)$

$$\text{For } X = 0 \text{ (no hydrogen) } R = \frac{0.114\,h^2}{G\,m_e(2m_H)^{5/3}}\frac{1}{M^{1/3}} \text{ so } R_S \propto M^{-1/3}$$

So for $M = 1\,M_\odot$, $R_S = 8800\,\text{km}$ (similar to size of Earth), $\rho_C = 4.2\times 10^9\,\text{kg/m}^3$.
Bizarre property — more massive white dwarf stars are smaller! (non-relativistic case).

$$\text{In the degenerate relativistic case, } P \propto \rho^{\frac{4}{3}} \qquad P \sim P_e \propto \left(\frac{\rho}{\mu_e\, m_H}\right)^{\frac{4}{3}} \sim \frac{G\,M_S^2}{R_S^4}$$

Which implies that radius is independent of mass! (total mass independent of density). Suggests there is a fixed mass for a star obeying the fully-relativistic equation of state. It is not possible to have more massive white dwarfs.

The Chandrasekhar mass – maximum mass supported against collapse by degeneracy pressure

Consider a star with radius R_S containing N fermions (electrons for a white dwarf, or neutrons for a neutron star).
Number density $n = 3N/(4\pi R_S^3)$ and from the uncertainty principle, $[\Delta x \Delta p]^3 \sim (h/2\pi)^3$.

So typical momentum: $p_F \sim n^{1/3}h/(2\pi)$ and energy $E = pc$ if relativistic: $E_F \sim n^{1/3}hc/(2\pi) \sim N^{1/3}h\,c/(2\pi R_S)$.

Now consider the gravitational energy of the star $E_G \approx G\,M_S^2/R_S$.
And the mass of the star $M_S = N\,\mu_f\,m_H$ where μ_f is the mean mass per fermion (in atomic mass units).

Total energy per fermion: $E_{\rm TOT} = E_F + E_G \sim N^{1/3}h\,c/(2\pi R_S) - G\,N(\mu_f\,m_H)^2/R_S$.

Maximum stable configuration if $E_{\rm TOT} = 0$ (for $E_{\rm TOT} < 0$, gravitational collapse).

$$\text{Maximum number of fermions: } N_{\rm max} \sim \left(\frac{h\,c}{2\pi G(\mu_f\,m_H)^2}\right)^{\frac{3}{2}} \approx 2 \times 10^{57}$$

Chandrasehkar mass for White Dwarfs (electron degeneracy, $\mu_f \approx 2$ if $X = 0$, no hydrogen).

If done properly, $M_{\rm Ch} \approx 1.46\left(\frac{2}{\mu_f}\right)^2 M_\odot$. None are observed heavier than this (most $\approx 0.6\,M_\odot$). For neutron stars (see below), mean mass per fermion (where the fermions are neutrons rather than the electrons in White Dwarfs) is instead $\mu_f = 1$.

Neutron Stars

Neutrons are also spin-1/2 fermions, obeying Pauli Exclusion. Usually free neutrons decay in ≈ 13 minutes half-life: $n \to p + e^- + \bar{\nu}_e$. The reverse reaction is not energetically favoured due to the mass deficit $p + e^- \to n + \nu_e$. However, in extremely high density environments, due to Pauli exclusion energy levels full and unavailable, so forming a neutron is energetically favourable for an electron rather than having to acquire a Fermi energy. This is called "Neutron Drip", and forms a neutron star if massive enough – supported by neutron degeneracy pressure. Essentially a single atom with an atomic number of $\sim 10^{57}$!

$r = 1.2 \times 10^{-15} A^{1/3} = 1.2 \times 10^{-15} \times (2 \times 10^{57})^{1/3}\,\text{m} \approx 10\,\text{km}$.

Neutrons effectively "touching" (separation the range of the strong nuclear force). Density is about $1.2 \times 10^{10}\,\text{kg/m}^3$. Escape velocity relativistic. Many rapidly spinning. Rotating magnetic dipole $\sim 10^9$ T – causes pulsars. Theoretical limit about $2\,M_\odot$ – above this neutron degeneracy can no longer support collapse, and it becomes a black hole.

S26 (Short Option) – Stars & Galaxies – Dr Andrew Bunker

Lecture 10: Galaxy Dynamics

Scope of Lecture 10: *Galaxy dynamics - galaxies treated as systems of stars in spherically symmetric gravitational potentials. The Collisionless Boltzmann Equation; Jeans' equations; moments of distributions.*

Gravity is a long-range attractive force - all stars attract all others (unlike in electromagnetism, no screening by negative "charges"). **Gravitational potential and force per unit mass** (acceleration) is:

$$\Phi(r) = -\frac{GM}{r} \qquad \text{acceleration} \quad \boldsymbol{a} = -\nabla\Phi(r) \qquad \boldsymbol{a} = -\frac{GM}{r^2}\hat{r}$$

Energy of an Orbit

Angular momentum $\boldsymbol{L} = \boldsymbol{r} \times (m\dot{\boldsymbol{r}})$; $\dot{\boldsymbol{L}} = \boldsymbol{r} \times (m\ddot{\boldsymbol{r}}) = \boldsymbol{r} \times \boldsymbol{F} \equiv \boldsymbol{G}$ which is the Torque

Kinetic energy $KE = \frac{1}{2}m\dot{\boldsymbol{r}}^2 = \frac{1}{2}m(\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}) = T$; $\dot{T} = m\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} = \dot{\boldsymbol{r}} \cdot \boldsymbol{F} = -m\dot{\boldsymbol{r}} \cdot \nabla\Phi$

Potential energy $PE \equiv m\Phi = W$

And if Φ is independent of time: $\dot{W} = m\left(\frac{d\Phi}{dx}\frac{dx}{dt} + \frac{d\Phi}{dy}\frac{dy}{dt} + \frac{d\Phi}{dz}\frac{dz}{dt}\right) = m\nabla\Phi \cdot \dot{\boldsymbol{r}}$

$\dot{E} = \dot{T} + \dot{W} = -m\dot{\boldsymbol{r}} \cdot \nabla\Phi + m\nabla\Phi \cdot \dot{\boldsymbol{r}} = 0 \therefore$ energy is conserved along the orbit

Orbits in Spherical Potentials

If $\Phi(\boldsymbol{r})$ is only dependent on $|\boldsymbol{r}|$, not $\boldsymbol{r}$, then

$$F = -m\nabla\Phi = -m\frac{d\Phi}{dr}\hat{\boldsymbol{r}}$$

then

$$\dot{\boldsymbol{L}} = \boldsymbol{r} \times \boldsymbol{F} = -m\frac{d\Phi}{dr}\boldsymbol{r} \times \hat{\boldsymbol{r}} = 0$$

⇒ angular momentum about origin is conserved

Hence, orbital motion is in a plane

$\boldsymbol{r}, \dot{\boldsymbol{r}}$ in the plane

specific angular momentum, $\boldsymbol{J} = \boldsymbol{L}$/m perpendicular to the plane

Many-Body Systems There is no general analytic solution for systems with N>2 point masses. We must turn to n*umerical simulations:* we calculate force on each body by summation over all other bodies, then move masses over a small time-step, then repeat this process to determine motions.

Another approach is a statistical treatment of dynamics, which is appropriate for large N systems (many particles). We can determine form of gravitational potential, and hence determine the orbits of individual objects in this fixed potential. From this we can calculate the evolution of system as a whole (e.g. collapse/expansion) as well as the probabilities of individual events (e.g. collisions of stars, sinking of massive objects to centre).

Consider system of N point masses m_i at positions R_i and relative radii $\boldsymbol{r}_{ij} = \boldsymbol{R}_i - \boldsymbol{R}_j$ $\boldsymbol{r}_{ij} = -\boldsymbol{r}_{ji}$

The equation of motion of the i^{th} particle is $m_i\ddot{\boldsymbol{R}}_i = G\sum_{j=1}^{N}\frac{m_i m_j}{r_{ij}^3}\boldsymbol{r}_{ij}$ $\quad (j \neq i,\ i = 1, N)$

As long as the system is isolated:

- The centre of mass of the system has a constant velocity
- Total angular momentum is conserved
- Total energy is conserved
- The angular momentum vector defines the vector of the "invariable plane"

The Virial Theorem (revisited)

We can derive (again) the virial theorem by considering the moment of inertia of the system:

$$\boldsymbol{I}=\sum_{i=1}^{N} m_i \boldsymbol{R_i^2} \qquad \dot{\boldsymbol{I}}=2\sum_{i=1}^{N} m_i \boldsymbol{R_i}\cdot\dot{\boldsymbol{R}}_i \qquad \left(\text{differentiating w.r.t time}\right)$$

$$\ddot{\boldsymbol{I}}=2\sum_{i=1}^{N} m_i \boldsymbol{R_i}\cdot\ddot{\boldsymbol{R}}_i+2\sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i^2 \quad \left(\text{differentiating w.r.t time again}\right)$$

Compare this to the kinetic and potential energies:

the kinetic energy is $T=\frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i^2$; potential energy is $W=-\frac{1}{2}G\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{m_i m_j}{r_{ij}} \quad i\neq j$.

Using the equation of motion of the ith body $m_i \ddot{\boldsymbol{R}}_i=-\nabla_i W$ where the subscript refers to the gradient experienced by the ith body. The second derivative of *I* is then $\ddot{\boldsymbol{I}}=-2\sum_{i=1}^{N} \boldsymbol{R_i}\cdot\nabla_i W+4\mathrm{T}$

but it can be shown that $\sum_{i=1}^{N} \boldsymbol{R_i}\cdot\nabla_i W=-W$ and thus $\ddot{I}=2\mathrm{W}+4\mathrm{T}$ which can also be written

$\ddot{I}=2\mathrm{W}+4\mathrm{T}=2\mathrm{E}+2T$ where E=T+W is the total energy of the system.

