

# GENERAL QUALIFYING EXAM SOLUTIONS: STELLAR ASTROPHYSICS

---

Jessica Campbell, Dunlap Institute for Astronomy & Astrophysics (UofT)

## Contents

<b>1 Stellar Astrophysics</b>	<b>2</b>
1.1 Question 1 . . . . .	2
1.2 Question 2 . . . . .	10
1.3 Question 3 . . . . .	11
1.4 Question 4 . . . . .	16
1.5 Question 5 . . . . .	18
1.6 Question 6 . . . . .	24
1.7 Question 7 . . . . .	27
1.8 Question 8 . . . . .	28
1.9 Question 9 . . . . .	38
1.10 Question 10 . . . . .	41
1.11 Question 11 . . . . .	44
1.12 Question 12 . . . . .	45
1.13 Question 13 . . . . .	49
1.14 Question 14 . . . . .	56
1.15 Question 15 . . . . .	57
1.16 Question 16 . . . . .	58
1.17 Question 17 . . . . .	59
1.18 Question 18 . . . . .	60
1.19 Question 19 . . . . .	61
1.20 Question 20 . . . . .	62
1.21 Resources . . . . .	63

# 1 Stellar Astrophysics

## 1.1 Question 1

Sketch out a Hertzsprung-Russell diagram. Indicate where on the main sequence different spectral classes lie. Draw and describe the post main-sequence tracks of both low- and high-mass stars.

### 1.1.1 Short answer

Answer.

### 1.1.2 Additional context

**Evolution of a  $9 M_{\odot}$  star:** Figure 1 shows the calculated evolutionary path on the HRD of a  $9 M_{\odot}$  star of solar composition. Letters mark critical points in the course of evolution. Specifically, the various points mark the following events:

- A: Beginning of steady hydrogen burning, ZAMS.
- C-C0: Exhaustion of hydrogen in the core, and ignition of hydrogen burning in a shell surrounding the hydrogen-exhausted core.
- E: Arrival on the Hayashi line, that is, the envelope is (almost) fully convective.
- F: Ignition of helium burning in the center.
- J: Exhaustion of helium in the core.
- JK: Ignition of helium burning in a shell surrounding the helium-exhausted core.
- L: Back to the Hayashi line (fully convective envelope).
- LM: Early asymptotic giant branch phase (E-AGB).

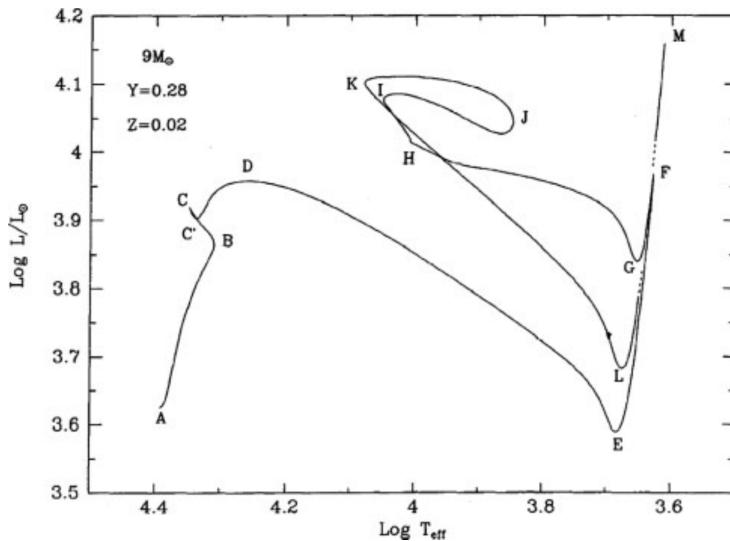


Figure 1: The evolutionary track in the HRD diagram of a  $9 M_{\odot}$  star with solar composition from the ZAMS all the way to the AGB phase (hydrogen and helium burning in two separate shells). What happens at the labeled points. Source: Renzini et al. (1992, *Astrophys. J.*, 400, 280). Figure taken from Gregg & Renzini (2011).

Before disclosing what happens at points BDGHI and K (not mentioned in the previous list) it is necessary to introduce a few concepts in the language of stellar model makers. The **core** is the innermost part of the star, where nuclear reactions have greatly altered the original composition. By **envelope** one means the outer part of the star, over (most of) which nuclear burning is negligible and whose composition may still be close to the original one.

Next is the concept of **thermal equilibrium**. *A star is said to be in thermal equilibrium (TE) when the total rate of nuclear energy generation ( $L_N$ ) is almost perfectly equal to its surface luminosity ( $L_S$ ).* TE means that the envelope is able to transfer outside and radiate into space quite precisely the same amount of energy which per unit time is produced by nuclear reactions in the core. If  $L_S \simeq L_N$  the star is in TE and its evolution proceeds on a nuclear timescale. Conversely, if  $L_S \neq L_N$  the star is out of TE, and its evolution proceeds on a thermal timescale, which usually is much shorter than the nuclear

timescale. TE is often broken when the core runs out of fuel or when a new fuel is ignited. A dramatic example of breaking TE is the helium ignition in a degenerate core, called a **helium flash**. At flash peak  $L_N \simeq 10^{10} L_\odot$  while  $L_S \simeq 2000 L_\odot$  (and decreases!). The energy that does not escape the star is used to expand the helium core, hence relieving its degeneracy.

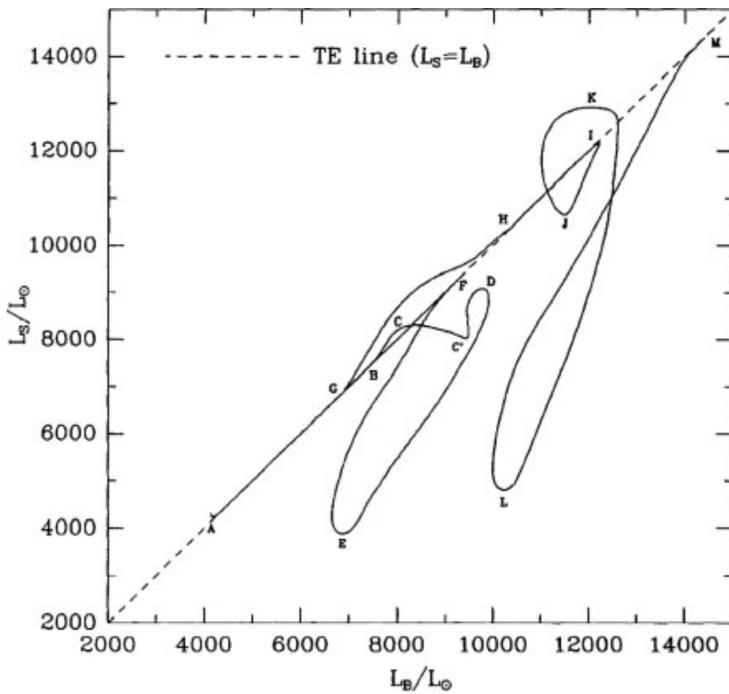


Figure 2: The luminosity radiated from the star's surface as a function of the luminosity released by the stellar core and entering the envelope from its base, for the same track shown in Figure 1. Labels of relevant points along the sequence are the same as in Figure 1. Source: Renzini et al. (1992, *Astrophys. J.*, 400, 280). Figure taken from Gregg & Renzini (2011).

Evolutionary phases in TE and those out of TE can be easily identified by plotting  $L_S$  versus  $L_N$ , or, equivalently, versus the luminosity impinging at the base of the envelope ( $L_B$ ), as shown in Figure 2 relative to the evolution of the same  $9 M_\odot$  star model whose HRD is shown in Figure 1. Letter labels in this figure mark the same events as in Figure 1, so that one can identify those phases that are in TE and those which are out of it.

Thus, phase AB (which is most of the main sequence phase) proceeds in strict TE, but as hydrogen approaches exhaustion in the core, nuclear burning starts to fall short of keeping in pace with the rate at which energy is being radiated away, the star starts departing from TE and begins to contract (point B). At point B, as the core is running out of fuel the envelope starts losing more energy than it receives, and contracts until at point C hydrogen is effectively exhausted over the central regions, and nuclear energy generation quickly shifts to a shell surrounding a hydrogen-exhausted helium core. The quick readjustment from core to shell burning leaves the star at point C', somewhat out of TE, but then the structure tends to approach again TE, until at point D this tendency is reverted and a major excursion away from TE begins. In stars a little less massive than the one considered here, TE is actually well restored shortly after point C', and yet (as here at point D) TE is broken and stars undergo an extensive loop in the  $L_S - L_B$  diagram, as shown in Figure 2. The journey across the HRD from point D to point E takes place on a thermal timescale, and the star expands to red giant dimensions. Clearly, a thermal instability erupts at point D. This is indeed a quite severe thermal instability, suffice to say from Figure 2 that at the peak of the instability the core releases  $\sim 7000 L_\odot$  but the envelope radiates away only  $\sim 4000 L_\odot$ , and  $\sim 3000 L_\odot$  are absorbed for its expansion.

The physical origin of this thermal instability is actually quite easy to understand. During phase C'D the luminosity provided by the hydrogen burning shell steadily increases as the shell moves out in the mass coordinate thanks to its own burning, and sits on a progressively more massive helium core. In response to the increasing luminosity impinging on its base ( $L_B$ ) the envelope slowly expands. By expanding the envelope cools, and by cooling heavy metal ions begin to recombine. Besides being scattered by free electrons, photons now begin to be absorbed by such heavy ions via bound-bound and bound-free transitions: *radiative opacity increases*. This opacity increase is the key factor that determines the onset of the thermal instability. At any point within the star the luminosity transmitted outwards by the radiation field is:

$$L_r = 4\pi r^2 \frac{4acT^3}{3\kappa\rho} \frac{dT}{dr} = 4\pi r^2 F_r \text{ [erg s}^{-1}\text{]},$$

where  $r$  is the distance from the center,  $T$  the temperature,  $\rho$  the density,  $\kappa$  the opacity, and  $F_r$  the

radiative energy flux. During phase C'D the envelope slowly expands, hence  $r^2$  increases while the flux  $F_r$  decreases, but their product still increases and the star is approaching TE. However, as this trend continues the increase in opacity accelerates and eventually the flux drops faster than  $r^2$  increases, and their product  $L_r$  starts to decrease. The decrease happens first near the stellar surface, and then (very quickly) through the whole envelope. At this point the envelope is transferring outwards and radiating away less energy than it receives from the stellar core, that is,  $\Delta L = L_B - L_S > 0$ . But as the envelope expands more ions recombine, opacity increases further, the flux drops even more, the thermal imbalance  $L_B - L_S$  increases and expansion accelerates: the stellar envelope is in a thermal runaway, as it becomes more and more unable to radiate away the energy it receives from the stellar interior.

At point E the surface luminosity starts to rise again,  $L_B - L_S$  begins to drop, and TE is rapidly restored. What relieves the instability and saves the star from literally falling apart is convection. As the expanding envelope cools, and opacity increases, eventually the radiative gradient  $\nabla_{\text{rad}}^1$  exceeds the adiabatic gradient  $\nabla_{\text{ad}}$ , first near the photosphere, where hydrogen is only partly ionized, and then rapidly through the whole envelope. Thus, convection replaces radiative transfer in carrying out energy through the envelope, and the thermal instability is quenched since it was intimately related to the radiative mode of energy transfer. Most of the energy flux in the envelope being now carried by convective motions, the envelope ceases to absorb energy, and the surface luminosity  $L_S$  starts to increase again: the star now ascends the Hayashi line, until at point F the helium burning reactions ignite in the core.

Following helium ignition, the helium core initiates a very slow expansion, which will last through a major fraction of the helium burning phase. This is because the helium burning core works like a breeder reactor during this stage, that is, it produces more new fuel than it burns. As triple- reactions produce fresh  $^{12}\text{C}$ , the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reactions release an increasing amount of energy, and the inner core is forced to expand slightly. Such a modest expansion of the core is sufficient to cause a decrease of temperature and density in the surrounding hydrogen burning shell and the stellar luminosity correspondingly starts to decrease. The star now evolves along the **Hayashi line** from F to G, burning helium in the core and hydrogen in the shell, until thermal stability is broken again at point G. As the envelope contracts it heats up gradually, and in particular near its base, heavy ions start losing electrons. Hence opacity decreases along with the radiative gradient. As  $\nabla_{\text{rad}}$  drops below  $\nabla_{\text{ad}}$  a radiative region appears at the base of the envelope and grows outwards in mass. As this growth progresses the envelope becomes more and more transparent to radiation, until it starts radiating away more energy than it receives from the core:  $L_B - L_S$  turns increasingly negative, the envelope starts deflating, but the more it contracts the more “transparent” it becomes, and the more energy it loses into space. The thermal instability bringing the star in its envelope deflation from point G to point H is precisely the reverse analog of the thermal instability that causes the runaway expansion from point D to point E. As envelope inflation can be ascribed to a runaway recombination of heavy elements in the envelope, envelope deflation is due to a runaway ionization of such heavy elements. A comparison of the three figures (Figures 1-3) helps in visualizing the onset of the thermal instability and of its results.

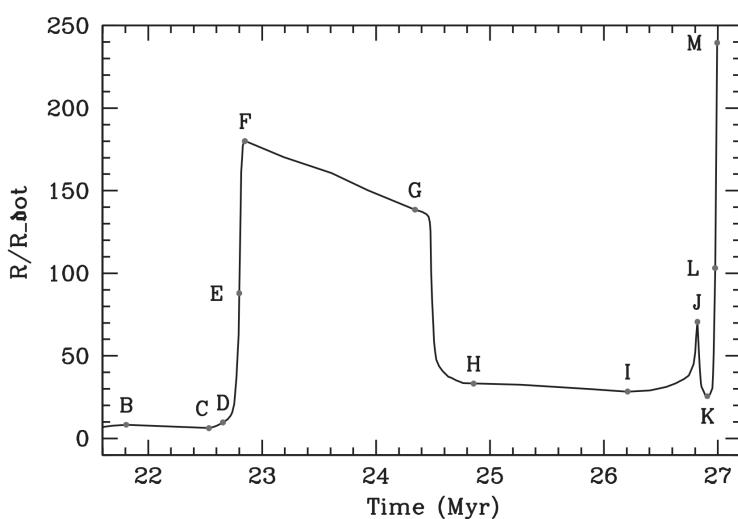


Figure 3: The stellar radius as a function of time for the same track shown in Figure 1. Labels of relevant points along the sequence are the same as in Figure 1. Source: Renzini et al. (1992, *Astrophys. J.*, 400, 280). Figure taken from Greggio & Renzini (2011).

Thus, in this star, the core helium burning phase in thermal equilibrium is spent in two distinct locations in the HRD: part on the Hayashi line as a red giant (F-G) and part as a blue giant (H-I) separated by a runaway contraction out of TE (G-H). This first blue loop (also called the **Cepheid blue loop**)

<sup>1</sup>Temperature gradients are defined as  $\nabla T(\rho, T, X_i) = (\partial \log T / \partial \log P)$ .

continues after point I when TE is broken again. Here the envelope is slowly expanding in response to the slowly increasing luminosity from the core and shell burning, when the envelope turns unstable again due to the same physical process already described for the DE phase.

However, the journey towards the Hayashi line is suddenly interrupted and inverted at point J, which marks the start of the so-called **second blue loop**. What happens at point J and after is a rather complex series of interconnected events: helium is exhausted in the core, the core contracts and helium burning rapidly shifts from the helium-exhausted CO core to a shell surrounding it; helium shell ignition is quite violent and causes expansion of the helium buffer above this shell: when the expansion front breaks through the hydrogen burning shell temperature and density in the shell drop; burning in the hydrogen shell (which was still providing most of the stellar luminosity) is effectively shut off completely causing a sudden drop of the luminosity impinging on the base of the envelope:  $L_B$  drops below  $L_S$  and the envelope stops expanding and contracts. In the meantime the strength of the helium burning shell steadily increases until it leads  $L_B$  to exceed  $L_S$  again, contraction stops and the star resumes at point K its runaway expansion that was temporarily stopped at J. The rest, from K to L to M is quite similar to the DEF phase, with convection replacing radiative transfer in the envelope and TE being rapidly restored, shortly before point M. Figure 3 clearly illustrates the dramatic effects of the envelope thermal instabilities on the overall structure of the star. Those phases which are in TE are nearly flat in this plot, that is, the stellar radius changes quite slowly with time. With one exception, those phases which are out of TE are instead nearly vertical, that is, the radius varies very rapidly during such runaway inflations or deflations. The exception is phase BC, which is only modestly out of TE, and contraction is relatively slow.

What happens past point M is still an open issue. A  $9 M_\odot$  star lies between two domains: massive stars that eventually undergo core collapse and supernova explosion, and intermediate-mass stars, which shedding all their hydrogen-rich envelope die as a white dwarf. One believes that in a  $9 M_\odot$  star carbon is ignited in the central core under only mildly degenerate conditions, and the star keeps ascending along the Hayashi line as a super asymptotic giant branch star experiencing a few thermal pulses in its deeper regions, where hydrogen and helium are still burning in two separate shells. Thus, if the envelope is completely lost in a wind the star leaves an ONeMg white dwarf. If instead mass loss is less severe then the core keeps growing in mass thanks to the active burning shells until electron captures in the core trigger a core collapse and we have a supernova explosion. Clearly the fate critically depends on the strength of the mass loss process.