If T>-E, then $\ddot{I}>0$ and the moment of inertia will increase, i.e. the system will expand.

If T<-E, then $\ddot{I}<0$ and the moment of inertia will decrease i.e. the system will contract.

For a system to be globally stable requires $\ddot{I}=0$. Thus, the Virial Theorem is that for a stable system $2\mathrm{T}+W=0$. Therefore, for a system in virial equilibrium: T= –W / 2 = – E.

In forming a virialised system, only half the available potential energy may be converted into *K.E.*

Radius and Velocity Dispersion

For a spherical system that lacks a sharp edge, we can define a 'gravitational radius', defined in terms of the total mass, M, of the system and its potential energy. $r_g=\frac{GM^2}{|W|}$

The kinetic energy of this system is $T=\frac{1}{2}M\langle v^2\rangle$ where $\langle v^2\rangle$ is the mean of the squares of the velocities of the objects "stars" otherwise known as the square of the "velocity dispersion"

The virial theorem then gives $\langle v^2\rangle=\frac{|W|}{M}=\frac{GM}{r_g}=3\langle v_r^2\rangle$

We generally only observe the radial (line-of-sight) velocity, v_r, from the Doppler shift, but assume from symmetry: $\langle v_r^2\rangle\approx\langle v_\theta^2\rangle\approx\langle v_\phi^2\rangle$ and $\langle v^2\rangle=\langle v_r^2\rangle+\langle v_\theta^2\rangle+\langle v_\phi^2\rangle$

Hence we can measure the mass of a star cluster from the velocity dispersion of its stars (assuming it's in virial equilibrium).

Collapse of a Cold System

Consider a system of stars initially at rest (dynamically `cold': zero kinetic energy, T=0).

The system is out of virial equilibrium, and will collapse. Initial Total energy E=W (E=T+W and T=0).

Radius of system is initially $r_{g,0}=-\frac{GM^2}{W}=-\frac{GM^2}{E}$

System collapses (gravitational attraction between stars). Total energy E is conserved: E=T+W.

The system reaches virial equilibrium when 2T=–W (i.e. E=T+W=W/2). Hence final W=2E and

$\frac{GM^2}{r_{g,\text{final}}}=2\frac{GM^2}{r_{g,0}} \qquad \Rightarrow r_{g,\text{final}}=\frac{r_{g,0}}{2}$ Thus, the system collapses to half its initial size.

Collisionless Systems: Galactic Dynamics

Observed many-body systems include stellar clusters (with 100 to 10^6 stars) and galaxies (10^{10}-10^{12} stars), and also clusters of galaxies containing a few thousand galaxies. These are generally in virial equilibrium since: age>>dynamical time (time for star to cross the cluster or orbit the galaxy).

The overall structure is not changing (only over many dynamical times), and we can treat systems as collisionless. This does not refer to actual collisions between stars (which are extremely rare), but instead refers to systems where the dynamics are determined by overall gravitational potential of the ensemble of stars.

In calculating orbits of stars in the potential, we Ignore individual gravitational interactions between stars. Close encounters will occur occasionally, and these will be important for the individual stars, but not the system as a whole (in this "collisionless approximation").

Gravitational Potentials

Poisson's equation relates density distribution ρ to potential Φ $\quad \nabla^2\Phi(\boldsymbol{r}) = 4\pi G\rho(\boldsymbol{r})$.

Newtonian gravity is linear: the total force is the vector sum of forces from all mass elements. We generally consider the potential - easier to visualise the scalar potential than the vector force field. Integrating Poisson's equation, we have Gauss' theorem:

$$\iiint_V \nabla^2\Phi(\boldsymbol{r})\,dV = \iiint_V 4\pi\mathrm{G}\rho(\boldsymbol{r})\,dV = 4\pi GM$$

$$\iiint_V \nabla^2\Phi(\boldsymbol{r})\,dV = \iint_S \nabla\Phi(\boldsymbol{r})\cdot d\boldsymbol{S} \qquad \therefore \iint_S \nabla\Phi(\boldsymbol{r})\cdot d\boldsymbol{S} = 4\pi GM$$

The integral over any closed surface of the normal component of the gradient of the potential is equal to $4\pi G$ times the mass contained within the surface.

Newton's Theorems

For spherical systems, Newton gave us two useful theorems for calculating the gravitational force:

1) A body that is inside a spherical shell of matter experiences no net gravitational force from the shell

2) A gravitational force on a body that lies outside a spherical shell of matter is the same as it would be if the shell's mass was concentrated at one point at the shell's centre

These rules allow us to calculate the potential of an arbitrary spherical mass by summing over shells,

$$\text{The potential inside a uniform sphere of radius } r_0 \text{ is } \quad \Phi_{\text{in}} = \frac{2}{3}\pi G\rho_0(r^2 - 3\mathrm{r}_0^2)$$

for example:

$$\text{The potential outside is } \Phi_{\text{out}} = -\frac{GM}{r}$$

Escape Speed and Circular Speed From Newton's 2nd theorem:

$$\boldsymbol{f} = -\frac{d\Phi}{dr}\hat{\boldsymbol{r}} = -\frac{GM}{r^2}\hat{\boldsymbol{r}} \quad \text{where} \quad M = \int_0^r \rho(r')r'^2\,dr' \quad \text{is the enclosed mass}$$

The circular speed is that for which the centrifugal force balances gravity

$$f = \frac{GM}{r^2} = \frac{v_c^2}{r} \qquad \Rightarrow \qquad v_c^2 = \frac{GM}{r}$$

The escape speed is that required for an object to escape the potential

$$\frac{1}{2}v_{\text{esc}}^2 = |\Phi(r)| = |-\frac{GM}{r}| \qquad \Rightarrow \qquad v_{\text{esc}}^2 = \frac{2GM}{r}$$

For a homogeneous sphere

$$M = \frac{4\pi}{3}\rho_0 r^3 \quad \rightarrow \quad v_c = \sqrt{\frac{4\pi\mathrm{G}\rho_0}{3}}\,r \quad \rightarrow \quad v_c \propto r \quad \text{circular speed increases with radius}$$

For a centrally-condensed sphere (e.g. Dark Matter halo of galaxy)

$$\text{If } \rho \propto 1/r^2 \qquad \rightarrow \qquad M = \int_0^r \frac{4\pi}{3}\frac{\rho_0}{r'^2}r'^2\,dr' \quad \propto r \quad \rightarrow \quad v_c \propto \sqrt{\frac{r}{r}} = \text{constant}$$

Orbits in Spherical Potentials

General orbits in spherical potentials are not closed, but form rosettes. Eventually the orbit passes through every point in the annulus between some inner and outer radii.

The two exceptions where orbits are closed are: (i) the spherical harmonic oscillator potential,

$\Phi=\Omega^2 r^2/2$ + constant (generated by a homogenous sphere); and (ii) the Keplerian potential

Φ=-GM/r (generated outside a point mass or any spherical mass).

The Collisionless Boltzmann Equation (CBE)

We can treat stars as we would atoms in a gas (i.e. determine density of stars in a region and their distribution of velocities, rather than following motions of individual stars).

Position $\underline{r}=x\underline{i}+y\underline{j}+z\underline{k}$ and velocity $\underline{v}=v_x\,\underline{i}+v_y\,\underline{j}+v_z\,\underline{k}$ both have 3 spatial coordinates, and distribution function *f(r,v,t)* gives probability density in this 6-dimensional phase space:

$f(\underline{r},\underline{v},t)\Delta x\Delta y\Delta z\Delta v_x\Delta v_y\Delta v_z$= number of particles in volume $\Delta x\Delta y\Delta z$ between $\underline{v}$ and $\underline{v}+\underline{\Delta v}$

If the number of stars is constant (no stellar births, deaths or stars leaving) then there should be an *equation of continuity*: the particles obey Hamiltonian dynamics, and the distribution function does not change with time (this is *Liouville's Theorem*).

$$\frac{df}{dt}=0=\lim_{dt\to 0}\frac{f\left(\underline{r}+\underline{\dot r}\,dt\,,\underline{v}+\underline{\dot v}\,dt\right)-f\left(\underline{r}\,,\underline{v}\,,t\right)}{dt}$$ Expand as a Taylor series:

$$f\left(\underline{r}+\underline{\dot r}\,\delta t\,,\underline{v}+\underline{\dot v}\,\delta t\right)=f\left(\underline{r}\,,\underline{v}\,,t\right)+\sum_i \delta t\;\dot r_i\frac{\partial f}{\partial r_i}+\sum_i \delta t\;\dot v_i\frac{\partial f}{\partial v_i}+\delta t\frac{\partial f}{\partial t}$$

Where summation is over 3 dimensions x,y,z $\sum_i \dot r_i\frac{\partial f}{\partial r_i}+\sum_i \dot v_i\frac{\partial f}{\partial v_i}+\frac{\partial f}{\partial t}=\frac{df}{dt}=0$

And using the gravitational scalar potential, Φ

$$\frac{F_i}{m}=\ddot r_i=\dot v_i=-\nabla\Phi \qquad\qquad \underline{v}\cdot\nabla f-\nabla\Phi\cdot\frac{\partial f}{\partial \underline{v}}+\frac{\partial f}{\partial t}=0$$

The CBE in Cylindrical Coordinates Rather than Cartesian coordinates, f($x,y,z,v_x\,v_y\,,v_z$) it is more useful to have the Collisionless Boltzmann Equation in polar coordinates - specifically, cylindrical coordinates for disk galaxies, and spherical coordinates for ellipticals.

$$v_r=\dot r\;,\;v_\theta=r\dot\theta\;\text{ and }\;v_z=\dot z \qquad\qquad KE=\tfrac{1}{2}m|\underline{v}|^2=\tfrac{1}{2}m\left(\dot r^2+r^2\dot\theta^2+\dot z^2\right)$$

Here we consider cylindrical coordinates $f(r,\theta,z,v_r\;,v_\theta\;,v_z)$ where

$$v_r\frac{\partial f}{\partial r}+\frac{v_\theta}{r}\frac{\partial f}{\partial\theta}+v_z\frac{\partial f}{\partial z}+\dot v_r\frac{\partial f}{\partial v_r}+\dot v_\theta\frac{\partial f}{\partial v_\theta}+\dot v_z\frac{\partial f}{\partial v_z}+\frac{\partial f}{\partial t}=\frac{df}{dt}=0$$

The Lagrangian of a star of mass=1 in a potential $\Phi(r,\theta,z)$:

$$L=\tfrac{1}{2}\left(\dot r^2+r^2\dot\theta^2+\dot z^2\right)-\Phi\left(r\,,\theta\,,z\right)\;\text{ and }\;\text{accel}=\frac{F}{m}=-\nabla\Phi \qquad\qquad \ddot z=-\frac{\partial\Phi}{\partial z}\Rightarrow\dot v_z=-\frac{\partial\Phi}{\partial z}$$

$$\ddot r-r\dot\theta^2=-\frac{\partial\Phi}{\partial r}\Rightarrow\dot v_r=\frac{v_\theta^2}{r}-\frac{\partial\Phi}{\partial r} \qquad\qquad \frac{d}{dt}\left(r^2\dot\theta\right)=\frac{d}{dt}\left(rv_\theta\right)=-\frac{\partial\Phi}{\partial\theta}\Rightarrow v_r v_\theta+r\dot v_\theta=-\frac{\partial\Phi}{\partial\theta}$$

$$v_r\frac{\partial f}{\partial r}+\frac{v_\theta}{r}\frac{\partial f}{\partial\theta}+v_z\frac{\partial f}{\partial z}+\left(\frac{v_\theta^2}{r}-\frac{\partial\Phi}{\partial r}\right)\frac{\partial f}{\partial v_r}-\left(\frac{v_r v_\theta}{r}+\frac{1}{r}\frac{\partial\Phi}{\partial\theta}\right)\frac{\partial f}{\partial v_\theta}-\frac{\partial\Phi}{\partial z}\frac{\partial f}{\partial v_z}+\frac{\partial f}{\partial t}=0$$

Jeans Equations

Unfortunately the Collisionless Boltzmann Equation is a function of seven variables $(r,\theta,z,v_r\,,v_\theta\,,v_z\,,\Phi)$ many of which are hard to observe, and it also depends on their derivatives. So not directly very useful! Simplify by considering a galaxy which is axisymmetric (i.e. derivatives with respect to θ are zero) and in steady state (time derivatives are zero):

$$v_r\frac{\partial f}{\partial r}+v_z\frac{\partial f}{\partial z}+\left(\frac{v_\theta^2}{r}-\frac{\partial \Phi}{\partial r}\right)\frac{\partial f}{\partial v_r}-\frac{v_r v_\theta}{r}\frac{\partial f}{\partial v_\theta}-\frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial v_z}=0 \quad \text{Eqn.}(1)$$

Consider a dynamic variable $q(r,\theta,z,v_r,v_\theta,v_Z)$ which has a value depending on the position in phase space (e.g. the energy). Average value $\langle q\rangle$ or stars in a unit volume:

$$n\langle q\rangle=\iiint qf dv_r dv_\theta dv_z \quad \text{where number density} \quad n=\iiint f dv_r dv_\theta dv_z$$

Take the CBE from Eqn.1, and multiply through by v_r then integrate over velocity space to get a useful equation involving average velocities. The third term is:

$$\iint\left(\frac{v_\theta^2}{r}-\frac{\partial \Phi}{\partial r}\right)dv_\theta dv_z\int v_r\frac{df}{dv_r}dv_r=-\frac{n}{r}\langle v_\theta^2\rangle+n\frac{\partial \Phi}{\partial r}$$

$$\text{since by parts:}\int v_r\frac{df}{dv_r}dv_r=\left[v_r f\right]_{-\infty}^{\infty}-\int f dv_r=-\int f dv_r$$

And similarly for the fourth term. The last term is zero. Hence we get one of the Jeans equations:

$$\frac{\partial}{\partial r}\left(n\langle v_r^2\rangle\right)+\frac{\partial}{\partial z}\left(n\langle v_r v_z\rangle\right)-\frac{n}{r}\left(\langle v_\theta^2\rangle-\langle v_r^2\rangle\right)=-n\frac{\partial \Phi}{\partial r} \quad .$$

Another of Jeans' Equations

Taking the CBE from previously (Eqn.1), and multiply through by v_z then integrate over velocity space to another important Jeans equation: $\frac{\partial}{\partial r}\left(n\langle v_r v_z\rangle\right)+\frac{\partial}{\partial z}\left(n\langle v_z^2\rangle\right)+\frac{n\langle v_r v_z\rangle}{r}=-n\frac{\partial \Phi}{\partial z}$

Easiest to measure velocity dispersion, σ $\quad \langle v^2\rangle=\langle v\rangle^2+\sigma^2$

Why are these equations useful? Can be used to measure the mass density in the Galactic disk, using motions of stars in the solar neighbourhood. The Sun is a long way from the Galactic centre (*r* is large), but relatively close the plane (*z* is small); the vertical gradient should be much stronger than the radial gradient. Plane is symmetric and so:

$\langle v_Z\rangle=0$. Hence $\langle v_r v_Z\rangle\approx 0$ $\quad \frac{\partial}{\partial z}\left(n\langle v_z^2\rangle\right)\approx -n\frac{\partial \Phi}{\partial z}=-ng_z$ vertical grav. field x number density

For a tracer population of stars (perhaps old K-dwarf stars, well-mixed so *f* and *n* not changing with time) measure their number density and velocity dispersion as a function of height above the plane, *z*. Use Poisson equation to get the density of all matter, ρ - Oort used this method to get a limit on the amount of interstellar material $\rho_{tot}\approx 10\times10^{-21}$ kg/m^3 but observed stars: $\rho_{star}\approx 4\times10^{-21}$ kg/m^3.

$$-\frac{\partial^2 \Phi}{\partial z^2}=4\pi G\rho\Rightarrow\frac{\partial}{\partial z}\left(\frac{1}{n}\frac{\partial}{\partial z}\left(n\langle v_z^2\rangle\right)\right)=\frac{\partial}{\partial z}\left(\frac{1}{n(z)}\frac{\partial}{\partial z}\left(n(z)\langle\sigma_z^2\rangle\right)\right)=-4\pi G\rho$$

References

Carroll & Ostlie chapter 2

Sparke & Gallagher chapter 3

Choudhuri chapter 7

See Binney & Tremaine "Galaxy Dynamics" for a much more detailed treatment of this field.

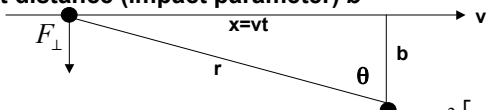
Collisionless Systems

- Stellar motions determined by overall gravitational potential
  - Can neglect individual stellar interactions in many cases
- Rational
  - Gravity is a long-distance force
    - c.f. motion of particles in a box determined by particle interactions
  - Newton's 1st theorem: a body experiences no net force inside a spherical shell
  - If the mass density is uniform, object experiences no net force
  - Net force dependent on gross shape (structure) of system!
  - Not individual forces
- Truly collisionless systems
  - Orbits conserve energy, angular momentum
  - No evolution of the global properties of the system - df(r,v)/dt=0
  - Virial theorem should be true for any system in steady state (not expanding or shrinking) - *virialized*, whether or not it is *collisionally relaxed*

But in what circumstances are collisions important? On what timescale will they "relax" the system - i.e. velocity distribution is set by collisions, erasing "memory" of initial conditions.

Stellar Interactions

- In reality, all stellar systems evolve through individual interactions
- Consider system of *N* stars. Evaluate deflection of a star as it crosses system
- Consider an encounter of a star with mass M with another star of mass *m* at distance (impact parameter) *b*



$$F_\perp = M\frac{dv_\perp}{dt} = \frac{GMm}{r^2}\cos\theta = \frac{GMmb}{(x^2+b^2)^{3/2}} \approx \frac{Gm^2}{b^2}\left[1+\left(\frac{vt}{b}\right)^2\right]^{-\frac{3}{2}}$$

The change in the perpendicular velocity is then

$$\Delta v_\perp = \int\left(\frac{F_\perp}{M}\right)\mathrm{d}t = \frac{Gm}{bv}\int_{-\infty}^{\infty}\frac{1}{(1+s)^{3/2}}\mathrm{d}s = \frac{2Gm}{bv} \qquad \text{where } s=\frac{vt}{b}$$

(Do this integral using the substitution $s=\tan\theta$ and the limits $\theta=-\pi/2 \rightarrow +\pi/2$)

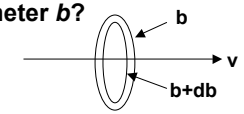
Same result from impulse approximation $\Delta v_\perp = F_\perp \Delta t$

Where $F_\perp$ is the force at closest approach and Interaction duration estimated as $\Delta t = 2b/v$

$$\Delta v_\perp = \frac{F_\perp}{m}\Delta t = \frac{Gm}{b^2}\frac{2b}{v} = \frac{2Gm}{bv}$$

Number of Interactions & Change in K.E.