**Evolution of a Solar-composition star:** Figure 1.4 shows stellar evolutionary tracks of solar composition, covering a wide range of initial masses, from a  $0.8 M_\odot$  star, whose MS evolutionary lifetime is longer than one Hubble time, up to a  $20 M_\odot$  model. The evolutionary tracks of stars with mass greater than  $20 M_\odot$  are heavily affected by mass loss.

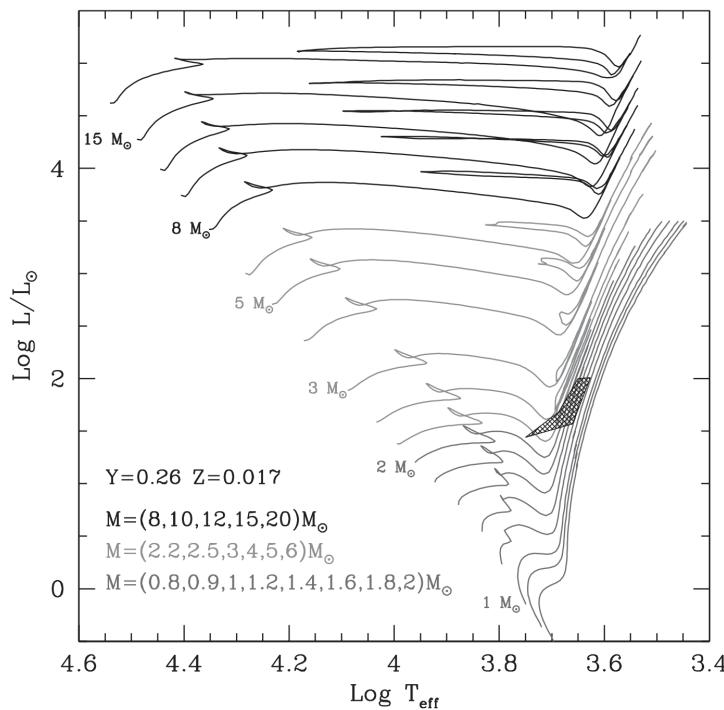


Figure 4: Evolutionary tracks of solar composition star. The shaded area shows the location of low-mass ( $0.55 \leq M/M_\odot \leq 2$ ) core helium burning models. Drawn using the YZVAR database (Bertelli, G. et al. 2008, Astron. Astrophys., 484, 815; 2009, Astron. Astrophys., 508, 355). Figure taken from Greggio & Renzini (2011).

Different gray shades in Figure 4 pertain to the different mass ranges as customarily distinguished in

stellar evolution: low-mass stars, up to  $\lesssim 2.2 M_{\odot}$ , intermediate-mass stars, up to  $\lesssim 8 M_{\odot}$ , and high-mass stars. These ranges correspond to different physical behavior during the evolution of the stars, and the mass limits depend on chemical composition. Low-mass stars develop an **electron degenerate helium core** after their MS evolution; intermediate mass stars ignite helium under non-degenerate conditions, but develop a **CO degenerate core** after central helium burning; massive stars experience all successive nuclear burnings up to the production of an **iron core**.

During their MS evolution, stars with mass below  $\sim 1 M_{\odot}$  burn hydrogen through the **p-p chain**, whose reaction rate is not extremely sensitive to temperature, so that these stars possess a radiative core. As evolution proceeds hydrogen is progressively depleted in the inner core, more in the center than in the periphery, leading to a smooth hydrogen profile from a central minimum up to the initial abundance in the outer layers. On the HRD, the stellar track climbs to higher luminosities and temperatures until the central hydrogen is severely depleted; at this point the effective temperature starts decreasing, producing the turnoff (TO), that is, the maximum temperature point on the MS, which is easily recognizable on the CMD of globular clusters. Shortly after the TO, hydrogen is completely exhausted in the center and the hydrogen shell burning phase starts as a natural progression from the previous core hydrogen burning phase. During shell hydrogen burning, the track forms the subgiant branch (SGB), evolving at (almost) constant luminosity and decreasing temperature until the external convection penetrates deeply inside, and the star becomes almost fully convective at the base of the red giant branch (RGB).

The MS stars (with  $M \gtrsim 1 M_{\odot}$ ) have a convective core, since at least part of the hydrogen burning occurs through the CNO cycle, whose energy generation rate is extremely sensitive to temperature. Because of central mixing, the hydrogen profile is characterized by an inner plateau; as evolution proceeds, the extension of the convective core progressively decreases, leaving behind a gradient of hydrogen abundance. In stars with a convective core, the fuel depletion affects a sizable central region, which starts rapid contraction when approaching fuel exhaustion. This happens at the local minimum effective temperature point during the MS evolution of stars with  $M > 1 M_{\odot}$  (Figure 4), which signals the beginning of the overall contraction phase. The evolutionary behavior across this phase and the following runaway expansion has been already described in detail in the previous section.

As already mentioned, the helium core of low-mass stars is (electron) degenerate: this implies that central helium ignition is delayed because core contraction does not lead to an efficient increase of the central temperature. The (almost) fully convective star climbs along the RGB, while the hydrogen burning shell progresses outward, thereby increasing the mass of the helium core. When the core reaches a critical limit, a helium flash occurs off-center, since neutrino losses induce a temperature inversion in the innermost layers. This event is not catastrophic for the star because local expansion removes the degeneracy; instead, a sequence of flashes occurring at progressively inner locations totally remove the core degeneracy. During this phase, which lasts 1 Myr, the star moves downward along the Hayashi line to settle on the core helium burning locus, either the red clump, or the horizontal branch (HB). The maximum luminosity reached on the RGB (RGB Tip, or TRGB) is a very important feature in the color-magnitude diagram (CMD) of stellar populations because it is virtually independent of mass, and then of evolutionary lifetime. This is because on the one hand the critical mass for helium ignition under degenerate conditions is almost constant ( $0.5 M_{\odot}$ ), and on the other, along the RGB there exists a core mass-luminosity relation. The evolutionary lifetimes of low-mass stars range from 1 Gyr up to the Hubble time, as seen in Figure 4, and all of them experience the helium flash at almost the same luminosity; therefore, the CMD of a stellar population with stars older than 1 Gyr will show a prominent TRGB feature whose luminosity is known, thus allowing us to determine the distance of the stellar population. The TRGB luminosity depends slightly on metallicity, being higher for higher metal content. However, by a fortunate combination, the absolute magnitude in the I-band of the TRGB does not depend much on metallicity, for metal-poor populations. This is because the effective temperature at the TRGB also depends on metallicity: the higher  $Z$ , the cooler the TRGB stars, and the higher the bolometric correction (in absolute value). The trend with metallicity of the bolometric correction to the I-band largely compensates that of the tip luminosity, so that the I-band absolute magnitude of the TRGB ( $M_{I,\text{TRGB}}$ ) is almost independent of age and weakly dependent on metallicity of the parent stellar population. This makes  $M_{I,\text{TRGB}}$  a very effective distance indicator in galaxies.

The luminosity of core helium burning low-mass stars, being fixed by the mass of their hydrogen-exhausted core, is also largely independent of their total mass, thereby providing another distance indicator. However, evolution during the core helium burning phase spreads the red clump stars over 0.6 magnitudes, as can be seen from the vertical size of the hatched region in Figure 4. In addition, the core mass at helium ignition is not a monotonic function of the total mass, and as the latter increases beyond the limit of the low-mass stars regime ( $M_{\text{Hef}}$ ), it decreases, reaches a minimum of about  $0.326 M_{\odot}$ , and then increases. As a result, stars with mass just above  $M_{\text{Hef}}$  start their core helium burning evolution at fainter luminosities, have the longest helium burning lifetimes, and cover a wider range of luminosities, compared to stars with  $M < M_{\text{Hef}}$ . Therefore, the core helium burners of a composite stellar population

$Z, Y$	$a$	H burning lifetime			Lifetime up to first pulse or C ignition				
		$b$	$c$	$\left\langle \frac{\Delta t_{\text{TO}}}{t_{\text{TO}}} \right\rangle$	$a$	$b$	$c$	$\left\langle \frac{\Delta t_{\text{tot}}}{t_{\text{tot}}} \right\rangle$	
0.0001, 0.23	0.8751	-3.240	9.754	0.041	0.8130	-3.194	9.809	0.022	
0.0001, 0.40	0.8248	-3.007	9.288	0.049	0.7784	-2.992	9.371	0.023	
0.001, 0.26	0.8192	-3.168	9.705	0.040	0.7679	-3.147	9.776	0.031	
0.001, 0.40	0.7590	-2.951	9.310	0.049	0.7363	-2.976	9.409	0.032	
0.004, 0.26	0.8085	-3.230	9.796	0.040	0.7534	-3.211	9.875	0.042	
0.017, 0.26	0.8071	-3.423	10.029	0.038	0.7260	-3.369	10.104	0.053	
0.04, 0.26	0.7466	-3.516	10.213	0.049	0.6694	-3.457	10.281	0.063	
0.04, 0.40	0.7649	-3.361	9.767	0.073	0.7324	-3.371	9.866	0.063	
0.04, 0.46	0.7362	-3.243	9.554	0.082	0.7427	-3.303	9.669	0.068	
0.07, 0.40	0.7322	-3.420	9.855	0.079	0.7065	-3.437	9.948	0.066	

Table 1: Coefficients of  $\log t$  for various chemical compositions, resulting from the least square fit of the hydrogen burning and the total lifetimes as a function of  $M_0$ . The fit covers the range  $0.6 \leq M_0/M_\odot \leq 20$ ; columns 5 and 9 report the average relative accuracy on the evolutionary lifetimes at the TO and at the first thermal pulse or central carbon ignition. The YZVAR database (Bertelli, G. et al. 2008, Astron. Astrophys., 484, 815; 2009, Astron. Astrophys., 508, 355) has been used to derive the coefficients. Table taken from Greggio & Renzini (2011).

with a sizable component at ages just below 1 Gyr are spread over a wide magnitude range, which limits the use of this feature for distance determinations.

The effective temperature of low mass core helium burning stars depends on the mass of their envelope, a dependence which is very pronounced when the envelope is thinner than for example  $0.2 M_\odot$  for  $Z \lesssim 0.1 Z_\odot$ . Below this threshold, the lower the envelope mass, the hotter the core helium burning stars. In a coeval and homogeneous stellar population the core helium burning stars all have virtually the same core mass and then luminosity; a spread of their envelope masses produces a feature on the HRD at about constant (bolometric) magnitude extending over a temperature range. The observational counterpart of this locus is the horizontal branch: a prominent feature in the HRD of globular clusters whose age is old enough to host core helium burning stars with low-mass envelopes. The existence of wide HBs in globular clusters was explained as due to a dispersion of mass lost during the RGB phase, but recently it became apparent that other effects are also at play. Evolutionary tracks during the core helium burning phase exhibit a wide variety of morphologies, depending on their mass, envelope mass, metallicity and helium abundance. At central helium exhaustion, a rapid core contraction leads to shell helium ignition; the model star expands and moves again towards the Hayashi line to start the asymptotic giant branch (AGB) evolutionary stage. However, if the envelope mass is small enough, the shell helium burning phase is entirely spent at high effective temperatures, as an AGB manqué star. The hot HB stars and their AGB manqué progeny are likely responsible for the UV emission from elliptical galaxies which host old stellar populations.

The evolution of intermediate and high-mass stars during the hydrogen and helium burning stages is very similar to what has already been described for the  $9 M_\odot$  star. But, following helium exhaustion, intermediate-mass stars behave similarly to low-mass stars. We only notice here that for stars with mass in the vicinity of  $M_{\text{Hef}}$  the core helium burning phase is spent in a red clump, which becomes a wider and wider blue loop as the model mass grows. Thus, the core helium burning phase of intermediate-mass stars is spent part in the blue and part in the red. The very occurrence of the loop, its extension and the fraction of lifetime spent on each side of the loop, are sensitive to a number of parameters describing the input physics of the models, like metallicity, opacity, convection and others. Therefore, the blue-to-red ratio in stellar populations with stars in this mass range is difficult to interpret. The luminosity of intermediate mass core helium burners is instead a more robust prediction of the models, and can be effectively used as an age indicator of stellar populations.

**Dependence on chemical composition:** The evolutionary tracks of given mass are very sensitive to the (initial) helium content ( $Y$ ) and metallicity ( $Z$ ), which control the energy generation rates and the opacity. We briefly illustrate some dependencies relevant to the interpretation of the HR diagram and of the spectral energy distribution of stellar populations.

At a given initial mass, evolutionary lifetimes become shorter as  $Y$  increases, because stars with higher molecular weight are more compact, hotter, brighter, hence faster in consuming the hydrogen fuel, of which less is available. Instead, lifetimes become longer as  $Z$  increases, because the hydrogen burning models get fainter due to the higher opacity, while the hydrogen fuel reservoir remains virtually unchanged. Table 1.2 lists the values of the coefficients of the following relation:

$$\log t = a \log^2 M_0 + b \log M_0 + c [\text{yr}],$$

adopted to describe the evolutionary lifetimes (in years) as a function of initial mass (in  $M_{\odot}$ ). The parabolic fit over the whole considered mass range is a rather drastic approximation; nevertheless these analytic relations can be useful to estimate evolutionary lifetimes and their dependence on composition. As mentioned previously, the values of mass defining the low, intermediate and high-mass range depend somewhat on chemical composition. For example  $M_{\text{He}}$  decreases with  $Y$  increasing and with  $Z$  decreasing. Therefore, an extended RGB on the HRD of a stellar population traces the presence of evolved stars with mass lower than  $2.1 M_{\odot}$  for solar composition, or with mass lower than  $1.5 M_{\odot}$  if  $Z = 0.001$ ,  $Y = 0.4$ . However, evolutionary lifetimes at given mass also depend on composition, and, by and large, an extended and well populated RGB is developed in stellar populations older than 1 Gyr almost irrespective of chemical composition.

At fixed initial mass, zero age main sequence models are hotter and brighter for a higher helium content, and/or a lower metallicity. Indeed, the locus of the corresponding stars on the HRD (ZAMS) is used to infer the composition of the target stellar population, and is virtually the only way to estimate the helium content of these stars, if  $Z$  is known from spectroscopy. The puzzling composition of multiple stellar populations in some galactic globular clusters has been derived from ZAMS fitting.

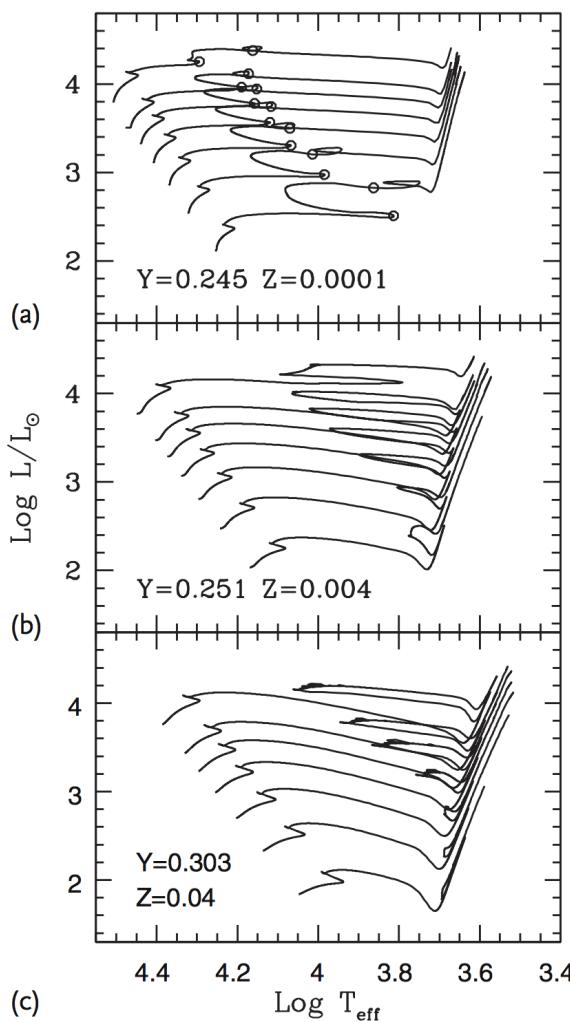


Figure 5: a-c) Evolutionary tracks of intermediate-mass stars ( $M/M_{\odot} = 3, 4, 5, 6, 7, 8, 10$ ) for different compositions, as labeled. In (a), the open circles mark start and the end of the core helium burning phase. Drawn using the BaSTI database (Pietrinferni, A. et al. 2004, *Astrophys. J.*, 612, 168). the Figure taken from Greggio & Renzini (2011).

Figure 5 shows some evolutionary tracks of intermediate-mass stars for three different initial compositions. The described trend of the ZAMS with metal content is readily visible, together with some other properties already mentioned: at very low metallicity (Figure 5a) the entire core helium burning phase occurs in the blue part of the HRD, and the thermal runaway in the envelope which brings the model star to the Hayashi track is delayed to the very latest stages. As metallicity increases, the luminosity decrease associated with the thermal runaway gets more and more pronounced: this reflects the progressively higher opacity, and then radiative energy trapping in the envelope. At the same time, at higher metallicity it is more difficult to produce extended blue loops: the  $3$  and  $4 M_{\odot}$  tracks at  $Z \sim 2 Z_{\odot}$  do not present a blue loop at all, and the loop of the  $5 M_{\odot}$  track is just alluded to. The models in Figure 5 are computed adopting classical recipes for the input physics, and even slight modifications of the assumptions lead to dramatic variations of the tracks shape. For example, intermediate-mass models with  $Z = 0.04$ ,  $Y = 0.40$  which adopt a modest overshooting from the convective core lack the loops completely, and the core helium

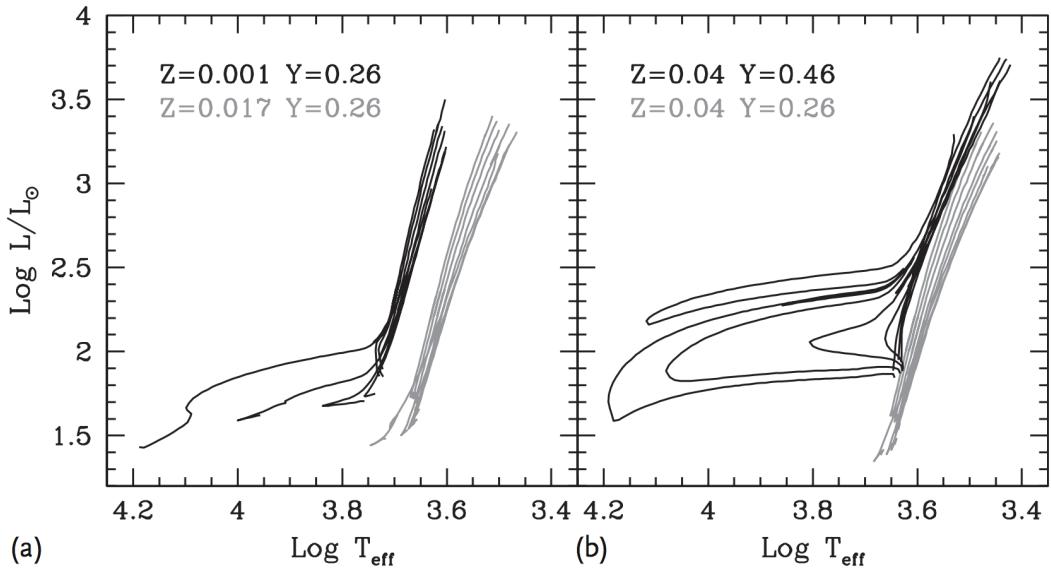


Figure 6: (a,b) Evolutionary tracks of low-mass stars during the core helium burning and early AGB phases. The chemical composition is labeled. Initial model masses are  $M/M_{\odot} = 0.55, 0.6, 0.65, 0.7, 1, 1.2, 1.4, 1.6$  except for the ( $Z = 0.04, Y = 0.46$ ) set, for which the  $1.6 M_{\odot}$  track ignites helium under non-degenerate conditions. Drawn using the YZVAR database (Bertelli, G. et al. 2008, Astron. Astrophys., 484, 815; 2009, Astron. Astrophys., 508, 355). Figure taken from Greggio & Renzini (2011).

burning phase is totally spent close to the Hayashi line.

Figure 6 illustrates the effect of chemical composition on core helium burners of low-mass. The gray tracks in Figure 6a show that at solar composition this evolutionary phase is completely spent in the red, even for masses as low as  $0.55 M_{\odot}$ . Conversely, at low metallicity (black lines), low-mass core helium burners are blue, opening the possibility of producing extended HBs in old stellar populations. The tracks in Figure 1.6b show the effect of enhancing the helium abundance: at high metallicity the core helium burning phase is completely spent close to the Hayashi line for a solar helium abundance, but if the helium abundance is high, a blueward excursion occurs during the core helium burning phase which is very wide for low-mass stars. Therefore the production of blue HB stars in old stellar populations can be achieved either assuming heavy mass loss on the RGB, or a high helium content, or both. Actually, the existence of multiple stellar populations with different helium content in NGC 2808 has been first suggested on the basis of its HB stars distribution, and confirmed later from the multiple MSs. Mass loss and helium abundance have indeed an important impact on the HRD of stellar populations, as well as on their spectral energy distribution, since stars in the core helium burning phase provide an important contribution to the total light.

- How does metallicity affect this diagram?
- Elaborate on the life sequence of low-mass stars.
- Elaborate on the life sequence of high-mass stars.
- Which direction does radius increase?
- What's the functional form for how radius increases as a function of temperature and luminosity?
- Why can you use the tip of the red giant branch for distance determination?
- How can you use the H-R diagram to determine ages? Which of the two techniques is more accurate?

## 1.2 Question 2

Sketch a plot of radius versus mass for various “cold” objects made of normal matter, including planets, brown dwarfs and white dwarfs. Explain the mass-size relationship for rocky and gaseous objects. Why is there an upper mass limit?

### 1.2.1 Short answer

Answer.

### 1.2.2 Additional context

Additional context.

### 1.2.3 Follow-up Questions

- How do you calculate the Chandrasekhar mass limit?
- Why is Saturn smaller than Jupiter? Or, why do we see a range of radii in extrasolar planets (e.g., hot Jupiters)?

### 1.3 Question 3

Describe the physical conditions that lead to the formation of absorption lines in stars' spectra. What leads to emission lines?

#### 1.3.1 Short answer

Answer.

#### 1.3.2 Additional context

If an absorber  $X$  is in a level  $\ell$  and there is radiation present with photons having an energy equal to  $E_u - E_\ell$ , where  $E_\ell$  and  $E_u$  are the energies of levels  $\ell$  (for “lower”) and  $u$  (for “upper”), the absorber can absorb a photon and undergo an upward transition:

$$\text{absorption : } X_\ell + h\nu \rightarrow X_u, \quad h\nu = E_u - E_\ell.$$

Suppose that we have number density  $n_\ell$  of absorbers  $X$  in level  $\ell$ . The rate per volume at which the absorbers absorb photons will obviously be proportional to both the density of photons of the appropriate energy and the number density  $n_\ell$ , so we can write the rate of change of  $n_\ell$  due to photo-absorption by level  $\ell$  as

$$\left( \frac{dn_u}{dt} \right)_{\ell \rightarrow u} = - \left( \frac{dn_\ell}{dt} \right)_{\ell \rightarrow u} = n_\ell B_{\ell u} u_\nu, \quad \nu = \frac{E_u - E_\ell}{h},$$

where  $u_\nu$  is the **radiation energy density per unit frequency**, and the proportionality constant  $B_{\ell u}$  is the **Einstein B coefficient** for the transition  $\ell \rightarrow u$ .

An absorber  $X$  in an excited level  $u$  can decay to a lower level  $\ell$  with emission of a photon. There are two ways this can happen:

$$\text{spontaneous emission : } X_u \rightarrow X_\ell + h\nu, \quad \nu = (E_u - E_\ell)/h,$$

$$\text{stimulated emission : } X_u + h\nu \rightarrow X_\ell + 2h\nu, \quad \nu = (E_u - E_\ell)/h.$$

**Spontaneous emission** is a random process, independent of the presence of a radiation field, with a probability per unit time  $A_{u\ell}$  – the **Einstein A coefficient**.

**Stimulated emission** occurs if photons of the identical frequency, polarization, and direction of propagation are already present, and the rate of stimulated emission is proportional to the density of these photons. Thus the total rate of depopulation of level  $u$  due to emission of photons can be written

$$\left( \frac{dn_\ell}{dt} \right)_{u \rightarrow \ell} = - \left( \frac{dn_u}{dt} \right)_{u \rightarrow \ell} = n_u (A_{u\ell} + B_{u\ell} u_\nu), \quad \nu = \frac{E_u - E_\ell}{h},$$

where the coefficient  $B_{u\ell}$  is the **Einstein B coefficient** for the downward transition  $u \rightarrow \ell$ . Thus we now have three coefficients characterizing radiative transitions between levels  $u$  and  $\ell$ :  $A_{u\ell}$ ,  $B_{u\ell}$  and  $B_{\ell u}$ . We will now see that they are not independent of one another.

In TE, the radiation field becomes the “blackbody” radiation field, with intensity given by the **blackbody spectrum**

$$B_\nu(T) = \frac{2h\nu^2}{c^2} \frac{1}{\exp(h\nu/k_B T) - 1} [\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1} \text{sr}^{-1}],$$

with specific energy density

$$(u_\nu)_{\text{LTE}} = \frac{4\pi}{c} B_\nu(T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/k_B T} - 1} [\text{erg s}^{-2} \text{cm}^{-1} \text{Hz}^{-1} \text{sr}^{-1}].$$

If we place absorbers  $X$  into a blackbody radiation field, then the net rate of change of level  $u$  is

$$\begin{aligned} \frac{dn_u}{dt} &= \left( \frac{dn_u}{dt} \right)_{\ell \rightarrow u} + \left( \frac{dn_u}{dt} \right)_{u \rightarrow \ell} \\ &= n_\ell B_{\ell u} \frac{8\pi h\nu^3}{c^3} \frac{1}{\exp(h\nu/k_B T) - 1} - n_u \left( A_{u\ell} + B_{u\ell} \frac{8\pi h\nu^3}{c^3} \frac{1}{\exp(h\nu/k_B T) - 1} \right). \end{aligned}$$

If the absorbers are allowed to come to equilibrium with the radiation field, levels  $\ell$  and  $u$  must be populated according to  $n_u/n_\ell = (g_u/g_\ell) e^{(E_\ell E_u)/k_B T}$ , with  $dn_u/dt = 0$ . From the equation above, it is easy to show that  $B_{u\ell}$  and  $B_{\ell u}$  must be related to  $A_{u\ell}$  by

$$B_{u\ell} = \frac{c^3}{8\pi h\nu^3} A_{u\ell} [\text{s}^{-1}],$$

$$B_{\ell u} = \frac{g_u}{g_\ell} B_{u\ell} = \frac{g_u}{g_\ell} \frac{c^3}{8\pi h\nu^3} A_{u\ell} [\text{s}^{-1}].$$

Thus the strength of stimulated emission ( $B_{u\ell}$ ) and absorption ( $B_{\ell u}$ ) are both determined by  $A_{u\ell}$  and the ratio  $g_u/g_\ell$ .

Rather than discussing absorption and stimulated emission in terms of the radiation energy density  $u_\nu$ , it is helpful to characterize the intensity of the radiation field by a dimensionless quantity, the photon occupation number  $n_\gamma$ :

$$n_\gamma \equiv \frac{c^2}{2h\nu^3} I_\nu \text{ [dimensionless]},$$

$$\bar{n}_\gamma \equiv \frac{c^2}{2h\nu^3} \bar{I}_\nu = \frac{c^3}{8\pi h\nu^3} u_\nu \text{ [dimensionless]},$$

where the bar denotes averaging over directions. With this definition of  $n_\gamma$ , we can rewrite equations for  $(dn_u/dt)_{\ell \rightarrow u}$  and  $(dn_\ell/dt)_{u \rightarrow \ell}$  as simply

$$\left( \frac{dn_\ell}{dt} \right)_{u \rightarrow \ell} = n_u A_{u\ell} (1 + \bar{n}_\gamma),$$

$$\left( \frac{dn_u}{dt} \right)_{\ell \rightarrow u} = n_\ell \frac{g_u}{g_\ell} A_{u\ell} \bar{n}_\gamma.$$

If the radiation field depends on frequency in the vicinity of the transition frequency  $\nu_{u\ell}$ , then  $n_\gamma$  needs to be averaged over the emission and absorption profiles in the above two equations.

From the first of the above two equations, we immediately see that the photon occupation number  $n_\gamma$  determines the relative importance of stimulated and spontaneous emission: stimulated emission is unimportant when  $\bar{n}_\gamma \ll 1$ , but should otherwise be included in analyses of level excitation.

**Absorption cross section:** Having determined the rate at which photons are absorbed by an absorber exposed to electromagnetic radiation, it is useful to recast this in terms of an absorption cross section. The photon density per unit frequency is just  $u_\nu/h\nu$ . Let  $\sigma_{\ell u}(\nu)$  be the cross section for absorption of photons of frequency  $\nu$  with resulting  $\ell \rightarrow u$  transition. The absorption rate is then

$$\left( \frac{dn_u}{dt} \right)_{\ell \rightarrow u} = n_\ell \int \sigma_{\ell u}(\nu) c \frac{u_\nu}{h\nu} d\nu \approx n_\ell u_\nu \frac{c}{h\nu} \int \sigma_{\ell u}(\nu) d\nu [\text{s}^{-1}],$$

where we have assumed that  $u_\nu$  (and  $h\nu$ ) do not vary appreciably over the line profile of  $\sigma_{\ell u}$ . Thus

$$B_{\ell u} = \frac{c}{h\nu} \int \sigma_{\ell u}(\nu) d\nu [\text{s}^{-1}],$$

and, using our equation that relates  $B_{\ell u}$  to  $A_{u\ell}$ , we obtain the integral over the absorption cross section:

$$\int \sigma_{\ell u}(\nu) d\nu = \frac{g_u}{g_\ell} \frac{c^2}{8\pi\nu_{\ell u}^2} A_{u\ell} [\text{Hz cm}^2].$$

Thus we may relate the monochromatic absorption cross section  $\sigma_{\ell u}(\nu)$  to a normalized line profile  $\phi_\nu$ :

$$\sigma_{\ell u}(\nu) = \frac{g_u}{g_\ell} \frac{c^2}{8\pi\nu_{\ell u}^2} A_{u\ell} \phi_\nu [\text{cm}^2], \quad \int \phi_\nu d\nu = 1.$$

**Intrinsic line profile:** The intrinsic line profile is characterized by a normalized profile function  $\phi_\nu^{\text{intr}}$ :

$$\sigma_\nu^{\text{intr}} = \frac{\pi e^2}{m_e c^2} f_{\ell u} \phi_\nu^{\text{intr}} [\text{cm}^2], \quad \int \phi_\nu^{\text{intr}} d\nu = 1.$$

The intrinsic line profile of an absorption line is normally described by the **Lorentz line profile** function:

$$\phi_\nu^{\text{intr}} = \frac{4\gamma_{u\ell}}{16\pi^2(\nu - \nu_{u\ell})^2 + \gamma_{u\ell}^2} [\text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1}],$$

where  $\nu_{u\ell} \equiv (E_u - E_\ell)/h$ . The Lorentz profile provides an accurate (but not exact)<sup>2</sup> approximation to the actual line profile. The Lorentz line profile has a **full width at half maximum** (FWHM) of

$$(\Delta\nu)^{\text{intr}}_{\text{FWHM}} = \frac{\gamma_{u\ell}}{2\pi} [\text{Hz}].$$

---

<sup>2</sup>The line profile is more accurately given by the **Kramers-Heisenberg formula**; Lee (2003) discusses application of this formula to the Lyman  $\alpha$  line.

The intrinsic width of the absorption line reflects the uncertainty in the energies of levels  $u$  and  $\ell$  due to the finite lifetimes of these levels<sup>3</sup> against transitions to all other levels, including both radiative and collisional transitions. If the primary process for depopulating levels  $u$  and  $\ell$  is spontaneous decay (as is often the case in the ISM), then

$$\gamma_{u\ell} \equiv \gamma_{\ell u} = \sum_{E_j < E_u} A_{uj} + \sum_{E_j < E_\ell} A_{\ell j}.$$

In the case of a **resonance line**, where  $\ell$  is the ground state, the second sum vanishes.

**Doppler broadening; The voigt line profile:** Atoms and ions are generally in motion, and the velocity distribution is often approximated by a Gaussian, this being of course the correct form if the velocities are entirely due to thermal motions:

$$p_v = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_v} e^{-(v-v_0)^2/2\sigma_v^2} = \frac{1}{\sqrt{\pi}} \frac{1}{b} e^{-(v-v_0)^2/b^2},$$

where  $p_v dv$  is the probability of the velocity along the line of sight being in the interval  $[v, v + dv]$ ,  $\sigma_v$  is the one-dimensional velocity dispersion, and  $b$  is the the **broadening parameter**,  $b \equiv \sqrt{2}\sigma_v$ .