- How many interactions occur at impact parameter *b*?
- Number density of stars = *n*
- In time *t* travel a distance *vt*
- Number of interactions within $b \rightarrow b+\mathrm{d}b$



$$\delta N = \left[\pi(b+\mathrm{d}b)^2 - \pi b^2\right] nvt \approx 2\pi b\, nvt\, \mathrm{d}b$$

Each interaction has an effect $\Delta v_\perp$, but each one is randomly orientated. Sum is zero: $\langle \Delta v_\perp \rangle = 0$

But kinetic energy $\propto v^2$, and summing over squares $\langle \Delta v_\perp^2 \rangle > 0$

Hence $$\langle \Delta v_\perp^2 \rangle = \left(\frac{2Gm}{bv}\right)^2 2\pi b n v t\, \mathrm{d}b$$

Considering interactions over all *b*: $$\Delta v_\perp^2 \approx \int_{b_{\min}}^{b_{\max}} \frac{8\pi G^2 m^2 nt}{v}\frac{\mathrm{d}b}{b} = \frac{8\pi G^2 m^2 nt}{v}\ln\left(\frac{b_{\max}}{b_{\min}}\right)$$

What is b_{min}? Cannot be 0, as db/b infinite.

Could be stellar radius, or better where $|\Delta v_\perp| < v$ straight-line approximation breaks down, at $b_{\min} \approx \frac{Gm}{v^2}$

What about b_{max}? Size of system - if in virial equilibrium, a system of *N* stars each of mass *m* has a size $b_{\max} \approx R_{vir} = \frac{GmN}{\langle v^2 \rangle}$ so $\frac{b_{\max}}{b_{\min}} = \Lambda \approx N$

Relaxation Time

- Thus, v^2 changes by Δv^2 each time the star crosses the system
- A star has not behaved in a collisionless manner when $v^2 \sim \Delta v^2$
  - This is when a system become virialized and will occur after n_{relax} crossings of the system
- The relaxation time of the system
  - Timescale on which a star's energy changes by order unity
  - Timescale on which "collisions" (interactions) erase memory of initial velocity distribution

$$\frac{\Delta v^2}{v^2} \approx 1 \Rightarrow t_{relax} = \frac{v^3}{8\pi G^2 m^2 n \ln\Lambda} \sim 10^{10}\frac{v^3}{n}\,\text{years}$$

where v in km/s and n in pc^{-3}

The crossing time of the system $t_{cross} \sim R_{vir}/v$ and $N \sim n 4\pi R_{vir}^3/3$

$$\frac{t_{relax}}{t_{cross}} \approx \frac{v^4}{8\pi G^2 m^2 n R_{vir}\ln\Lambda} \approx \frac{v^4 R_{vir}^2}{6G^2 m^2 N \ln N} = \frac{N}{6\ln N} \sim \frac{N}{30} \rightarrow \frac{N}{150} \quad \text{for } N = 100 \rightarrow 10^{11}$$

Implications: if N~10^{11} stars, like our Galaxy, the relaxation time ~$10^8 t_{cross}$~10^{13}years - longer than age of Universe! So NOT important, unrelaxed and collisionless (smooth grav potential) good approx. BUT smaller star clusters have had time to relax.

$$t_{\text{relax}} = n_{\text{relax}} t_{\text{cross}} \approx \frac{N}{6\ln N} t_{\text{cross}}$$

S26 Stars & Galaxies (Short Option) Summary

Often use Solar Units $L_\odot, M_\odot, R_\odot$ - on formula sheet! Stars: 10^{-4}-10^{6} $L_\odot$, 0.1-100 $M_\odot$, 10^{-2}-10^{3} $R_\odot$

DISTANCE SCALE: Astronomical unit (AU) : Mean distance between the Earth and the Sun

1AU=1.50x10^{11} m. Parsec (pc) : Distance at which 1 AU subtends an angle of 1 arc second

1pc=3.26 light-years=3.09x10^{16} m. Nearest star~1pc, galaxy~10^{6}pc; Milky Way radius~10^{4}pc

Distance to nearby stars by trigonometric parallax, motion of Earth around Sun produces angular shift, *p* (in arcsec): distance (in parsec) *d=1/p* . Only reliable to <~a few kpc.

Beyond this use other techniques to measure distance (the "distance ladder"), standard candles such as RRLyrae stars (variable stars all same luminosity, recognisable light curve with 13hr period) and brighter Cepheid variables (period related to luminosity).

From RRLyraes, Galactic centre is 8kpc away; Cepheid distances to nearby galaxies to determine Hubble constant (current expansion rate of universe, $v=H_0 d$, H_0=71km/s/Mpc). At greater cosmological distance, use very luminous Supernova Type Ia for distance.

S26 Stars & Galaxies (Short Option) Summary

Observed flux *f* related to intrinsic luminosity $L=4\pi d^2 f$ Use magnitude system to measure log(flux); stars visible by eye have m=0→6 (bright→faint). Star Vega has *m=0* (definition) in any waveband.

$m_1 - m_2 = -2.5 \times \log_{10}(f_1/f_2)$. Apparent magnitude (m) related to the observed flux;

absolute Magnitude (M) is the mag that would be recorded at a distance of 10pc (related to luminosity)

Distance modulus $m_\lambda - M_\lambda = 5\log_{10} d - 5 + A_\lambda$ where *A* is extinction in magnitudes and λ is the filter in which observation is done in (UV,blue,visible,red,IR called **UBVRI**).

Interstellar dust causes extinction (A) - greater for blue wavelengths (i.e. reddening).

Can relate the total bolometric flux (integrated over all wavelengths) to that measured in a particular filter by bolometric correction (B.C.) for that filter (in mags): $m_{bol}=m_\lambda$**+B.C.**

Stellar Classification

Stellar photospheres can be approximated as black bodies:
$L=4\pi r^2\sigma T_{eff}^4$ (effective temp.)

Wien's displacement law: $\lambda_{max} = 2.898 \times 10^{-3}/T$ - hotter stars are blue, cooler red.

Define a colour - ratio of fluxes (or difference in mags) in two filters: $m_B-m_V=(B-V)$

Bigger colour means redder, and cooler surface temperature. Dust can increase the colour (reddening): colour excess E compared to the intrinsic colour $(B-V)_0$: $E(B-V)=(B-V)-(B-V)_0$
and for dust in the Milky Way: $A_V=3.2x\ E(B-V)$
[you would be given this relation in an exam]

Stars are classified by their instrinic colours, or surface temperatures, on the Harvard classification **OBAFGKM** where O is hottest (30000-50000K blue) and M coolest (3000K red)
With sub-divisions 0-9 in each (our Sun is a G2 star with T_{eff}=6000K, and Vega is A0).

Stellar Classification

Sequence originally alphabetical with strength of hydrogen balmer absorption dropping A→M

A-stars have T_{eff}=10000K; in cooler stars, not much H in n=2 level, in hotter stars much of H is ionized (HII). Strength of absorption lines governed by Boltzmann equation (population of energy levels) and Saha equation (ionization state). In cooler stars (G,K) see neutral metal absorption lines, even molecular in M stars. Ionized He in hot O stars

For stars, plotting luminosity (y-axis) against temperature (x-axis reversed, T_{eff} traditionally increased to left) in log units is the **Hertzsprung-Russell (HR) diagram**.

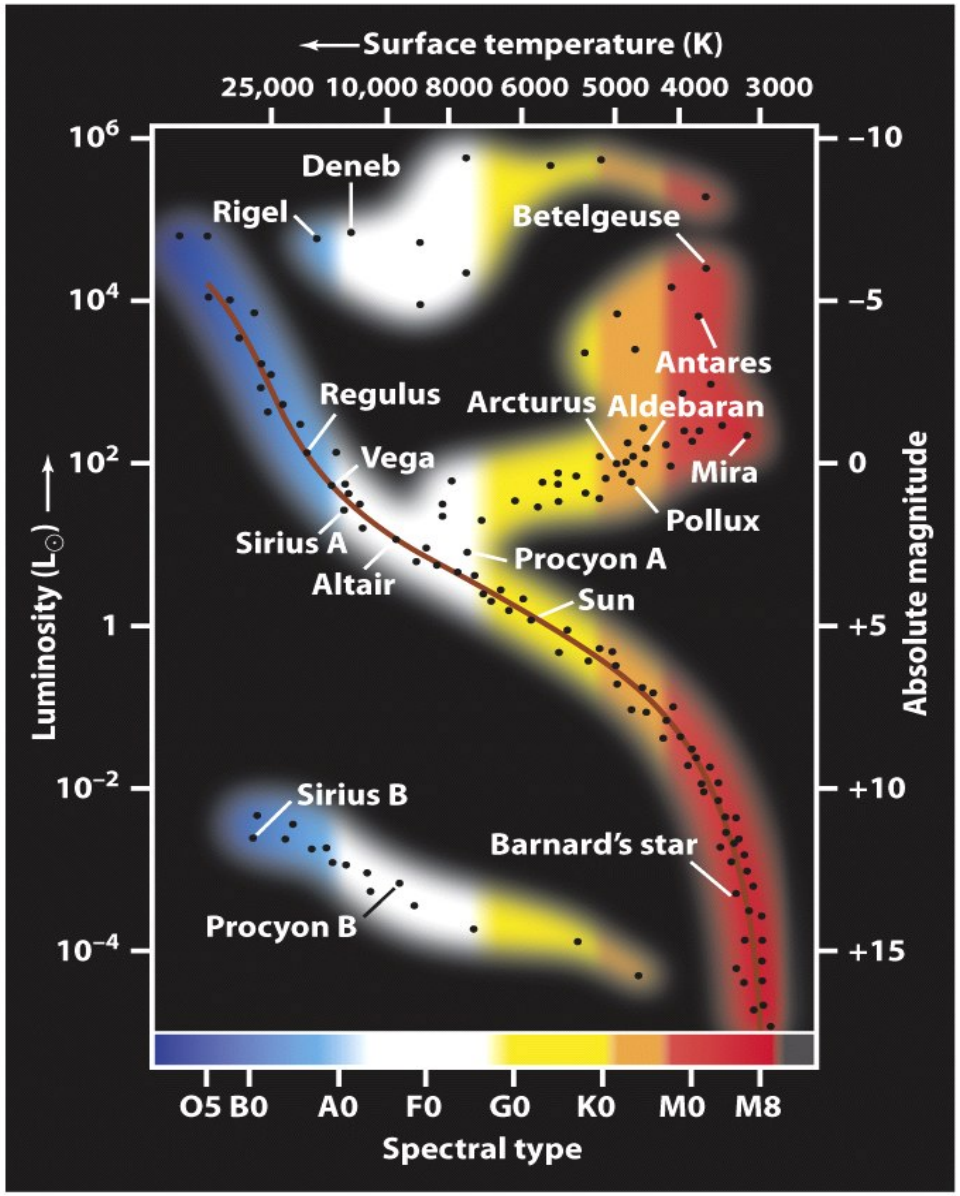
Stars not randomly distributed in HR diagram - fall in a few regions, depending on luminosity class classified from surface gravity from width of spectral lines, Morgan-Keenan class I (supergiants) - VI (sub-dwarfs).