The width of the velocity distribution is also sometimes specified in terms of the FWHM; for a Gaussian distribution of velocities, this is just

$$(\Delta v)_{\text{FWHM}} = \sqrt{8 \ln 2} \sigma_v = 2\sqrt{\ln 2} b [\text{m s}^{-1}].$$

If the velocity dispersion is entirely due to thermal motion with kinetic temperature  $T = 10^4 T_4$  K, then

$$\begin{aligned} \sigma_v &= \left( \frac{k_B T}{M} \right)^{1/2} = 9.12 \left( \frac{T_4}{M/\text{amu}} \right)^{1/2} [\text{km s}^{-1}] \\ b &= \left( \frac{2k_B T}{M} \right)^{1/2} = 12.90 \left( \frac{T_4}{M/\text{amu}} \right)^{1/2} [\text{km s}^{-1}] \\ (\Delta v)_{\text{FWHM}}^{\text{therm}} &= \left[ \frac{(8 \ln 2) k_B T}{M} \right]^{1/2} = 21.47 \left( \frac{T_4}{M/\text{amu}} \right)^{1/2} [\text{km s}^{-1}]. \end{aligned}$$

The intrinsic absorption line profile  $\phi_\nu^{\text{intr}}$  must be convolved with the velocity distribution of the absorbers to obtain the line profile

$$\phi_\nu = \int p_v(v) \frac{4\gamma_{u\ell}}{16\pi^2[\nu - (1-v/c)\nu_{u\ell}]^2 + \gamma_{u\ell}^2} dv [\text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1}],$$

where  $p_v dv$  is the probability of the absorber having radial velocity in the interval  $[v, v + dv]$ . If the absorbers have a **Maxwellian** (i.e., Gaussian) one-dimensional velocity distribution  $p_v$ , then the absorption line will have a so-called **Voigt line profile**:

$$\phi_\nu^{\text{Voigt}} \equiv \frac{1}{\sqrt{2\pi}} \int \frac{e^{-v^2/2\sigma_v^2}}{\sigma_v} p_v(v) \frac{4\gamma_{u\ell}}{16\pi^2[\nu - (1-v/c)\nu_{u\ell}]^2 + \gamma_{u\ell}^2} dv [\text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1}].$$

Unfortunately, the Voigt line profile cannot be obtained analytically except for limiting cases.<sup>5</sup> However, if, as is generally the case, the one-dimensional velocity dispersion  $\sigma_v \gg (\Delta v)_{\text{FWHM}}^{\text{intr}}$ , the central core of the line profile is well-approximated by treating the intrinsic line profile as a  $\delta$ -function, so that the central core of the line has a Maxwellian profile:

$$\phi_\nu \approx \frac{1}{\pi} \frac{c}{\nu_{u\ell}} \frac{1}{b} e^{-v^2/b^2} [\text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1}], \quad b \equiv \sqrt{2}\sigma_v.$$

**Selection rules for radiative transitions:** Some energy levels are connected by strong radiative transitions; in other cases, radiative transitions between the levels may be extremely slow. The strong transitions always satisfy what are referred to as the **selection rules for electric dipole transitions**. Here, we summarize the selection rules for the strong electric dipole transitions, and we also give the selection rules for **inter-system** and **forbidden transitions** that do not satisfy the electric dipole selection rules but nevertheless are strong enough to be astrophysically important. We will use the ion NII as an example; the first nine energy levels of NII are shown in Figure 7.

**Allowed: Electric dipole transitions:** The strongest transitions are electric dipole transitions. These are transitions satisfying the following selection rules:

---

<sup>3</sup>The **Heisenberg uncertainty principle**  $\Delta E \Delta t \geq \hbar$  implies that an energy level  $u$  has a width  $\Delta E_u \approx \hbar/\tau_u$ , where  $\tau_u$  is the level lifetime.

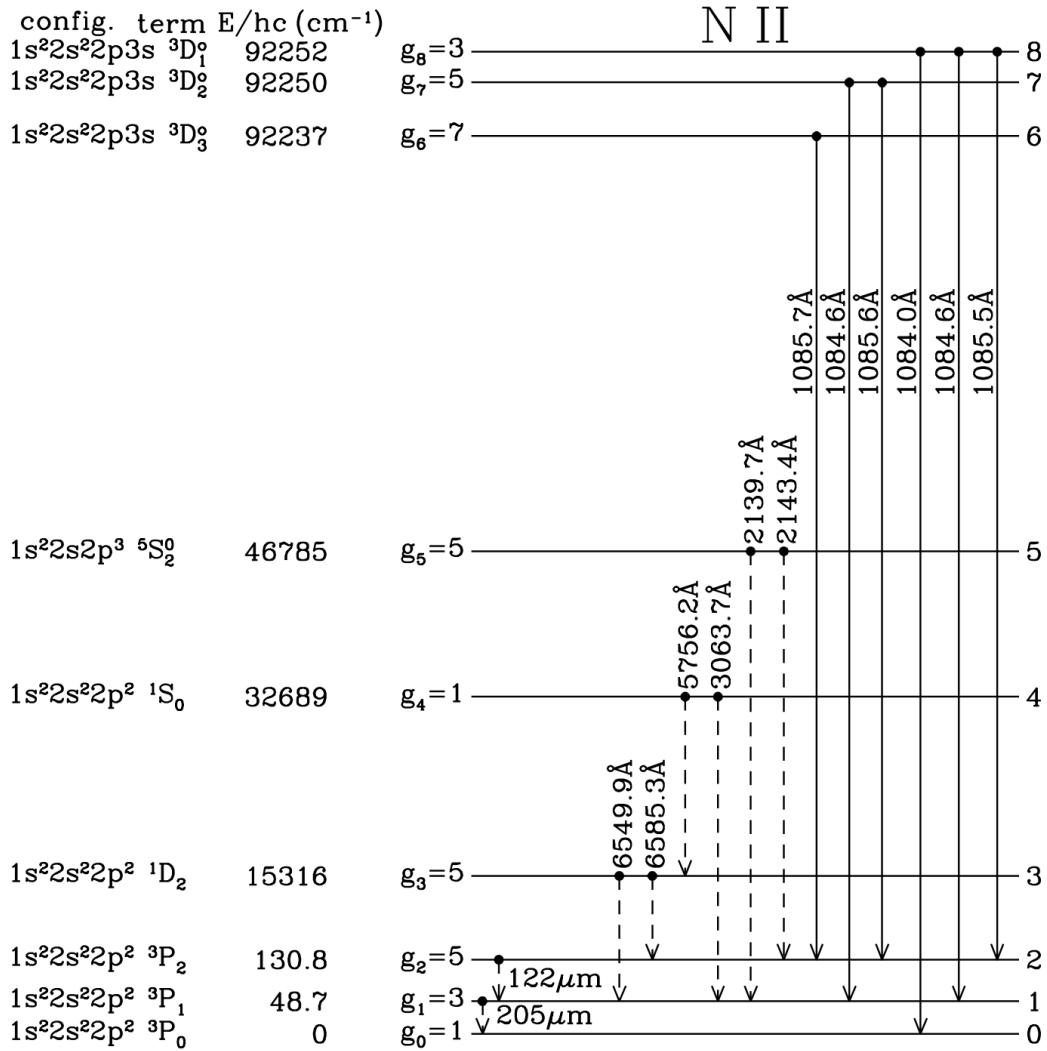


Figure 7: First nine energy levels of NII. Forbidden transitions are indicated by broken lines, and allowed transitions by solid lines; forbidden decays are not shown from levels that have permitted decay channels. Fine-structure splitting is not to scale. Hyperfine splitting is not shown. Figure taken from Draine (2015).

1. Parity must change.
2.  $\Delta L = 0, \pm 1$ .
3.  $\Delta J = 0, \pm 1$ , but  $\Delta J = 0 \rightarrow 0$  is forbidden.
4. Only one single-electron wave function  $n\ell$  changes, with  $\Delta\ell = \pm 1$ .
5.  $\Delta S = 0$ : Spin does not change.

An allowed transition is denoted *without* square brackets, for example,

$$\text{NII } 1084.0\text{\AA} \ ^3P_0 - \ ^3D_1^o.$$

This is a transition between the  $\ell = 1s^2 2s^2 2p^2 3P_0$  and  $u = 1s^2 2s^2 2p3s 3D_1^o$  levels of NII, with a wavelength  $\lambda_{ul} = 1084.0 \text{\AA}$ . The transition has  $A_{ul} = 2.18 \times 10^8 \text{ s}^{-1}$ . This decay is very fast – the lifetime of the  $3D_1^o$  level against this decay is only  $1/A_{ul} = 4.6 \text{ ns}$ !

**Spin-Forbidden or inter-system transitions:** These are transitions that fulfill the electric dipole selection rules 1 to 4 but have  $\Delta S \neq 0$ . These transitions are considerably weaker than allowed transitions. Such transitions are sometimes referred to as **semi-forbidden**, or **inter-combination**, or **inter-system transitions**; the latter is the terminology that we will use. An inter-system transition is denoted with a single right bracket – for example,

$$\text{NII}] 2143.4\text{\AA} \ ^3P_2 - \ ^5S_2^o.$$

This is a transition between the  $\ell = 1s^2 2s^2 2p^2 3P_2$  and  $u = 1s^2 2s^2 2p^3 {}^5S_2^o$  levels of NII, with a wavelength  $\lambda_{ul} = 2143.4 \text{ \AA}$  and  $A_{ul} = 1.27 \times 10^2 \text{ s}^{-1}$ .

**Forbidden transitions:** Forbidden transitions are those that fail to fulfill at least one of the selection rules 1 to 4. The transition probabilities vary widely, depending on the values of the electric quadrupole or magnetic dipole matrix elements between the upper and lower states. A forbidden transition is denoted with two square brackets – for example,

$$[\text{NII}] 6549.9 \text{\AA} \, {}^3P_1 - {}^1D_2^o.$$

This is a transition between the  $\ell = 1s^2 2s^2 2p^2 3P_1$  and  $u = 1s^2 2s^2 2p^2 {}^1D_2^o$  levels of NII, with a wavelength  $\lambda_{ul} = 6549.9 \text{ \AA}$  and  $A_{ul} = 9.20 \times 10^{-4} \text{ s}^{-1}$ . This fails rule 1 (parity is unchanged) and it fails rule 4 (single electron wave functions are unchanged). This is an example of a **magnetic dipole transition**.

Another example of a forbidden transition is the **electric quadrupole transition**

$$[\text{NII}] 5756.2 \text{\AA} \, {}^1D_2 - {}^1S_0^o.$$

This is a transition between the  $\ell = 1s^2 2s^2 2p^2 1D_2$  and  $u = 1s^2 2s^2 2p^2 {}^1S_0^o$  levels of NII, with a wavelength  $\lambda_{ul} = 5756.2 \text{ \AA}$  and  $A_{ul} = 1.17 \text{ s}^{-1}$ . This fails rules 1 (parity is unchanged) and 4 (single electron wave functions are unchanged) and it fails rules 2 and 3 ( $\Delta L = -2$  and  $\Delta J = -2$ ), yet its transition probability is three orders of magnitude larger than the magnetic dipole transition [NII] 6549.9\AA!

We see then that there is a hierarchy in the transition probabilities: very roughly speaking, inter-system lines are  $\sim 10^6$  times weaker than permitted transitions, and forbidden lines are  $\sim 10^2 - 10^6$  times weaker than inter-system transitions.

Despite being very “weak,” forbidden transitions are important in astrophysics for the simple reason that every atom and ion has excited states that can only decay via forbidden transitions. At high densities, such excited states would be depopulated by collisions, but at the very low densities of interstellar space, collisions are sufficiently infrequent that there is time for forbidden radiative transitions to take place.

### 1.3.3 Follow-up Questions

- Why aren’t emission and absorption lines delta functions?
- How does this relate to population levels and excitation temperatures?
- Are there emission lines in the Sun? Why is there emission from the Calcium doublet?
- Write down the heat transfer equation. What do solutions look like?

## 1.4 Question 4

Describe these important sources of stellar opacity: electron scattering, free-free, bound-free, and the H<sup>-</sup> ion.

### 1.4.1 Short answer

Answer.

### 1.4.2 Additional context

**Electron scattering:** If an EM wave passes an electron, the electric field makes the electron oscillate. The oscillating electron represents a classical dipole that radiates in other directions, (i.e., the electron scatters part of the energy of the incoming wave). The weakening of the original radiation due to scattering is equivalent to that by absorption, and we can describe it by way of a cross section at frequency  $\nu$  per unit mass ( $\kappa_\nu$ ). This can be calculated classically giving the result

$$\kappa_\nu = \frac{8\pi}{3} \frac{e_e^2}{\mu_e m_u} = 0.20(1+X) \text{ [cm}^2 \text{ g}^{-1}\text{]},$$

where  $r_e$  is the classical electron radius,  $X$  the mass fraction of hydrogen, and the constant is in  $\text{cm}^2 \text{ g}^{-1}$ . The term  $\mu_e m_u$  arises because  $\kappa_\nu$  is taken per unit mass; and  $\mu_e$  is replaced by (4.30). Since  $\kappa_\nu$  does not depend on the frequency, we immediately obtain the **Rosseland mean for electron scattering**:

$$\kappa_{\text{sc}} = 0.20(1+X) \text{ [cm}^2 \text{ g}^{-1}\text{]}.$$

The Thomson scattering just described neglects the exchange of momentum between electron and radiation. If this becomes important, then  $\kappa_\nu$  will be reduced compared to the value given above, though this effect plays a role only at temperatures sufficiently high for the scattered photons to be very energetic. In fact, during the scattering process the electron must obtain such a large momentum that its velocity is comparable to  $c$ , say  $v \gtrsim 0.1c$  for the equation above to become a bad approximation. The momentum of the photon is  $h\nu/c$ , which after scattering is partly transferred to the electron,  $m_e v \sim h\nu/c$ . Therefore relativistic corrections (**Compton scattering**) become important if the average energy of the photons is  $h\nu \gtrsim 0.1 m_e c^2$ . For  $h\nu$  we take the frequency at which the Planck function has a maximum; then according to Wien's law this is at  $h\nu = 4.965 k_B T$ , and the full Compton scattering cross section has to be taken into account if  $T > 0.1 m_e c^2/(4.965 \text{ K})$ , or roughly  $T > 10^8 \text{ K}$ . In fact even at  $T = 10^8 \text{ K}$  Compton scattering reduces the opacity by only 20% of that given by the equation above.

**Free-free:** If during its thermal motion a free electron passes an ion, the two charged particles form a system which can absorb and emit radiation. This mechanism is only effective as long as the electron and ion are sufficiently close. Now, the mean thermal velocity of the electrons is  $v \sim \sqrt{T}$ , and the time during which they form a system able to absorb or emit is proportional to  $1/v \sim 1/\sqrt{T}$ ; therefore, if in a mass element the numbers of electrons and ions are fixed, the number of systems temporarily able to absorb is proportional to  $1/\sqrt{T}$ .

The absorption properties of such a system have been derived classically; the absorption coefficient per system is proportional to  $Z^2 \nu^{-3}$ , where  $Z$  is the charge number of the ion. We therefore expect the absorption coefficient  $\kappa_\nu$  of a given mixture of (fully ionized) matter to be

$$\kappa_\nu \sim Z^2 \rho T^{-1/2} \nu^{-3} \text{ [cm}^2 \text{ g}^{-1}\text{]}.$$

Here the factor  $\rho$  appears because for a given mass element the probability that two particles are accidentally close together is proportional to the density.

For the determination of the Rosseland mean  $\kappa$  of this absorption coefficient we make use of a simple theorem: a factor  $\nu^\alpha$  contained in  $\kappa_\nu$  gives a factor  $T^\alpha$  in  $\kappa$ . With this and the above equation for  $\kappa_\nu$  we find

$$\kappa_{\text{ff}} \sim \rho T^{-7/2} \text{ [cm}^2 \text{ g}^{-1}\text{]}.$$

All opacities of this form are called **Kramers opacities** and give only a classical approximation. One normally multiplies the Kramers formula by a correction factor  $g$ , the so-called **Gaunt factor**, in order to take care of the quantum-mechanical correction. In the equation for  $\kappa_{\text{ff}}$ , we have still omitted the factor  $Z^2$  which appears in  $\kappa_\nu$ . In general, one has a mixture of different ions, and therefore one has to add the contributions of the different chemical species. The (weighted) sum over the values of  $Z^2$  is taken into the constant of proportionality in  $\kappa_{\text{ff}}$ , which then depends on the chemical composition. For a fully ionized mixture a good approximation is given by

$$\kappa_{ff} = 3.8 \times 10^{22}(1 + X)[(X + Y) + B]\rho T^{-7/2} [\text{cm}^2 \text{g}^{-1}],$$

with the numerical constant in cgs. The mass fractions of H and He are  $X$  and  $Y$ , respectively. Here the factor  $(1 + X)$  arises, since  $\nu_{ff}$  must be proportional to the electron density which is proportional to  $(1 + X)\rho$ . The term  $(X + Y)$  in the brackets can be understood in the following way: there are  $X/m_u$  hydrogen ions and  $Y/(4m_u)$  helium ions. The former have the charge number 1, the latter the charge number 2. But since  $\kappa_\nu \sim Z^2$ , when adding the contributions of H and He to the total absorption coefficient, we obtain the factor  $X/m_u + 4Y/(4m_u) = (XY)m_u$ . Correspondingly the term  $B$  gives the contribution of the heavier elements:

$$B = \sum_i \frac{X_i Z_i^2}{A_i} \text{ [dimensionless]},$$

where the summation extends over all elements higher than helium and  $A_i$  is the atomic mass number.