90% of stars are class V, the **Main Sequence** (MS) - a narrow strip going from hot, luminous stars (high mass short-lived) to cool, faint stars (low mass long-lived). Stars spend 90% of their lives on the MS, fusing H in their core to He. Knowing absolute mags, *M*, of stars on MS for each temperature (colour), can get the distance modulus *m-M* and hence distance by plotting the apparent magnitude, *m*, against colour (temperature) on an H-R diagram for a cluster and shifting up/downto match theoretical H-R MS – *main sequence fitting*.

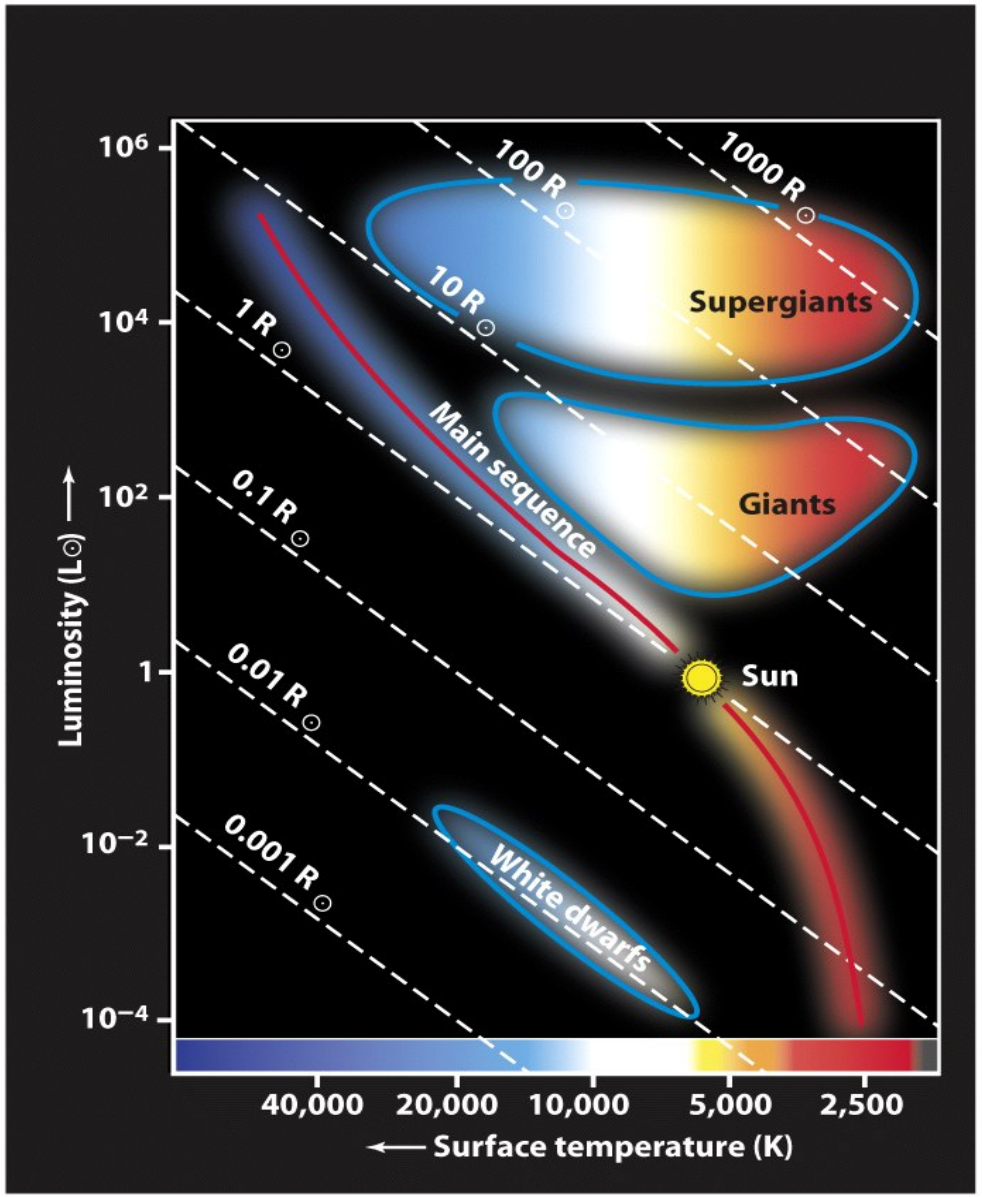
Stars remain at same approx place on MS. Location determined by mass: $L_{MS} \propto M^{\nu}$, $\nu=3 \rightarrow 5$

H-R diagram

spans a factor of 10^{10} in luminosity and from T=100,000K to 2000K in temperature



(a) A Hertzsprung-Russell (H-R) diagram



(b) The sizes of stars on an H-R diagram

Figure 17-15
Universe, Eighth Edition
© 2008 W. H. Freeman and Company

Remember: $L=4\pi R_s^2 \sigma T_{eff}^4$

Stellar Structure

Hydrostatic equilibrium - balance inward gravitational pull on a shell at radius r with pressure gradient (learn derivation) $\frac{\partial P}{\partial r} = -\frac{GM(r)\rho(r)}{r^2}$

Eqn of mass continuity: $\frac{\partial M}{\partial r} = 4\pi r^2 \rho(r)$

Dynamical (free-fall) time $\approx \frac{1}{\sqrt{2G\bar{\rho}}}$ (~40min for Sun)

Approximate $dP/dr \sim -P_C/R_S$
Estimate central pressure from point $r=R_S/2$ and $M(r)=M_S/2$.

Central temp from ideal gas law $P=\rho k_B T/(\mu m_H)$
where μ is mean particle mass (1/2 if ionized H). $T_C \sim 2\times10^7$K

Virial Theorem: $3(\gamma-1)U+\Omega=0$, U=thermal energy, Ω=grav. energy.

Learn derivation: multiply hydrostat.eqm. By $4\pi r^3$, integrate by parts, use internal energy per particle $u=fk_BT/2$, f=degrees of freedom, $\gamma=C_P/C_V=(f+2)/f$. For ideal gas, $\gamma=5/3\rightarrow 2U+\Omega=0$

Stellar Structure

Stars have negative heat capacity; if lose energy, contract & heat up. Self-regulates fusion. From U (Virial theorm) work out thermal timescale=$U/L_S=GM_S^2/(2R_SL_S)$~1.5×10^7yr for Sun.

Sun has shone for longer, require nuclear energy, ε=nuc.energy/s/kg

$$dL_r/dr = 4\pi r^2 \rho_r \epsilon_r$$

Energy Transport often by radiative transfer - photons scattering, random walk out of star in 30,000yr (Sun), degraded from X-rays to visible wavelengths.

Radiation pressure $P_{rad}=aT^4/3$
κ=opacity (in m^2/kg) & mean free path=$(\kappa\rho)^{-1}$. Use rate of momentum absorbed in a shell:

$$\frac{L_r}{c}\kappa\rho dr = -\frac{dP_{\rm rad}}{dr}4\pi r^2 dr = 4\pi r^2 dr\,\frac{\rm d}{dr}\left(\frac{aT^4}{3}\right) \Rightarrow L_r = -T_r^3\,\frac{dT_{\rm r}}{dr}4\pi r^2\,\frac{4ac}{3\kappa_r\rho_r}$$

Usually given this, but learn derivation

For massive stars, κ~constant. Get $L\propto M^3$ (MS - learn derivation). Lifetime ~M/L~M^{-2}
At lower temp $\kappa\sim\rho T^{-3.5}$. Leads to $L\propto M^5$ with assumption core T similar for stars on MS.
Very massive stars (>$100M_\odot$) supported against collapse by radiation pressure

Stellar Structure

Convective transport - bulk motion of matter; important where steep temp gradients (cores massive stars), or where neutral/ionized mix $\gamma \rightarrow 1$ (outer envelopes lower-mass stars).

Assume adiabatic $PV^{\gamma}=\text{const}' \rightarrow P/\rho^{\gamma}=\text{const}$, & with $P=\rho k_B T/(\mu m_H)$ derive buoyancy condition

Schwarzschild stability criterion for convection to occur (*learn derivation)*

$$\left|\frac{dT}{dr}\right| > \left(\frac{\gamma-1}{\gamma}\right)\frac{T}{P}\left|\frac{dP}{dr}\right|$$

Slow turbulent motions (hydrostatic equilib. holds).

Equation of state $P=\rho k_B T/(\mu m_H)$ ideal gas law.

If fully ionized, $4/\mu=(6X+Y+2)$

Where mass fraction X(hydrogen), Y(helium), Z(other elements, "metals"). X+Y+Z=1

Solve stellar structure equations with BCs: $M_r=0$ at r=0; at $r=R_S$ P=0 and $T=T_{eff}$ (small)

Stellar Structure

Nuclear energy production fusion of H to He in cores of main sequence (MS) stars.