**Bound-free:** We first consider a (neutral) hydrogen atom in its ground state, with an ionization energy of  $\chi_0$  (i.e., a photon of energy  $h\nu > X_0$  can ionize the atom). Energy conservation then demands that

$$h\nu = \chi_0 + \frac{1}{2}m_e v_e^2 \text{ [eV]},$$

where  $v_e$  is the velocity of the electron released (relative to the ion, which is assumed to be at rest before and after ionization).

If we define an absorption coefficient  $a_\nu$  per ion ( $a_\nu = \kappa_\nu \rho / n_{\text{ion}}$ ), we expect  $a_\nu = 0$  for  $\nu < \chi_0/h$  and  $a_\nu > 0$  for  $\nu \geq \chi_0/h$ . Classical considerations similar to those which lead to the Kramers dependence of  $\kappa_\nu$  for freefree transitions give a  $a_\nu \sim \nu^{-3}$  for  $\nu \geq \chi_0/h$ . Quantum-mechanical corrections can again be taken into account by a Gaunt factor. But if we have neutral hydrogen atoms in different stages of excitation, the situation is different: an atom in the first excited stage has an absorption coefficient  $a_\nu = 0$  for  $h\nu < \chi_1$ , where  $\chi_1$  is the energy necessary to ionize a hydrogen atom from the first excited state, while  $a_\nu \sim \nu^{-3}$  for  $h\nu \geq \chi_1$ . The absorption coefficient  $\kappa_\nu$  for a mixture of hydrogen atoms in different states of excitation is a superposition of the  $a_\nu$  for different stages of excitation. The resulting  $\kappa_\nu$  is a saw-tooth function. In order to obtain  $\kappa_\nu$  for a certain value of the temperature  $T$ , one has to determine the relative numbers of atoms in the different stages of excitation by the Boltzmann formula; then their absorption coefficients  $a_\nu$ , weighted with their relative abundances, are to be summed.

If there are ions of different chemical species with different degrees of ionization, one has to sum the functions  $a_\nu$  for all species in all stages of excitation and all degrees of ionization before carrying out the Rosseland integration. An important source of opacity are bound-free transitions of neutral hydrogen atoms, in which case the opacity must be proportional to the number of neutral hydrogen atoms and  $\kappa$  can be written in the form

$$\kappa_{bf} = X(1 - x)\tilde{\kappa}(T) [\text{cm}^2 \text{g}^{-1}].$$

Here  $\tilde{\kappa}(T)$  is obtained by Rosseland integration over (weighted) sums of functions  $a_\nu$  for the different stages of excitation, while  $x$  is the degree of ionization.

**H<sup>-</sup> ion:** Hydrogen atom has a bound state for a second electron in the field of the proton, though it has a very low ionization potential,  $E_{H^-} = 0.75 \text{ eV}$ . The number density of negative hydrogen ions will be proportional to the electron density, which, in all but the most metal-poor stars, will be set by ionization of the metals (which have much lower ionization potentials than hydrogen and helium). Thus, the H<sup>-</sup> opacity will scale as  $\kappa_{H^-} \propto \rho X Z$  at low temperatures; H<sup>-</sup> is of course easily ionized at higher temperatures, and at very low temperatures even metals will not be ionized, so there will be no electrons to form H<sup>-</sup> by combining with H.

## 1.5 Question 5

Describe the processes that can cause pulsations in a star's luminosity, and provide at least one example of a class of stellar pulsation.

### 1.5.1 Short answer

Classifications of stellar pulsations include:

Some classifications of stellar pulsations include:

- **Cepheids:**
- **RR Lyrae variables:**
- **Mira variables:** The Mira variable stars are defined as M-type stars with periods of 80–1000 days and amplitudes greater than 2.5 magnitudes. The spectra of these giant pulsating stars show dramatic changes during the pulsation cycle, and even show spectral differences from cycle to cycle. Shock waves in the pulsating atmospheres of these stars produce emission lines in their spectra. The hydrogen lines are usually in emission, with the emission strength increasing toward maximum light, although the Balmer decrement is quite often unusual.

For a star to begin and continue to pulsate, there must be some mechanism which converts radiant energy into energy of motion in a timely way, overcoming friction and building up the motion. Such a mechanism can exist in a star: it is the thermodynamic effect of an ionization zone in a star. In a star of moderate surface temperature, the layers below the surface have temperatures up to a few tens of thousands of degrees. The helium tends to be singly-ionized; it has lost one of its two electrons but not the second. If the star contracts, it heats slightly which causes the helium to lose its second electron. This requires energy which comes from the radiation flowing outward through the gas; it is absorbed by the opacity of the gas. The energy then becomes stored in the form of ionization energy. When the star subsequently expands and cools, the electrons recombine with the helium atoms, and release the stored-up ionization energy. This is the ‘push’ which maintains the pulsation in all of the pulsating stars in the Cepheid instability strip.

### 1.5.2 Additional context

Variable stars are those that change brightness. Their variability may be due to geometric processes such as rotation, or eclipse by a companion star, or physical processes such as vibration, flares, or cataclysmic explosions. In each case, variable stars provide unique information about the properties of stars, and the processes that go on within them.

Variable stars are stars which change in brightness. The change may be as small as a few parts in a million, or it may be a factor of a thousand or more. It may occur in a second or less, or it may take years, decades, or centuries. These are extremes, but astronomers have developed an array of techniques for discovering, measuring, and analyzing the full range of possible variable stars. Why? Because the variations provide important and often unique information about the nature and evolution of the stars. This information can be used to deduce even more fundamental knowledge about our universe in general. The variations may be due to the rotation of a spotted star, or to an eclipse of a star by a companion star, or even by an unseen planet. The variations may be due to the vibrations of a star; if they are complex enough (as they are in our Sun), they may provide an internal ‘picture’ of the star. The variations may be due to eruptions on a star (flares), an accretion disc (dwarf novae), major explosions on a star (novae), or to the total disruption of a star in a supernova. SNe are the most violent events in our Universe, yet we owe our existence to them, because they help to recycle the atoms, created in stars, into space where some of them became part of our Sun, our planet, and our biosphere. Also, the elements heavier than iron were mostly created in SNe explosions. SNe may be dramatic and extreme, but they represent only one of the many roles that variable stars play in modern astrophysics, and in our understanding of the Universe and the processes which govern it.

**Classification of variable stars:** The earliest classification scheme dates back nearly two centuries: Pigott divided variables according to the nature of their light curve into novae, long-period variables, and short-period variables. A century later, Pickering devised a more detailed scheme: (Ia) normal novae, primarily nearby ones in our own galaxy; (Ib) novae in nebulae: now known to be primarily SNe in distant galaxies; (IIa) long-period variables, cool, large-amplitude pulsating variables; (IIb) U Geminorum stars, dwarf novae; (IIc) R Coronae Borealis stars, stars which suddenly and unpredictably decline in brightness; (III) irregular variables, a motley collection; (IVa) short-period variables, such as Cepheids

and later including the cluster-type or RR Lyrae stars; (IVb) Beta Lyrae type eclipsing variables; and (V) Algol type eclipsing variables.

Pickering's classification scheme contains some hints as to the nature and cause of the variability. Classification and explanation go hand in hand (in principle), and in the late nineteenth and early twentieth centuries, progress was made in understanding the physical nature and the physical processes in variable stars, and in stars in general.

**Pulsating variable stars:** Pulsation is the astronomer's word for *vibration or oscillation*. Every physical object has natural patterns or modes of vibration, each with a corresponding period – the time required for one vibration.

**Pulsation modes:** Stars are, to a first approximation, spherical. The simplest form of pulsation is radial pulsation – a simple, spherically symmetric in-and-out expansion and contraction. A star has an infinite number of modes of radial pulsation. The simplest is called the **fundamental mode**. In this mode, all parts of the star expand together and contract together, in unison. The next simplest mode is the **first overtone**. In this mode, there is a nodal sphere in the star, where the material remains at rest. When the part of the star outside this sphere is expanding, the part inside is contracting, and vice versa. In the **second overtone mode**, there are two nodal spheres, where the material remains at rest. The large-amplitude pulsating stars – **Cepheids**, **RR Lyrae variables**, and **Mira variables** – pulsate primarily in radial modes. Radial pulsation produces substantial changes in luminosity (and therefore brightness), temperature (and therefore colour), and radial velocity; these are observed, and provide direct evidence for the radial pulsation. Through techniques such as the **Baade-Wesselink method**, they can be used to determine the average radius of the star.

It is less easy to tell which radial mode a star is pulsating in. Most, but not all, large-amplitude variables pulsate in the fundamental mode. If the radius and mass of the star can be estimated, the observed period can be compared with the theoretical or expected period for each mode. If the star pulsates in two or more radial modes, then the period ratio can be used to deduce the modes.

In **non-radial pulsations**, the star changes shape, not volume (Figure 8). There is a triply-infinite set of modes, corresponding to the three different co-ordinate axes on the star. These are chosen to be: the distance from the centre, the angular distance ('latitude') above or below the star's equator, and the angular distance ('longitude') around the star's equator. Non-radial modes can be divided into (i) **p (pressure) modes**, in which the motion is primarily radial and the restoring force is pressure (as it is in radial modes), and (ii) **g (gravity) modes**, in which the motion is primarily horizontal, and the restoring force is buoyancy or gravity (as it is in water waves). A wide assortment of stars pulsate non-radially. Non-radial pulsation tends to produce smaller amplitudes than radial modes in the brightness and colour variation, but, if they can be observed, their relative amplitudes and phases (along with the periods themselves) can be used to identify the precise non-radial mode. Non-radial pulsation also produces characteristic absorption line profile variations. With modern high S/R spectrographs, these can and have been observed and studied over time. Also, there are often very close non-radial periods, whose differences are a clue to the mode identification. In non-rotating stars, these close periods are actually identical or degenerate, but in a rotating star, they become different; the difference is proportional to the amount of rotation. The situation is complicated by the fact that many of these non-radially pulsating stars rotate rapidly, and the theory of rapidly rotating pulsating stars is not well understood.

**Modelling stellar pulsation:** The theory of stellar pulsation begins with a model of the star which gives the physical properties of the star as a function of distance from the centre. The first models of stars were primitive (and so were the deductions from stellar pulsation theory). Now, with improved understanding of the physics of stars, and powerful computers to implement this knowledge, the models are much more detailed and accurate.

The simplest level of modelling is the **linear adiabatic theory** (LAT) of radial pulsation of a spherically symmetric, non-rotating, non-magnetic star. This was formulated by Arthur Eddington, early in the twentieth century. It assumes that the pulsations are infinitesimally small, and that there is no transfer of energy between parts of the star. It provides reasonable estimates of the periods of the radial modes and their dependence on the physical properties of the star. For instance, it shows that the period of any mode is approximately inversely proportional to the square root of the mean density of the star. Alternatively, the period times the square root of the density is approximately constant; we call it the **Q-value, or pulsation constant**. It is this relationship which leads to the period-luminosity relation; since both period and luminosity are related to the radius of the star, and since the temperatures of pulsating stars are about the same, period and luminosity will be related to each other. The LAT also shows that the relative amplitude of pulsation tends to be largest in the outer layers, where the density is lowest. Corresponding calculations can be done for non-radial pulsation. The LAT assumes that the pulsation takes the form of a standing wave in the star (i.e., a wave which is perfectly reflected from the centre and the surface).

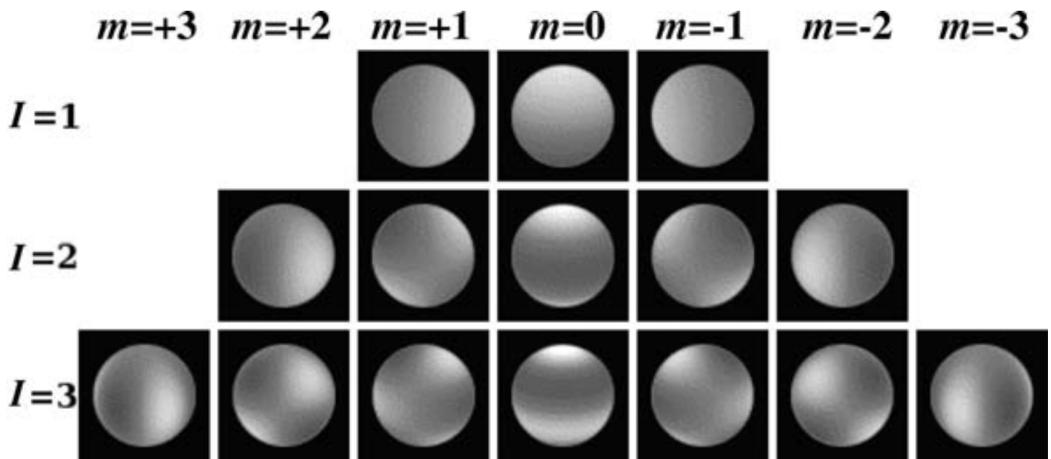


Figure 8: Illustration of non-radial pulsation modes. When the light-coloured regions are moving downward, the dark-coloured regions are moving upward. In radial pulsation, all parts of the star move in and out in unison. (Source: Whole Earth Telescope). Figure taken from Percy (2007).

The next level of complexity is the **linear non-adiabatic theory** (LNAT). This also assumes that the pulsations are infinitesimally small, but it allows for the possibility that any layer of the star may gain or lose energy. In particular, the amplitude of the infinitesimal pulsations can increase or decrease, and we can tell what modes may be unstable, and grow. We cannot tell what the eventual amplitudes may be (i.e., which mode(s) may grow to become observable). The LNAT gives periods which are slightly more realistic than the LAT, though they are about the same in most cases.

The next level of complexity is the **non-linear, non-adiabatic theory**, which amounts to a full, time-dependent, hydrodynamical model of the star. This includes the equation which relates the motion of a layer in the star to the forces acting on it. The calculation begins with the static, ‘equilibrium’ model of the star. Infinitesimal random motions are imposed on the model (such as from the round-off errors in the calculations), to see if they grow or decay. After several hundred cycles, the model (usually) reaches some constant, stable amplitude in one or more modes. The non-linear theory accurately predicts the amplitudes of Cepheids and RR Lyrae stars, which pulsate in one or two radial modes. It has not been able to explain stars like Delta Scuti stars, which can and do pulsate in many radial and non-radial modes. An ongoing problem in stellar pulsation is: what determines what mode(s) a star will pulsate in, and with what amplitude?

These models are non-rotating and non-magnetic. Non-linear analysis of rotating magnetic models is at the frontier of present research. But we know that all stars rotate (some of them very rapidly) and that many of them have significant magnetic fields. In reality, there is no such thing as a truly spherically symmetric star; for most of them, however, that is a reasonable approximation.

**Non-linear effects:** The pulsation waves in a star are not reflected perfectly from the surface, but move outward in an atmosphere of decreasing density. In many variables, such as Mira stars, this creates **shock waves** as the atmospheric layers pile up against each other. In Mira stars, this creates layers of higher density where gases condense into dust. The dust absorbs the star’s radiation and is pushed outward, carrying gas along with it. The result is mass loss which has a profound effect on the star and the space around it. In the largest, most luminous supergiants, the outward radiation pressure almost balances or perhaps even exceeds gravity; the star borders on instability; and non-linear effects may be significant.

**The instability strip(s):** Normally, stars are stable. The inward pull of gravity is balanced by the outward pressure of the hot gas in the star. If the star begins to contract, its internal pressure and temperature increase and reverse the contraction. If the star begins to expand, its internal pressure and temperature decrease and gravity restores the initial balance. Furthermore, if a star expands or contracts, the energy of motion is dissipated by friction. Eventually the star returns to rest.

For a star to begin and continue to pulsate, there must be some mechanism which converts radiant energy into energy of motion in a timely way, overcoming friction and building up the motion. Such a mechanism can exist in a star: it is the thermodynamic effect of an ionization zone in a star.

In a star of moderate surface temperature, the layers below the surface have temperatures up to a few tens of thousands of degrees. The helium tends to be singly-ionized; it has lost one of its two electrons but not the second. If the star contracts, it heats slightly which causes the helium to lose its second electron. This requires energy which comes from the radiation flowing outward through the gas; it is absorbed by the opacity of the gas. The energy then becomes stored in the form of ionization energy.

When the star subsequently expands and cools, the electrons recombine with the helium atoms, and release the stored-up ionization energy. This is the ‘push’ which maintains the pulsation in all of the pulsating stars in the Cepheid instability strip (Figure 9). If there is sufficient mass in the ionization zone to store up significant radiant energy during the pulsation cycle, then the star will pulsate. This mechanism is referred to as **self-exciting**, in contrast to the pulsations of the Sun and other Sun-like stars, which are excited by the turbulence in their convection zones.

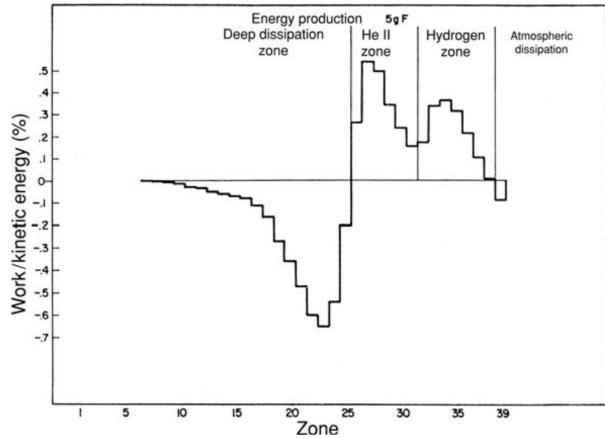


Figure 9: Diagram showing the zones in an unstable star which convert radiant energy into kinetic energy, and vice versa, in a model of a pulsating RR Lyrae star, expressed as the fractional energy production (positive) or dissipation (negative) per period. The centre of the star is at the left, the atmosphere at the right. The two energizing zones (the hydrogen zone and the ionized helium (He II) zone) are shown, along with the principal dissipation zone, deeper in the star. (From Christy, 1966.) Figure taken from Percy (2007).

But what about the other types of pulsating stars? The Beta Cephei stars were a real mystery, because they are much hotter than the Cepheid instability strip. Only in the 1980s was it realized that the ionization of iron, at temperatures near 150,000 K deep within the star had sufficient effect to maintain the pulsation. This realization came about as a result of physicists’ recalculation of the opacity of iron at these temperatures. This solved the long-standing mystery of the Beta Cephei stars, and some other nagging problems in pulsation theory as well.

Figure 10 shows the location of various types of pulsating variable stars in the HR diagram. There is a misconception often included in astronomy textbooks that pulsation is something that happens only in evolved stars such as giants and supergiants. In fact, the Cepheid instability strip extends from the top of the HR diagram to the bottom, at approximately constant temperature. That is because the instability mechanism occurs at a particular temperature – the temperature at which there is a helium ionization zone in the outer layers of the star. So supergiant Cepheids pulsate, main-sequence Delta Scuti stars pulsate, and white dwarfs pulsate – all as a result of the same mechanism.

The Cepheid instability strip is not exactly vertical: its temperature is cooler at its high-luminosity end. That is because more luminous stars have outer regions of lower density; at lower density, ionization can occur at a lower temperature because there is lots of ‘room’ for the electrons that are released.

Stars can evolve into the instability strip in different ways (figure 6.4). Massive young stars can evolve quickly from left to right in the HR diagram; old low-mass stars can evolve slowly to the same position. So there may be Population I and II stars in the same region of the instability strip.

Stars can evolve into the instability strip in different ways (Figure 11). Massive young stars can evolve quickly from left to right in the HR diagram; old low-mass stars can evolve slowly to the same position. So there may be Population I and II stars in the same region of the instability strip.

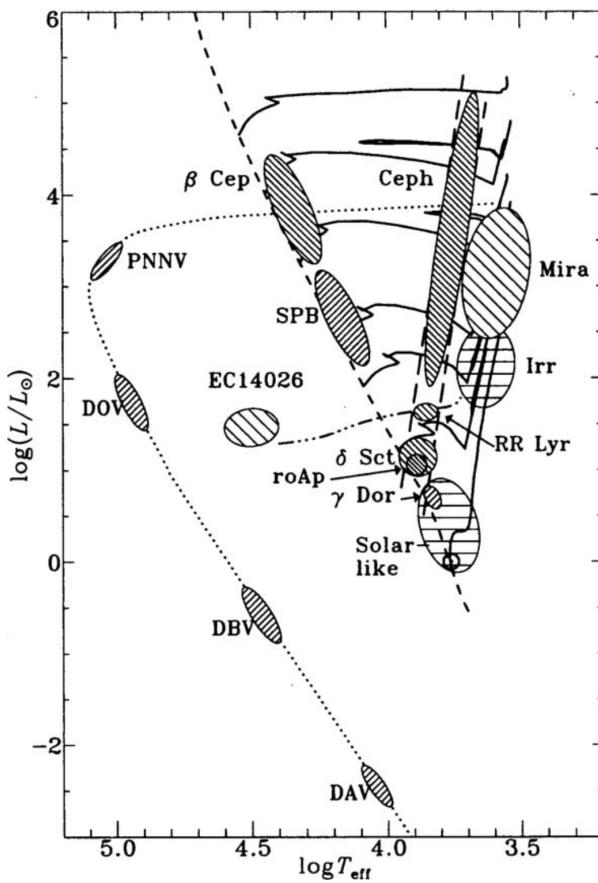


Figure 10: The location of various types of pulsating star on the HR diagram. The main instability regions are the Cepheid strip, including its intersection with the main sequence (dashed line), the region of the coolest stars (Miras and their relatives), and the region of the Beta Cephei and SPB stars. There are also instability regions along the white dwarf cooling track (dotted line) at the bottom of the diagram, including the nuclei of planetary nebulae (PNNV) and the DOV, DBV, and DAV pulsating white dwarfs. (From J. Christensen-Dalsgaard, private communication.) Figure taken from Percy (2007).

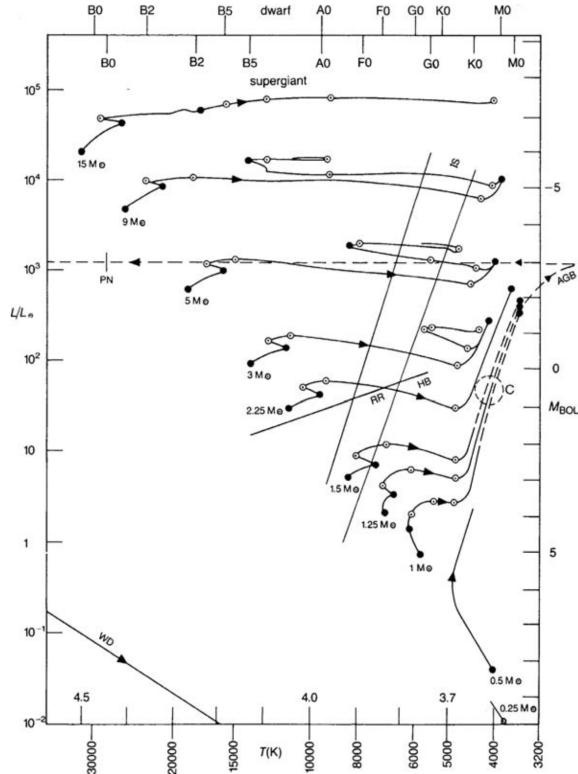


Figure 11: Evolution tracks on the HR diagram, showing the changes which take place in a star as it evolves. The tracks are marked by the mass of the star in solar units. The Cepheid pulsation instability strip is marked IS, and the horizontal branch where RR Lyrae (RR) variables are found is marked HB. The time scales for evolution depend on the mass: the more massive the star, the faster the evolution. (From Kaler, 1997.) Figure taken from Percy (2007).

### 1.5.3 Follow-up Questions

- What about the instability strip? RR Lyrae?
- What is the period-luminosity relation?
- What is the form of the period-luminosity relation?

- How would you derive the time scale of pressure waves in a star?
- How would you order-of-magnitude estimate the period for a pulsation?

## 1.6 Question 6

Briefly describe the sources of thermal energy for stars and planets.

### 1.6.1 Short answer

The sources of thermal energy of stars include:

- gravitational energy:
- rotational energy:
- nuclear energy:

The sources of thermal energy of planets include:

- radioactive decay:
- differentiation:
- solar radiation:
- accretion/impact:
- tidal heating:

### 1.6.2 Additional context

#### Stars:

One of the great mysteries of the late nineteenth and early twentieth centuries was the source of the energy required to sustain the luminosity of the sun. By then, the defining Solar parameters of mass, radius, and luminosity were known with sufficient precision to attempt to relate them. For instance, it was clear that if the Sun derived its energy from chemical processes typically yielding less than  $10^{12} \text{ erg g}^{-1}$ , it could shine no longer than about 10,000 years at its current luminosity. It is said that Lord Kelvin, in noting that the liberation of gravitational energy could only keep the sun shining for about 10 million years, found it necessary to reject Charles Darwin's theory of evolution because there would have been insufficient time for natural selection to provide the observed diversity of species.

**Gravitational energy:** It is generally conceded that the Sun has shone at roughly its present luminosity for at least the past 2 billion years and has been in existence for nearly 5 billion years. With this in mind, let us begin our study of the sources of stellar energy with an inventory of the stores of energy available to the sun. Perhaps the most obvious source of energy is that suggested by Lord Kelvin, namely gravitation:

$$E_p \leq -\frac{3}{5} \frac{GM^2}{R} [\text{J}].$$

The right-hand side of the inequality is the gravitational potential energy for a uniform density sphere, which provides a sensible upper limit for the energy. Remember that the gravitational energy is considered negative by convention; a rather larger magnitude of energy may be available for a star that is more centrally concentrated than a uniform-density sphere. We may acquire a better estimate of the gravitational potential energy by using the results for a polytrope. Chandrasekhar obtains the following result for the gravitational potential energy of a polytrope:

$$E_p = -\frac{3}{5-n} \frac{GM^2}{R} [\text{J}].$$

For a star in convective equilibrium (that is,  $n = 3/2$ ) the factor multiplying  $GM^2/R$  becomes  $6/7$  or nearly unity. Note that for a polytrope of index 5,  $U_g \rightarrow -\infty$  implying an infinite central concentration of material. This is also one of the polytropes for which there exists an analytic solution. Thus, one has the picture of a mass point surrounded by a massless envelope of infinite extent. This equation also tells us that as the polytropic index increases, so does the central concentration.

It is not at all obvious that the total gravitational energy would be available to permit the star to shine. Some energy must be provided in the form of heat, to provide the pressure which supports the star. We may use the Virial theorem [equation (1.2.35)] to estimate how much of the gravitational energy can be utilized by the luminosity. Consider a star with no mass motions, so that the macroscopic kinetic energy  $T$  in equation (1.2.35) is zero. Let us also assume that the equilibrium state is good enough that we can replace the time averages by the instantaneous values. Then the **Virial theorem becomes**

$$2E_k + E_p = 0 \text{ [J].}$$

Recall that  $E_k$  is the total internal kinetic energy of the gas which includes all motions of the particles making up the gas. Now we know from thermodynamics that not all the internal kinetic energy is available to do work, and it is therefore not counted in the internal energy of the gas.

For simplicity, let us assume that  $\gamma$  is constant through out the star. Remembering that the total energy  $E$  is the sum of the internal energy and the gravitational energy, we can express the Virial theorem in the following ways:

$$\begin{aligned} E_k &= -\frac{E_p}{3(\gamma - 1)} \text{ [J]} \\ E_{\text{tot}} &= -(3\gamma - 4)E_{\text{th}} \text{ [J]} \\ E_{\text{tot}} &= \frac{3\gamma - 4}{3(\gamma - 1)} E_p \text{ [J].} \end{aligned}$$

It is clear that for  $\gamma > 4/3$  (that is,  $n < 3$ ), the total energy of the star will be negative. This simply says that the star is gravitationally bound and can be in equilibrium. So we can look for the physically reasonable polytropes to have indices less than or equal to 3. The case of  $n = 3$  is an interesting one for it represents radiation dominated gas. In the limit of complete radiation dominance, the total energy of the configuration will be zero.

**Rotational energy:** Stars do rotate, and we should not forget to count the rotational energy in the inventory of energies. We may place a reasonable upper limit on the magnitude of the rotational energy that we can expect by noting that (1) the moment of inertia of the star will always be less than that of a sphere of uniform density and (2) there is a limit to the angular velocity  $\omega_c$  at which the star can rotate. Thus, for a centrally condensed star

$$\begin{aligned} \omega^2 &= \frac{8GM}{27R^3} \text{ [km}^2 \text{s}^{-2} \text{kpc}^{-2}\text]} \\ I_z &< \frac{2}{5}MR^2 \text{ [kg km}^2\text]}, \end{aligned}$$

which implies that the rotational energy must be bounded by

$$E_{\text{rot}} = \frac{1}{2}I_z\omega^2 < \frac{8}{135}\frac{GM^2}{R} \text{ [J].}$$

One may quibble that we have used the angular velocity limit for a centrally condensed star and the moment of inertia for a uniform-density star, but the fact remains that it is extremely difficult for a star to have more than about 10 percent of the magnitude of its gravitational energy stored in the form of rotational energy.

**Nuclear energy:** Of course, the ultimate upper limit for stored energy is the energy associated with the rest mass itself. It is also the common way of estimating the energy available from nuclear sources. Indeed, that fraction of the rest mass which becomes energy when four hydrogen atoms are converted to one helium atom provides the energy to sustain the solar luminosity. Below in Table 2 is a short table giving the mass loss for a few common elements involved in nuclear fusion processes.

Form of Energy	Amount, erg	$\%M_\odot$	Lifetime for constant $L_\odot$ , years
Chemical ( $10^{12}$ erg/g)	$1.98 \times 10^{45}$	$10^{-7}$	14,000
Rotational ( $8 \Omega/81$ )	$2.21 \times 10^{47}$	$10^{-5}$	1.4 million
Gravitational [ $3GM^2/(5R)$ ]	$2.24 \times 10^{48}$	$10^{-4}$	14 million
Nuclear ( $0.0106M_\odot c^2$ )	$1.89 \times 10^{52}$	1	140 billion
Rest mass energy ( $M_\odot c^2$ )	$1.78 \times 10^{54}$	100	$1.4 \times 10^{13}$

Table 2: Possible sources of Solar energy. Table taken from Kippenhahn, Weigert & Weiss (2012).

Clearly most of the energy to be gained from nuclear fusion occurs by the conversion of hydrogen to helium and less than one-half of that energy can be obtained by all other fusion processes that carry helium to iron. Nevertheless, 0.7% of  $mc^2$  is a formidable supply of energy. Table 2 is a summary of the energy that one could consider as being available to the Sun. All these entries are generous upper limits. For example, the sun rotates at less than 0.5% of its critical velocity, it was never composed of 100% hydrogen and will begin to change significantly when a fraction of the core hydrogen is consumed,

and not all the gravitational energy could ever be converted to energy for release. In any event, only nuclear processes hold the promise of providing the solar luminosity for the time required to bring about agreement with the age of the solar system as derived from rocks and meteorites. However, the time scales of Table 2 are interesting because they provide an estimate of how long the various energy sources could be expected to maintain some sort of equilibrium configuration.

**Planets:**

### 1.6.3 Follow-up Questions

- Why do different nuclear reaction pathways have different temperature sensitivities?
- If I assume a constant core temperature on the main sequence, how does stellar radius depend on mass?
- What are some other thermal sources, like say for neutron stars?

## 1.7 Question 7

Describe the process by which supernovae produce light. Why are Type Ia supernovae generally brighter than Type II events?

### 1.7.1 Short answer

Answer.

### 1.7.2 Additional context

Additional context.

### 1.7.3 Follow-up Questions

- When a star goes supernova, how much of the luminous energy generated at the rebound is available for heating the gas? As in, where does the heat come from?
- Are there cases where the rebound shock wave can't blow up the star? Why, and what happens then?

## 1.8 Question 8

Describe the condition for a star's envelope to become convective. Why are low mass stars convective in their outer envelopes while high mass stars are convective in their inner cores?

### 1.8.1 Short answer

In general, convection will occur when

1. the stellar opacity is large, implying that an unachievable steep temperature gradient would be necessary for radiative transport;
2. a region exists where ionization is occurring, causing a large specific heat and a low adiabatic temperature gradient; and
3. the temperature dependence of the nuclear energy generation rate is large, causing a steep radiative flux gradient and a large temperature gradient.

In the atmosphere of many stars, the first two conditions can occur simultaneously, whereas the third condition would occur only deep in stellar interiors. In particular, the third condition can occur when the highly temperature-dependent CNO cycle or triple alpha processes are occurring.

The Sun is purely radiative below  $r/D_\odot = 0.714$  and becomes convective above that point. Physically this occurs because the opacity in the outer portion of the Sun becomes large enough to inhibit the transport of energy by radiation. Stars only slightly more massive than the Sun are convective in their centers because of the stronger temperature dependence of the CNO cycle as compared to the pp chain.

### 1.8.2 Additional context

Three different energy transport mechanisms operate in stellar interiors. **Radiation** allows the energy produced by nuclear reactions and gravitation to be carried to the surface via photons, the photons being absorbed and re-emitted in nearly random directions as they encounter matter. This suggests that the opacity of the material must play an important role, as one would expect. **Convection** can be a very efficient transport mechanism in many regions of a star, with hot, buoyant mass elements carrying excess energy outward while cool elements fall inward. Finally, **conduction** transports heat via collisions between particles. Although conduction can play an important role in some stellar environments, it is generally insignificant in most stars throughout the majority of their lifetimes.