Need high temp and pressure to overcome Coulomb repulsion, barrier $E=Z_1Z_2/(4\pi\varepsilon_0 r)$

Quantum mechanical tunneling, probability $P \propto \exp(-(E_G/E)^{1/2})$, E_G is Gamov energy, to get within range of strong nuclear force ($\sim 10^{-15}$m)

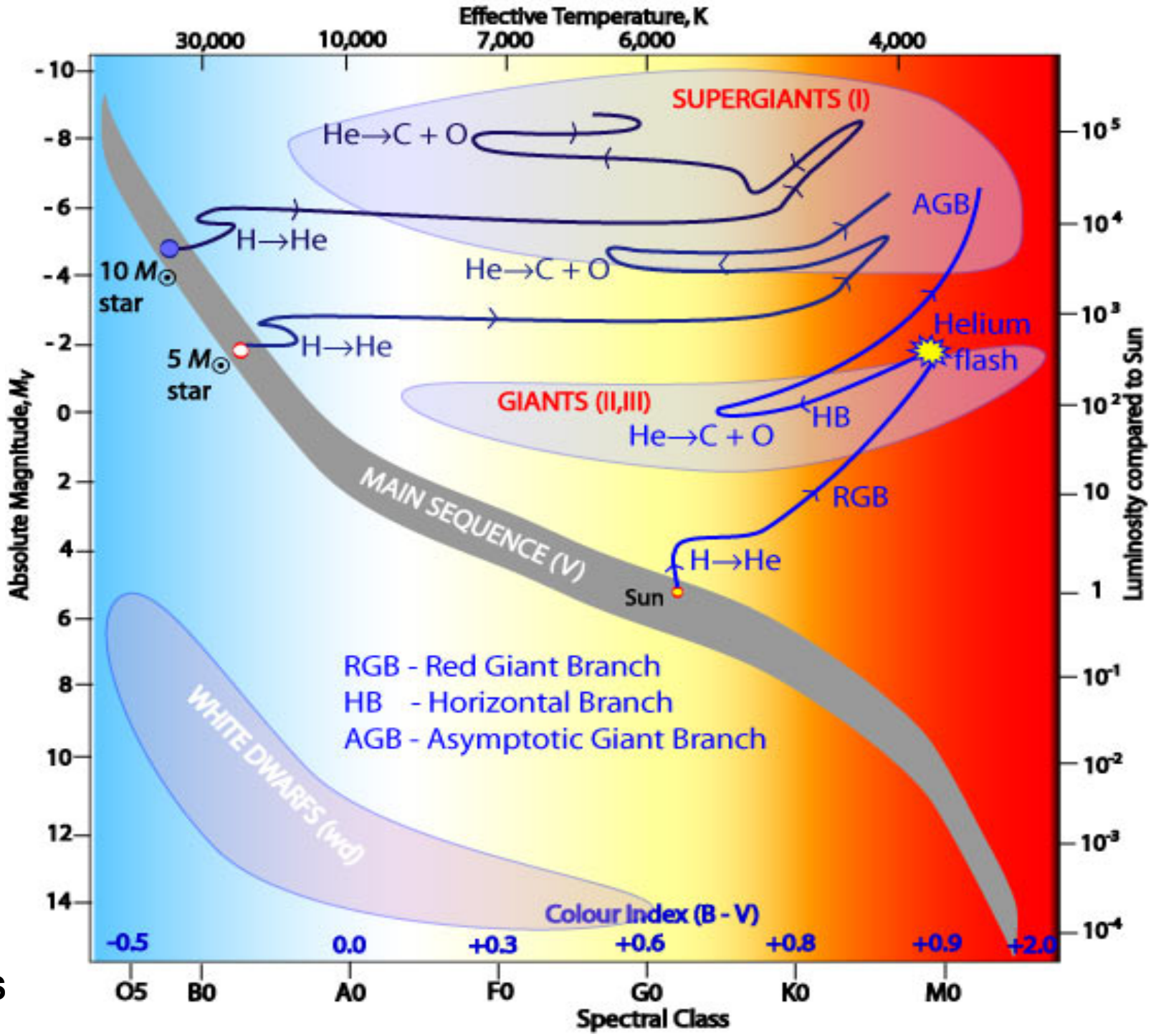
In stars up to $1.5M_\odot$ **PP chain (proton-proton chain reaction)** main energy on MS.

- $p + p \rightarrow {}^2D + e^+ + \nu_e$ +1.44MeV ; typical reaction time 1.4×10^{10}yr BOTTLENECK
- ${}^2D + p \rightarrow {}^3He + \gamma$ + 5.49MeV ; t~6 seconds
- ${}^3He + {}^3He \rightarrow {}^4He + 2p$ + 12.85MeV ; t ~10^6 years (has two other branches, PP-II&III)

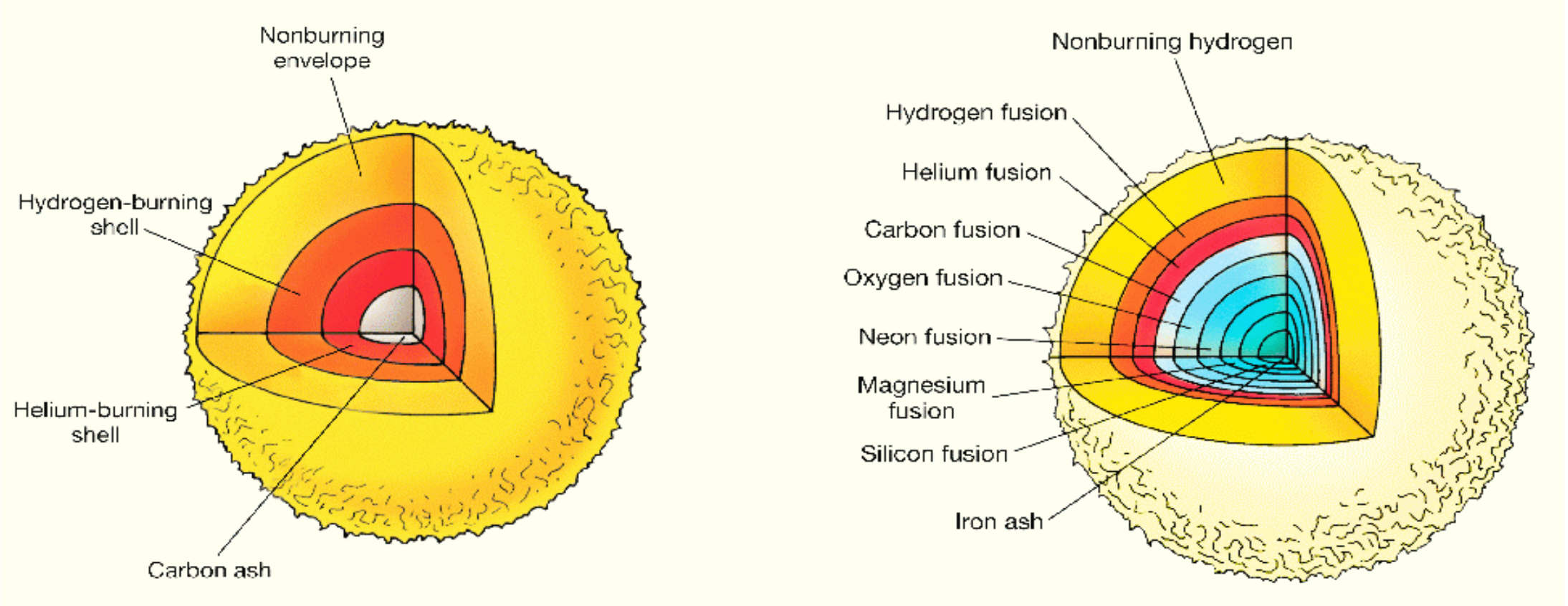
For PP chain, energy production $\varepsilon_{pp} \propto T^4$ .In more massive stars, the CNO cycle (those elements acts as catalysts) achieves H fusion, but need hotter temperatures (Coulomb barrier higher due to greater atomic charge Z on CNO).
$\varepsilon_{CNO} \propto T^{17}$, large temperature gradients mean convective cores for >$1.5M_\odot$.

After main sequence, core shrinks and heats up, enough energy for He fusion to carbon - triple-α process (no stable A=8 nucleus, so need high temp & pressure to get another He fusion quickly): $\varepsilon_{3\alpha} \propto T^{40}$

Evolutionary Tracks off the Main Sequence



Heavier element burning



White Dwarf Stars

No nuclear processes, supported against collapse by degenerate electron pressure rather than ideal gas pressure.

$$\Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z = h^3$$

There can only be 2 electrons (spin up/down) per volume of phase space size.

Electron conduction v. effective - isothermal

Pauli exclusion for spin-1/2 electron (Fermion). Such high density that all levels up to the Fermi level (momentum p_F) are full.

N_e=total number of electrons

n_e=density

Use the relation

$$N_e = \frac{8\pi V}{h^3}\int_0^{\pi_F} p^2\, dp = \frac{8\pi p_F^3 V}{3\mathrm{h}^3} \qquad p_F = \left(\frac{3\mathrm{N}_e h^3}{8\pi V}\right)^{1/3} = \left(\frac{3\mathrm{n}_e h^3}{8\pi}\right)^{1/3}$$

and p=mv if non-relativistic to get

$$P = \frac{1}{3}\int_0^{\infty} \frac{\mathrm{N_e}}{\mathrm{V}} p v_\mathrm{e} dp$$

$$P = \frac{1}{20}\left(\frac{3}{\pi}\right)^{2/3} \frac{h^2 n_e^{5/3}}{m_e} \propto \rho^{5/3}$$

White Dwarf Stars

If the electrons are relativistic, p~E/c and obtain:

$$P = \frac{1}{8}\left(\frac{3}{\pi}\right)^{1/3} h n_e^{4/3} \propto \rho^{4/3}$$

White dwarfs mostly carbon & oxygen, so:
where X = mass fraction of hydrogen

$$n_e = \frac{\rho X}{m_H} + \frac{\rho(1-X)}{2\mathrm{m}_H} = \frac{\rho(1+X)}{2\mathrm{m}_H} \approx \frac{\rho}{2\mathrm{m}_H}$$

White dwarf cooling: radiates thermal energy as luminosity

$$-\frac{d}{dt}\left(\frac{M_S}{Am_H}\frac{3}{2}k_B T_C\right) = \mathrm{const\ x}\ {T_\mathrm{C}}^{7/2} = 4\pi R_S^2 \sigma T_{eff}^4$$

Mass-radius relation from $P_C \approx \frac{GM_S^2}{R_S^4}$ and $\bar{\rho} = \frac{M_S}{\frac{4}{3}\pi R_S^3}$

for non-relativistic, $P \propto \rho^{5/3}$ so $R \propto M^{-1/3}$ (heavier WDs smaller!)

For relativistic case ($P \propto \rho^{4/3}$) no relation between M and R. Maximum mass when gravitational binding energy equals thermal energy, Chandrasekhar mass

$$M_{Ch} \sim 1.46\left(\frac{2}{\mu_e}\right)^2 M_{Sun}$$

The Milky Way - Our Galaxy

A spiral galaxy - main components: disk (~10kpc radius, ~300pc thick), bulge (~2kpc), halo.

Population I stars - ***younger, high metallicity***, found mainly in disk (Galactic plane)
Population II stars - ***older, metal poor***, found in halo and globular clusters (these are gravitationally-bound clusters of 10^4-10^6 old stars, about 150 orbiting Milky Way in halo).
Halo stars on random inclined orbits ("hot"), disk stars in ordered differential rotation ("cold")

Galaxy composed of stars, dark matter and the **interstellar medium** (ISM) dust and gas.

Gas can be atomic hydrogen (HI in disk, 21cm hyperfine transition radio emission, T~100K), molecular clouds (compact, cool T~10K, dense - regions of star formation) and ionized (due to hot stars T~5x10^4 - HII regions, or supernova remnants [coronal gas], v.hot 10^6K).
Map 21cm emission radial velocity (V_r) as a function of Galactic longitude (l):

$$V_r = R_0(\omega - \omega_0) sin\, l = \left(\frac{R_0}{R}V - V_0\right) sin\, l$$

where solar orbit has R_0, ω_0 .

Points on solar circle: constant separation: $V_r = 0$

$$\text{At } r \to \infty, \omega \to 0 \text{ and } V_r = 0 - V_0 sin\, l = -V_0 \text{ at } l\text{=90deg}$$

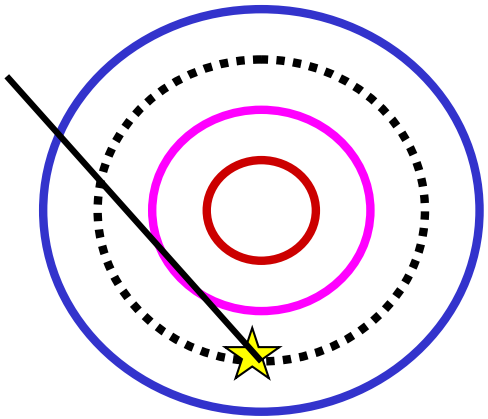
Tangent point: max radial velocity in line of sight $V_{max} = V_{ring} - V_0 \sin l$

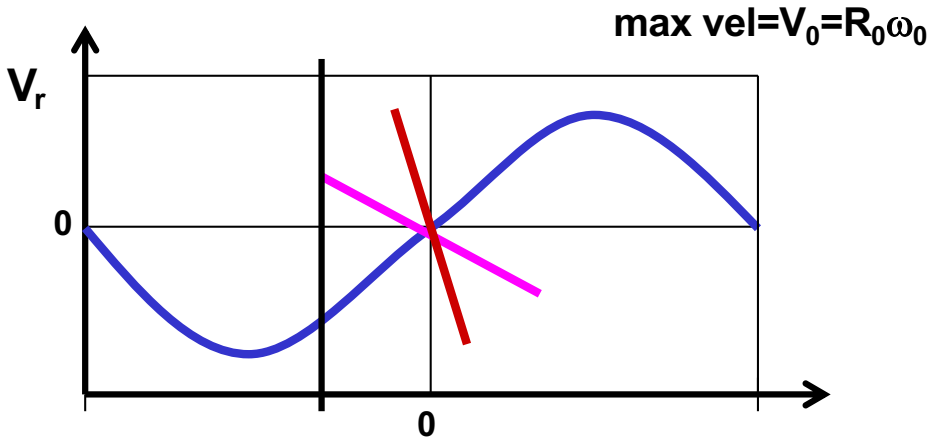
Emission from Rings

Consider 3 rings of emitting gas

- with $R \ll R_0$
- with $R < R_0$
- with $R > R_0$

$$V_r = R_0(\omega - \omega_0)\sin l = \left(\frac{R_0}{R}V - V_0\right)\sin l$$

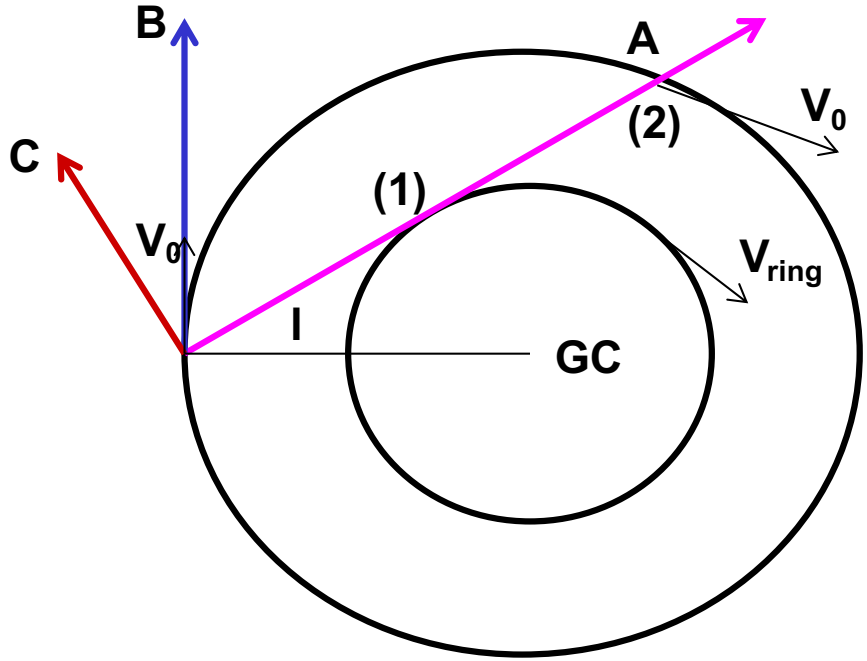


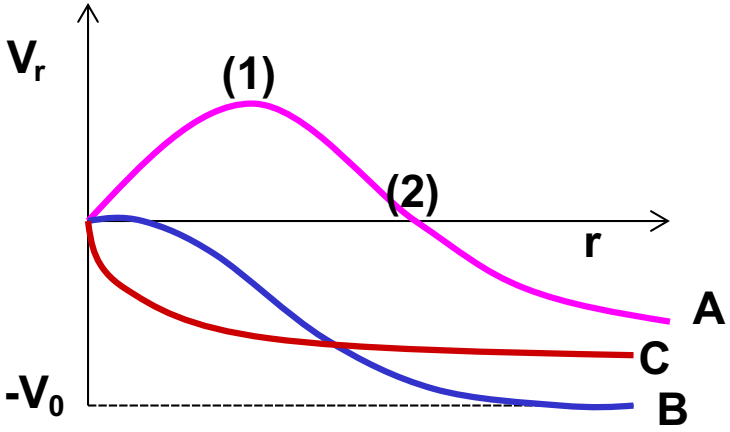


- Inner rings: highest velocity at tangent point
  - Positive velocities for l>0 (ring rotates faster than Sun)
- Outer rings: full sin curve
  - Negative velocities for l>0 (ring rotates slower than Sun)
  - Max velocity at R=inf is rotation speed of local standard of rest (LSR)

Emission from Galactic Disc

Consider 3 lines of sight:





A: (1) Tangent point: max radial velocity in line of sight
$V_{max}=V_{ring} - V_0 \sin l$
(2) Points on solar circle: constant separation: $V_r=0$

B: $\text{At } r \to \infty, \omega \to 0 \text{ and } V_r = 0 - V_0 \sin l = -V_0$

Galactic Components

Disc

- Population I stars
- Interstellar medium (ISM): Gas, dust
- Thick disc and thin disc

Halo

- Population II stars (old)
- Globular Clusters (ensembles of 10^4-10^6 stars held together by mutual gravity, ~150 in Galaxy, most within 3 kpc of centre, some out to ~25kpc)