**Radiative transport of energy: basic estimates** Rough estimates show important features of the radiative transfer in stellar interiors and justify an enormous simplification of the formalism. Let us first estimate the mean free path  $\ell_{\text{mfp}}$  of a photon at an “average” point inside a star like the Sun:

$$\ell_{\text{mfp}} = \frac{1}{\kappa\rho} [\text{m}].$$

where  $\kappa$  is a mean absorption coefficient (i.e., a radiative cross section per unit mass averaged over frequency). Typical values for stellar material are of order  $\kappa \approx 1 \text{ cm}^2 \text{ g}^{-1}$ ; for the ionized hydrogen in stellar interiors, a lower limit is certainly the value for electron scattering,  $\kappa \approx 0.4 \text{ cm}^2 \text{ g}^{-1}$ . Using this and the mean density of matter in the Sun,  $\bar{\rho}_\odot = 1.4 \text{ g cm}^{-3}$ , we obtain a mean free path of only

$$\ell_{\text{mfp}} = \frac{1}{\kappa\rho} = \frac{1}{1 \text{ cm}^2 \text{ g}^{-1} \cdot 0.4 \text{ g cm}^{-3}} \approx 2 [\text{cm}],$$

i.e., stellar matter is very opaque.

The typical temperature gradient in the star can be roughly estimated by averaging between centre ( $T_c \approx 10^7 \text{ K}$ ) and surface ( $T_0 \approx 10^4 \text{ K}$ ):

$$\frac{\Delta T}{\Delta r} \approx \frac{T_c - T_0}{R_\odot} \approx 1.4 \times 10^{-4} [\text{K cm}^{-1}].$$

The radiation field at a given point is emitted from a small, nearly isothermal surrounding, the differences of temperature being only of order  $\Delta T = \ell_{\text{mfp}}(dT/dr) \approx 3 \times 10^{-4} \text{ K}$ . Since the energy density of radiation is  $u \sim T^4$ , the relative anisotropy of the radiation at a point with  $T = 10^7 \text{ K}$  is  $\Delta T/T \sim 10^{-10}$ . The situation in stellar interiors must obviously be very close to TE, and the radiation very close to that of a blackbody. Nevertheless, the small remaining anisotropy can easily be the carrier of the stars' huge luminosity: this fraction of  $10^{-10}$  of the flux emitted from  $1 \text{ cm}^2$  of a blackbody of  $T = 10^7 \text{ K}$  is still  $10^3$  times larger than the flux at the solar surface ( $6 \times 10^{10} \text{ erg cm}^{-2} \text{ s}^{-1}$ ). Radiative transport of energy occurs via the non-vanishing net flux (i.e., via the surplus of the outwards-going radiation emitted from somewhat hotter material below over the inwards-going radiation emitted from less-hot material above).

**Diffusion of radiative energy:** The above estimates have shown that for radiative transport in stars, the mean free path  $\ell_{\text{mfp}}$  of the “transporting particles” (i.e., photons) is very small compared to the characteristic length  $R$  (stellar radius) over which the transport extends:  $\ell_{\text{mfp},\odot}/R_\odot \approx 3 \times 10^{-11}$ . In this case, the transport can be treated as a diffusion process, which yields an enormous simplification of the formalism. We derive the corresponding equation by analogy to those for particle diffusion.

The diffusive flux  $j$  of particles (per unit area and time) between places of different particle density  $n$  is given by

$$\vec{j} = -D \nabla n \quad [\text{???}],$$

where  $D$  is the coefficient of diffusion,

$$D = \frac{1}{3} v \ell_{\text{mfp}} \quad [\text{s}^{-1}],$$

determined by the average values of mean velocity  $v$  and mean free path  $\ell_{\text{mfp}}$  of the particles.

In order to obtain the corresponding diffusive flux of radiative energy  $\vec{F}$ ; we replace  $n$  by the energy density of radiation  $U$ ;

$$u = aT^4 = \frac{4\sigma}{c} T^4 \quad [\text{erg cm}^{-3}].$$

Here,  $a = 7.57 \times 10^{-15} \text{ erg cm}^3 \text{ K}^{-4}$  is the **radiation density constant**. Owing to the spherical symmetry of the problem,  $\vec{F}$  has only a radial component  $F_r = |\vec{F}| = F$  and  $\nabla U$  reduces to the derivative in the radial direction

$$\frac{\partial u}{\partial r} = 4aT^3 \frac{\partial T}{\partial r} \quad [\text{erg cm}^{-4}].$$

Using this result along with the equation for  $\vec{j}$ , this immediately gives us

$$F = -\frac{4ac}{3} \frac{T^3}{\kappa\rho} \frac{\partial T}{\partial r} \quad [\text{erg s}^{-1} \text{ cm}^{-2}].$$

Note that this can be interpreted formally as an **equation for heat conduction** by writing

$$\vec{F} = -k_{\text{rad}} \nabla T \quad [\text{erg s}^{-1} \text{ cm}^{-2}],$$

where

$$k_{\text{rad}} = \frac{4ac}{3} \frac{T^3}{\kappa\rho} \quad [\text{erg cm}^{-1} \text{ K}^{-1} \text{ s}^{-1}]$$

represents the **coefficient of conduction** for this radiative transport.

We solve  $F$  for the gradient of the temperature and replace it by the usual local luminosity  $L = 4\pi r^2 F$ ; then

$$\frac{\partial T}{\partial r} = -\frac{3}{16\pi ac} \frac{\kappa\rho L}{r^2 T^3} \quad [\text{K cm}^{-1}]$$

Of course, this neat and simple equation becomes invalid when one approaches the surface of the star. Because of the decreasing density, the mean free path of the photons will there become comparable with (and finally larger than) the remaining distance to the surface; hence the whole diffusion approximation breaks down, and one has to solve the far more complicated full set of transport equations for radiation in the stellar atmosphere (these equations indeed yield our simple diffusion approximation as the proper limiting case for large optical depths).

**The Rosseland mean for  $\kappa_\nu$ :** The above equations are independent of the frequency  $\nu$ ;  $F$  and  $L$  are quantities integrated over all frequencies, so that the quantity must represent a “proper mean” over  $\nu$ . We shall now prescribe a method for this averaging.

In general the absorption coefficient depends on the frequency  $\nu$ . Let us denote this by adding a subscript  $\nu$  to all quantities that thus become frequency dependent:  $\kappa_\nu$ ,  $\ell_\nu$ ,  $D_\nu$ ,  $u_\nu$  etc.

For the diffusive energy flux  $\vec{F}_\nu$  of radiation in the interval  $[\nu, \nu + d\nu]$ , we write now

$$\begin{aligned} \vec{F}_\nu &= -D_\nu \nabla u_\nu \quad [\text{erg s}^{-1} \text{ cm}^{-2}], \\ D_\nu &= \frac{1}{3} c \ell_\nu = \frac{c}{3\kappa_\nu \rho} \quad [\text{m}^2 \text{ s}^{-1}], \end{aligned}$$

while the energy density in the same interval is given by

$$u_\nu = \frac{4\pi}{c} B(\nu, T) = \frac{8\pi h}{c^3} \frac{\nu^3}{e^{h\nu/k_B T} - 1} \quad [\text{erg cm}^{-3}].$$

$B(\nu, T)$  denotes here the **Planck function** for the intensity of blackbody radiation (differing from the usual formula for the energy density simply by the factor  $4\pi/c$ ). From this equation for  $u_\nu$ , we have

$$\nabla u_\nu = \frac{4\pi}{c} \frac{\partial B}{\partial T} \nabla T,$$

which together with  $D_\nu$  is inserted into  $F_\nu$ , the latter then being integrated over all frequencies to obtain the total flux  $F$ :

$$\vec{F} = - \left[ \frac{4\pi}{3\rho} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial B}{\partial T} d\nu \right] \nabla T \text{ [m}^2 \text{s}^{-1}\text{].}$$

We have thus regained  $\vec{F} = -k_{\text{rad}} \nabla T$  with

$$k_{\text{rad}} = \frac{4\pi}{3\rho} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial B}{\partial T} d\nu \text{ [erg cm}^{-1} \text{K}^{-1} \text{s}^{-1}\text{].}$$

Equating this expression for  $k_{\text{rad}}$  with that in the averaged form, we have immediately the proper formula for averaging the absorption coefficient:

$$\frac{1}{\kappa} = \frac{\pi}{acT^3} \int_0^\infty \frac{\partial B}{\partial T} d\nu \text{ [cm}^2 \text{g}^{-1}\text{].}$$

This is the so-called **Rosseland mean**. Since

$$\int_0^\infty \frac{\partial B}{\partial T} d\nu = \frac{acT^3}{\pi},$$

the Rosseland mean is formally the harmonic mean of with the weighting function  $\partial B/\partial T$ , and it can simply be calculated, once the function  $\kappa_\nu$  is known from atomic physics.

In order to see the physical interpretation of the Rosseland mean, we rewrite  $\vec{F}_\nu = -D_\nu \nabla u_\nu$  as

$$\vec{F}_\nu = - \left( \frac{1}{\kappa_\nu} \frac{\partial B(\nu, T)}{\partial T} \right) \frac{4\pi}{3\rho} \nabla T \text{ [m}^2 \text{s}^{-1}\text{].}$$

This shows that, for a given point in the star ( $\rho$  and  $\nabla T$  given), the integrand in  $1/\kappa$  is at all frequencies proportional to the net flux  $\vec{F}_\nu$  of energy. The Rosseland mean therefore favours the frequency ranges of maximum energy flux. One could say that an average *transparency* is evaluated rather than an *opacity* – which is plausible, since it is to be used in an equation describing the transfer of energy rather than its blocking.

One can also easily evaluate the frequency where the weighting function  $\partial B/\partial T$  has its maximum. One finds that, for given a temperature,  $\partial B/\partial T \sim x^4 e^x (e^x - 1)^{-2}$  with  $x = h\nu/k_B T$ . Differentiation with respect to  $x$  shows that the maximum of  $\partial B/\partial T$  is close to  $x = 4$ .

The way we have defined the Rosseland mean  $\kappa$ , which is a kind of weighted harmonic mean value, has the uncomfortable consequence that the opacity  $\kappa$  of a mixture of two gases having the opacities  $\kappa_1, \kappa_2$  is not the sum of the opacities:  $\kappa \neq \kappa_1 \kappa_2$ . Therefore, in order to find  $\kappa$  for a mixture containing the weight fractions  $X$  of hydrogen and  $Y$  of helium, the mean opacities of the two single gases are of no use. Rather one has to add the frequency-dependent opacities  $\kappa_\nu = X_{\kappa_\nu, \text{H}} + Y_{\kappa_\nu, \text{He}}$  before calculating the Rosseland mean. For any new abundance ratio  $X/Y$  the averaging over the frequency has to be carried out separately.

In the above we have characterized the energy flux due to the diffusion of photons by  $F_\nu$ . Since in the following we shall encounter other mechanisms for energy transport, from now on we shall specify this radiative flux by the vector  $\vec{F}_{\text{rad}}$ . Correspondingly we shall use  $\kappa_{\text{rad}}$  instead of  $\kappa$ , etc.

**Conductive transport of energy:** In heat conduction, energy transfer occurs via collisions during the random thermal motion of the particles (electrons and nuclei in completely ionized matter, otherwise atoms or molecules). A basic estimate shows that in “ordinary” stellar matter (i.e., in a non-degenerate gas), conduction has no chance of taking over an appreciable part of the total energy transport. Although the collisional cross sections of these charged particles are rather small at the high temperatures in stellar interiors ( $10^{-18} - 10^{-20} \text{ cm}^2$  per particle), the large density ( $\rho = 1.4 \text{ g cm}^{-3}$  in the Sun) results in mean free paths several orders of magnitude less than those for photons; and the velocity of the particles is only a few per cent of  $c$ . Therefore the coefficient of diffusion is much smaller than that for photons.

The situation becomes quite different, however, for the cores of evolved stars where the electron gas is highly degenerate. The density can be as large as  $10^6 \text{ g cm}^{-3}$ . But degeneracy makes the electrons

much faster, since they are pushed up close to the Fermi energy; and degeneracy increases the mean free path considerably, since the quantum cells of phase space are filled up such that collisions in which the momentum is changed become rather improbable. Then the coefficient of diffusion (which is proportional to the product of mean free path and particle velocity) is large, and heat conduction can become so efficient that it short-circuits the radiative transfer.

The energy flux  $\vec{F}_{cd}$  due to heat conduction may be written as

$$\vec{F}_{cd} = -k_{cd}\nabla T \text{ [m}^2\text{s}^{-1}\text{].}$$

The sum of the conductive flux  $\vec{F}_{cd}$  and the radiative flux  $\vec{F}_{rad}$  is

$$\vec{F} = \vec{F}_{rad} + \vec{F}_{cd} = -(k_{rad} + k_{cd})\nabla T \text{ [m}^2\text{s}^{-1}\text{],}$$

On the other hand, we can just as well write the conductive coefficient  $k_{cd}$  formally in analogy to  $k_{rad}$  as

$$k_{cd} = \frac{4ac}{3} \frac{T^3}{\kappa_{cd}\rho} \text{ [cm}^2\text{g}^{-1}\text{],}$$

hence defining the **conductive opacity**  $\kappa_{cd}$ . Then the flux becomes

$$\vec{F} = -\frac{4ac}{3} \frac{T^3}{\rho} \left( \frac{1}{\kappa_{rad}} + \frac{1}{\kappa_{cd}} \right) \nabla T,$$

which shows that we arrive formally at the same type of equation as in the pure radiative case if we replace  $1/\kappa$  by  $1/\kappa_{rad} + 1/\kappa_{cd}$ . Again the result is plausible, since the mechanism of transport that provides the largest flux will dominate the sum (i.e., the mechanism for which the stellar matter has the highest “transparency”).

The basic equation for energy transport which, if we define properly, holds for radiative and conductive energy transport, can be rewritten in a convenient form. Starting with

$$\frac{\partial T}{\partial r} = -\frac{3}{16\pi ac} \frac{\kappa\rho L}{r^2 T^3} \text{ [K cm}^{-1}\text{],}$$

we can transform to the independent variable mass  $m$  using the fact that  $\partial r / \partial m = (4\pi r^2 \rho)^{-1}$ :

$$\frac{\partial T}{\partial m} = -\frac{3}{64\pi^2 ac} \frac{\kappa L}{r^4 T^3} \text{ [K g}^{-1}\text{].}$$

Assuming hydrostatic equilibrium, we can divide this by  $\partial P / \partial r = -(Gm\rho)/r^2$  and obtain

$$\frac{(\partial T / \partial m)}{(\partial P / \partial m)} = \frac{3}{16\pi ac G} \frac{\kappa L}{m T^3} \text{ [K P g}^{-2}\text{].}$$

We call the ratio of the derivatives on the left  $(dT/dP)_{rad}$ , and we mean by this the variation of  $T$  in the star with depth, where the depth is expressed by the pressure, which increases monotonically inwards. In this sense, in a star which is in hydrostatic equilibrium and transports the energy by radiation (and conduction),  $(dT/dP)_{rad}$  is a gradient describing the temperature variation with depth. If we use the customary abbreviation

$$\nabla_{rad} \equiv \left( \frac{d \ln T}{d \ln P} \right)_{rad} \text{ [K P}^{-1}\text{],}$$

which can be written in the form

$$\nabla_{rad} = \frac{3}{16\pi ac G} \frac{\kappa L P}{m T^4} \text{ [K P}^{-1}\text{],}$$

in which conduction effects are now included.  $\nabla_{rad}$  means a spatial derivative (connecting  $P$  and  $T$  in two neighbouring mass shells), while  $\nabla_{ad}$  describes the thermal variation of one and the same mass element during its adiabatic compression. Only in special cases  $(d \ln T / d \ln P)$  and  $\nabla_{ad}$  will have the same value, and we then speak of an **adiabatic stratification**.

**Transport of energy by convection:** Convective transport of energy means an exchange of energy between hotter and cooler layers in a dynamically unstable region through the exchange of macroscopic mass elements (“blobs”, “bubbles”, “convective elements”), the hotter of which move upwards while the cooler ones descend. The moving mass elements will finally dissolve in their new surroundings and thereby deliver their excess (or deficiency) of heat. Owing to the high density in stellar interiors, convective transport can be very efficient. However, this energy transfer can operate only if it finds a sufficient driving mechanism in the form of the buoyancy forces.