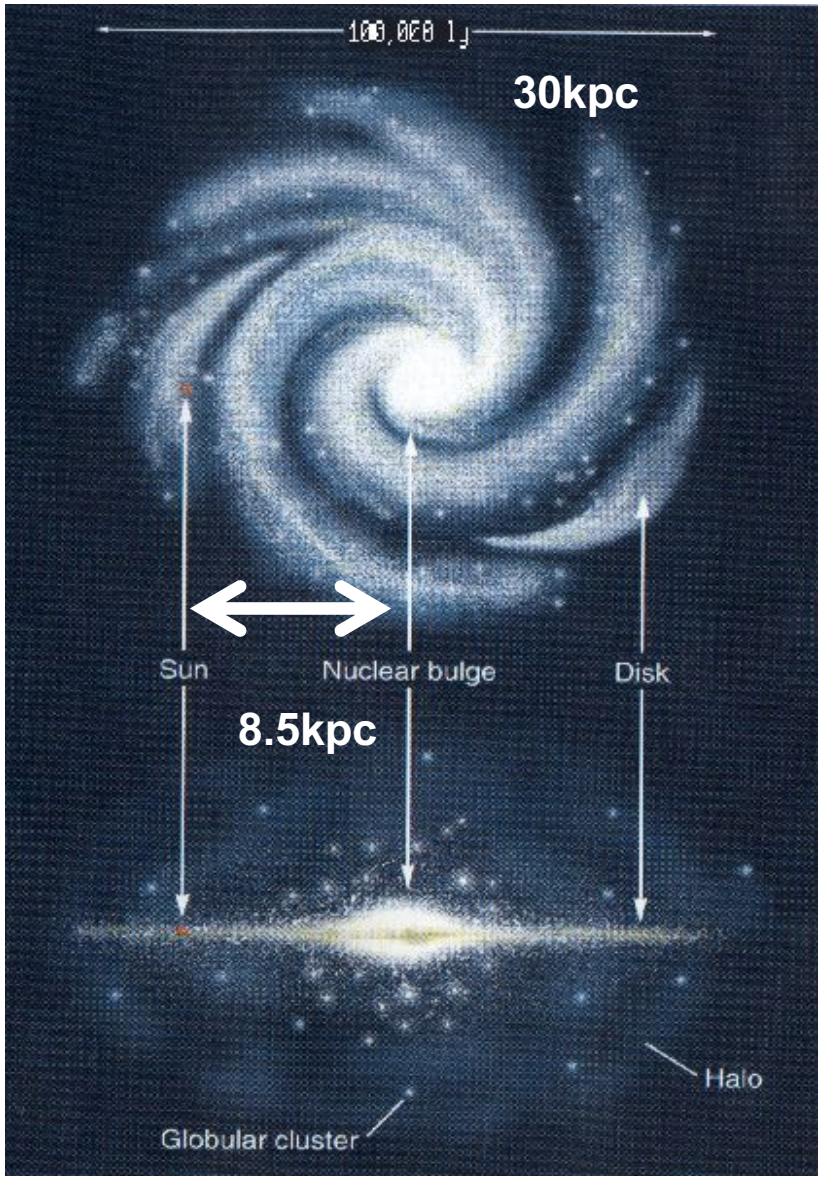
Bulge

Galactic Centre (Nucleus)

- Black hole

Dark matter; gas and dust (the interstellar medium)

What do the individual components tell us about Galactic structure and formation?



Gene Smith, UCSD

The Milky Way - Our Galaxy

Disk is **exponential light profile** $I=I_0\,exp(-r/r_d)$ disk scale length in Galactic plane r_d~3kpc

Also more rapid exponential decline perpendicular to plane (in z direction), H~100pc

Around Sun, disk has projected surface density 100stars/pc^2, & on average each star 0.3$M_\odot$

Estimate disk mass πR^2 and R~10kpc to be $10^{10}M_\odot$ increasing to 4x $10^{10}M_\odot$ including proper integration of exponential profile (r_0=8kpc for Sun) and stellar remnants (dead stars).

BUT circular velocity of Sun in orbit v_0=220km/s, and Galactic mass $M \sim v_0^2 r_0/G$ (centripetal)

So $M_{dynamical}=10^{11}M_\odot$ - need 60% of mass within Sun's orbit to be invisible **dark matter**

For constant speed rotation curve (derived from HI 21cm observations at different Galactic longitudes, l) need $\rho(r) \propto r^{-2}$

Galaxies

Hubble Tuning Fork Diagram classify shape/morphology: elliptical/S0-lenticular/[barred]spiral

Elliptical galaxies - triaxial. Called E0-E7 (sphere to cigar shaped)

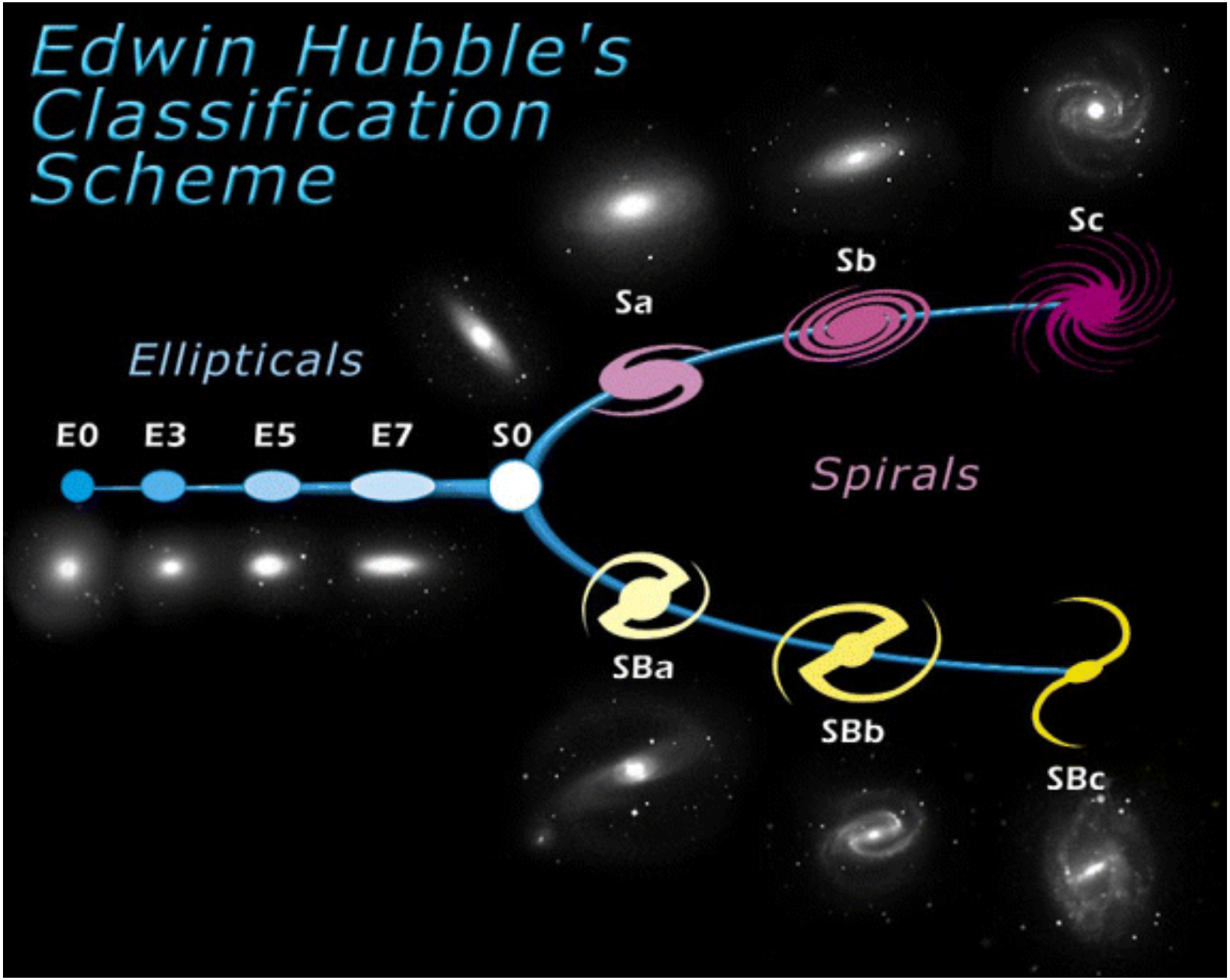
Apparent minor/major axial ratio=b/a and En class, n=10x(1-b/a).

Older stars, redder stellar populations (no short-lived masive blue stars), no current star formation, little gas or dust.

Orbits are predominantly random - ellipticity NOT due to rotation (vel.dispersion $\sigma >> v_{rot}$)

Surface brightness profile (isophotes), de Vaucouleur's $R^{1/4}$ law: $\log_{10}(I/I_e) = -3.33[(r/R_e)^{1/4} - 1]$ where R_e encloses half the light.

Faber-Jackson law: luminosity $L \propto \sigma^4$. Mass from virial theorem: $\langle \sigma^2 \rangle \sim GM/R$


Edwin Hubble's Classification Scheme
Ellipticals
E0
E3
E5
E7
S0
Sa
Sb
Sc
Spirals
SBa
SBb
SBc

Galaxies

Spiral galaxies Sa (tightly-wound, bulge is prominent), Sb (intermediate), Sc (spiral arms loosely would, disk dominates over bulge).

Star formation greatest in Sc galaxies (and irregular dwarf galaxies), also has most gas.
S0 lenticulars have a disk but no arms.
Spirals have exponential surface brightness profile with about same central brightness (Freeman law).
Well-ordered rotation. Have flat rotation curves $v_{rot}(r)$ demands dark matter.
Scaling relation in spiral galaxies, total mass $M \propto v_{rot}^4$ (Tully-Fisher relation)
Spiral arms - winding problem (stars near centre complete thousands of orbits, but only a few for outer stars over age of galaxy, so why are arms not wound up?). Spiral arms are density wave perturbations ("grand-design" spirals), except in a few "stochastic" or "flocculent" spirals with short-lived filamentary arms.

Barred spirals - bar has older stars like bulge. Bar supported by disk self-gravity.
Bars and spiral arms transport matter and angular momentum from the outer parts to the center - bulge grows with time, and spiral arms wind-up (so Sc→Sb→Sa evolution).

Luminosity function - galaxy number density in a luminosity interval, Schechter function
$n(x)\,\delta x = \phi_* \, x^{-\alpha} \, e^x \, \delta x$ where $x = L/L^*$ $\alpha \sim 1.3$ and $L^* \sim 10^{10} L_\odot$, $\phi^* \sim 0.01 \mathrm{Mpc}^{-3}$