A thorough theoretical treatment of convective motions and transport of energy is extremely difficult. It is the prototype of the many astrophysical problems in which the bottleneck preventing decisive progress

is the difficulty involved in solving the well-known hydrodynamic equations. For simplifying assumptions, solutions are available that may even give reasonable approximations for certain convective flows in the laboratory. Unfortunately, convection in stars proceeds under rather malicious conditions: turbulent motion transports enormous fluxes of energy in a very compressible gas, which is stratified in density, pressure, temperature, and gravity over many powers of ten. Nevertheless, large efforts have been made over many years to solve this notorious problem, and they have partly arrived at promising results. None of the so-called Reynolds stress models, however, have reached a stage where it could provide a procedure easy enough to be handled in everyday stellar-structure calculations, and at the same time would describe the full properties of convection accurately enough. On the other hand, full two- and three-dimensional hydrodynamical simulations have also made large progress, thanks to the impressive advances in supercomputer technology and efficient numerical algorithms. They give valuable hints to the true nature of convection and often serve as numerical experiments to test the dynamical methods. Nevertheless, these numerical simulations are still limited in their size and thus can follow convection in most cases only for a limit time and only for thin convection zones. But even if these restrictions can be foreseen to get relaxed with time, such full hydrodynamical simulations will never be used in full stellar evolution models, as they would unnecessarily follow the star's evolution on a dynamical timescale, which is so much shorter than the dominant nuclear one. Therefore, we limit ourselves exclusively to the description of the old so-called **mixing-length** theory. The reason for this is not that we believe it to be sufficient, but it does provide at least a simple method for treating convection locally, at any given point of a star. Moreover, empirical tests of the resulting stellar models show a surprisingly good agreement with observations. And, finally, even this poor approximation shows without any doubt that in the very deep interior of a star, a detailed theory is normally not necessary.

Note that in the following we are dealing only with convection in stars that are in hydrostatic equilibrium. We furthermore assume that the convection is time independent, which means that it is fully adjusted to the present state of the star. Otherwise, a convection theory for rapidly changing regions (time-dependent convection) has to be developed.

The gradient  $\nabla_{\text{rad}}$  is what would be maintained in a star if the whole luminosity  $L$  had to be transported outwards by radiation only. If convection contributes to the energy transport, the actual gradient  $\nabla$  will be different (namely smaller). In the following, we will estimate  $\nabla$  in the case of convection.

**The basic picture:** The mixing-length theory goes back to Ludwig Prandtl, who in 1925 modelled a simple picture of convection in complete analogy to molecular heat transfer: the transporting "particles" are macroscopic mass elements ("blobs") instead of molecules; their mean free path is the **mixing length** after which the blobs dissolve in their new surroundings. Prandtl's theory was adapted for stars afterwards.

The radiation pressure gradient is given by

$$\frac{dP_{\text{rad}}}{dr} = -\frac{\kappa\rho}{c} F_{\text{rad}} \quad [\text{P m}],$$

where  $F_{\text{rad}}$  is the outward radiative flux. The radiation pressure may also be expressed as

$$\frac{dP_{\text{rad}}}{dr} = \frac{4}{3} a T^3 \frac{dT}{dr} \quad [\text{P m}].$$

Equating the two, we have

$$\frac{dT}{dr} = -\frac{3}{4ac} \frac{\kappa\rho}{T^3} F_{\text{rad}} \quad [\text{K m}^{-1}].$$

If we use the expression for the radiative flux written in terms of the local radiative luminosity of the star at radius  $r$ ,

$$F_{\text{rad}} = \frac{L_r}{4\pi r^2} \quad [\text{erg s}^{-1} \text{ cm}^{-2}],$$

the temperature gradient for radiative transport becomes

$$\frac{dT}{dr} = -\frac{3}{4ac} \frac{\kappa\rho}{T^3} \frac{L_r}{4\pi r^2} \quad [\text{K m}^{-1}].$$

As either the flux or the opacity increases, the temperature gradient must become steeper (i.e., more negative) if radiation transport is to carry all of the required luminosity outward. The same situation holds as the density increases or the temperature increases.

We now consider a situation where a hot convective bubble of gas rises and expands **adiabatically**, meaning that the bubble does not exchange heat with its surroundings. After it has travelled some distance, it finally **thermalizes**, giving up any excess heat as it loses its identity and dissolves into the surrounding gas. Differentiating the **ideal gas law**

$$P = \frac{\rho k_B T}{\mu m_H} [\text{P}]$$

yields an expression involving the bubble's **temperature gradient**

$$\frac{dP}{dr} = -\frac{P}{\mu} \frac{d\mu}{dr} + \frac{P}{\rho} \frac{d\rho}{dr} + \frac{P}{T} \frac{dT}{dr}.$$

Using the adiabatic relationship between pressure and density

$$PV^\gamma = K [\text{P m}^3]$$

and recalling that  $V \equiv 1/\rho$  is the **specific volume**, we have that

$$P = K\rho^\gamma [\text{P}].$$

Differentiating and re-writing, we obtain

$$\frac{dP}{dr} = \gamma \frac{P}{\rho} \frac{d\rho}{dr} [\text{P m}^{-1}].$$

If we assume for simplicity that the mean molecular weight  $\mu$  is constant, we can combine equations for  $dP/dr$  to give the **adiabatic temperature gradient**

$$\left( \frac{dT}{dr} \right)_{\text{ad}} = \left( 1 - \frac{1}{\gamma} \right) \frac{T}{P} \frac{dP}{dr} [\text{K m}^{-1}].$$

Using the equation of **hydrostatic equilibrium**

$$\frac{dP}{dr} = -\frac{GM_r \rho}{r^2} = -\rho g [\text{P m}^{-1}], \quad g \equiv \frac{GM_r}{r^2} [\text{m s}^{-2}]$$

and the ideal gas law, we finally obtain

$$\left( \frac{dT}{dr} \right)_{\text{ad}} = \left( 1 - \frac{1}{\gamma} \right) \frac{\mu m_H}{k} \frac{GM_r}{r^2} [\text{K m}^{-1}].$$

If the star's actual temperature gradient is *steeper* than the adiabatic temperature gradient,

$$\left| \frac{dT}{dr} \right|_{\text{act}} > \left| \frac{dT}{dr} \right|_{\text{ad}} [\text{K m}^{-1}],$$

the temperature gradient is said to be **super-adiabatic** (recall that  $dT/dr < 0$ ). In the deep interior of a star, if the actual temperature gradient is just *slightly* larger than the adiabatic temperature gradient, this may be sufficient to carry nearly all of the luminosity by convection. Consequently, it is often the case that either radiation or convection dominates the energy transport in the deep interior of stars, while the other energy transport mechanism contributes very little to the total energy outflow. The particular mechanism in operation is determined by the temperature gradient. However, near the surface of the star, the situation is much more complicated: both radiation and convection can carry significant amounts of energy simultaneously.

This condition for convection can be used to find another useful relation. Since  $dT/dr < 0$  and  $1/\gamma - 1 < 0$  (recall that  $\gamma > 1$ ),

$$\frac{T}{P} \left( \frac{dT}{dr} \right)^{-1} \frac{dP}{dr} < -\frac{1}{\gamma^{-1} - 1} [\text{dimensionless}],$$

which may be simplified to give

$$\frac{T}{P} \frac{dP}{dT} < \frac{\gamma}{\gamma - 1} [\text{dimensionless}],$$

or, for convection to occur,

$$\frac{d \ln P}{d \ln T} < \frac{\gamma}{\gamma - 1} [\text{P K}^{-1}].$$

For an ideal monatomic gas,  $\gamma = 5/3$  and convection will occur in some region of a star when  $d \ln P/d \ln T < 2.5$ . In that case, the temperature gradient ( $dT/dr$ ) is given approximately by

$$\left( \frac{dT}{dr} \right)_{\text{ad}} = -\left( 1 - \frac{1}{\gamma} \right) \frac{\mu m_H}{k} \frac{GM_r}{r^2} [\text{K m}^{-1}].$$

By comparing this with the temperature gradient for radiative transport

$$\frac{dT}{dr} = -\frac{3}{4ac} \frac{\bar{\kappa}\rho}{T^3} \frac{L_r}{4\pi r^2} [\text{K m}^{-1}]$$

together with the condition for convection written in terms of the temperature gradient,

$$\left| \frac{dT}{dr} \right|_{\text{act}} > \left| \frac{dT}{dr} \right|_{\text{ad}} [\text{K m}^{-1}],$$

it is possible to develop some understanding of which conditions are likely to lead to convection over radiation. In general, convection will occur when

1. the stellar opacity is large, implying that an unachievable steep temperature gradient would be necessary for radiative transport;
2. a region exists where ionization is occurring, causing a large specific heat and a low adiabatic temperature gradient; and
3. the temperature dependence of the nuclear energy generation rate is large, causing a steep radiative flux gradient and a large temperature gradient.

In the atmosphere of many stars, the first two conditions can occur simultaneously, whereas the third condition would occur only deep in stellar interiors. In particular, the third condition can occur when the highly temperature-dependent CNO cycle or triple alpha processes are occurring.

**Convective regions:** Knowledge of the extension of convective regions is very important in view of their influence on the ensuing chemical evolution. A rough overview can be obtained from Fig. 22.7, where  $m/M$  and  $\log M/M_\odot$  are ordinate and abscissa. For any given stellar mass  $M$  along a line parallel to the ordinate it is indicated what conditions we would encounter when drilling a radial borehole from the surface to the centre. In particular, one can see whether the corresponding mass elements are convective or radiative. Aside from the stars of smallest mass ( $M < 0.25 M_\odot$ ), we can roughly distinguish between two types of models:

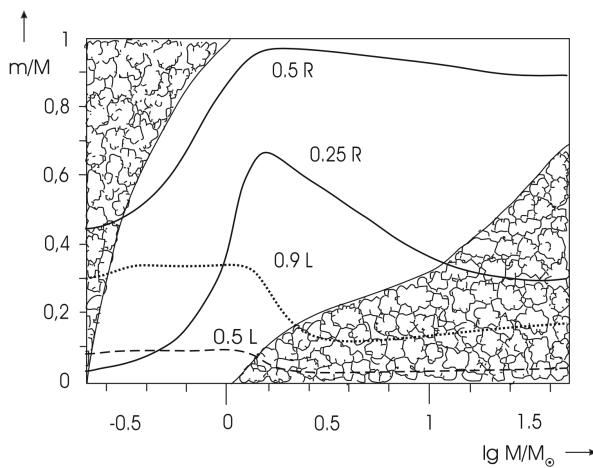
$$\begin{aligned} &\text{radiative core + convective envelope (lower MS)} \\ &\text{convective core + radiative envelope (upper MS).} \end{aligned}$$

The transition from one type to the other occurs near  $M = 1 M_\odot$ .

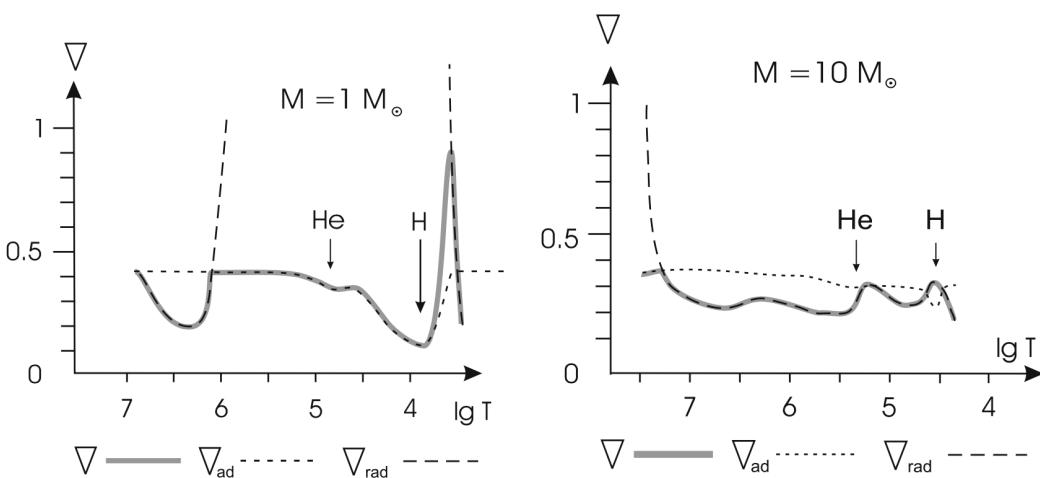
The distinction between convective and radiative regions is made here by using the **Schwarzschild criterion**, which predicts convection if the radiative gradient of temperature  $\nabla_{\text{rad}}$  exceeds the adiabatic gradient  $\nabla_{\text{ad}}$ . The variation of these gradients (together with that of the actual gradient  $\nabla$ ) throughout the star is plotted in Fig. 22.8 for  $M = 1 M_\odot$  and  $10 M_\odot$ . For the abscissa,  $\log T$  is chosen, since this conveniently stretches the scale in the complicated outer layers.

Let us start with the simpler situation concerning the convective core. When comparing Figure 13a, b, we see that the convective core in the more massive models is caused by a steep increase of  $\nabla_{\text{rad}}$  towards the centre. The reason for this is that the dominating CNO cycle, with its extreme temperature sensitivity, concentrates the energy production very much towards the centre (cf. the curve l=L D 0:5 in Fig. 22.7, and Fig. 22.4e). Therefore we find in these stars very high fluxes of energy ( $L = 4\pi r^2$ ) at small  $r$ , which produce large values of  $\nabla_{\text{rad}}$ . Figure 12 shows a remarkable increase in the extent of the convective core for increasing  $M$ . The core covers as much as 65% of the stellar mass in a star of  $50 M_\odot$ , an increase caused by the increasing radiation pressure which depresses the value of  $\nabla_{\text{ad}}$  well below its standard value of 0.4 for an ideal monatomic gas. In the centre of the  $50 M_\odot$  model, roughly  $1/3$  of  $P$  is radiation pressure, and  $\nabla_{\text{ad}} \approx 0.27$ . From Figure 13b it is clear that a depression of  $\nabla_{\text{ad}}$  in the central region will shift the intersection with  $\nabla_{\text{rad}}$  (i.e., the border of the convective core) outwards to smaller  $T$ . When we increase  $M$  to much larger values still, the top of the convective core will finally approach the surface such that we should obtain fully convective stars. We then approach models of the so-called **supermassive stars**. In less massive stars, the pp chain with its smaller temperature sensitivity dominates. This distributes the energy production over a much larger area, so that the flux ( $L/4\pi r^2$ ) and  $\nabla_{\text{rad}}$  are much smaller in the central region, which thus remains radiative.

Outer convective envelopes can generally be expected to occur in stars of low effective temperature. When studying the different gradients in the outer layers of cool stars (Figure 13a), one finds a variety of complicated details. The variation of  $\nabla_{\text{ad}}$  clearly shows depressions in those regions where the most abundant elements, hydrogen ( $T \gtrsim 10^4 \text{ K}$ ) and helium ( $T \approx 10^5 \text{ K}$ ), are partially ionized. The most striking feature is that  $\nabla_{\text{rad}}$  reaches enormous values (more than  $10^5$ ). This is due to the large opacity  $\kappa$ , which here increases by several powers of 10. Therefore the Schwarzschild criterion indicates convective instability: the models have an outer convective zone. In the largest part of it, the density is so high that convection is very effective and the actual gradient  $\nabla$  is close to  $\nabla_{\text{ad}}$ . Convective transport becomes



**Figure 12:**  
The mass values  $m$  from centre to surface are plotted against the stellar mass  $M$  for the same ZAMS models as in Fig. 22.1. “Cloudy” areas indicate the extension of convective zones inside the models. Two solid lines give the  $m$  values at which  $r$  is  $1/4$  and  $1/2$  the total radius  $R$ : The dashed and dotted lines show the mass elements inside which  $50\%$  and  $90\%$  of the total luminosity  $L$  are produced. Figure taken from Kippenhahn, Weigert & Weiss (2012).



**Figure 13:**  
The grey solid lines show the actual temperature gradient  $\nabla = d \ln T / d \ln P$  over the temperature  $T$  (in K) inside two ZAMS models of  $1 M_\odot$  (left panel) and  $10 M_\odot$  (right panel). The corresponding adiabatic gradients  $\nabla_{ad}$  (dotted lines) and radiative gradients  $\nabla_{rad}$  (dashed lines) are also plotted, and the location of the ionization zones of hydrogen and helium are indicated (arrows). The chemical composition of the models is the same as for those of Figure 12. Figure taken from Kippenhahn, Weigert & Weiss (2012).

ineffective only in the outermost, super-adiabatic part, where  $\nabla$  is clearly above  $\nabla_{ad}$ . Scarcely anything of all these features appears in the hot envelope of the  $10 M_\odot$  star (Figure 13b).  $\nabla_{rad}$  remains nearly at the same level; even the photosphere is too hot for hydrogen to be neutral, and only the small dip from the second He ionization is seen immediately below the photosphere. This causes such a shallow zone with convective instability that only for special cases, depending on the detailed chemical composition, convective motions set it.

The outer convection zone gradually penetrates deeper into the star with decreasing  $T_{eff}$ . Its lower border finally reaches the centre at  $M \lesssim 0.25 M_\odot$  (left end of Figure 12), such that the main-sequence stars of even smaller masses are fully convective.

**The present-day interior structure of the Sun:** Consistent with the current age of the Sun, a **solar model** may be constructed for the present-day Sun. Table 3 gives the values of the central temperature, pressure, density, and composition for one such solar model, and a schematic diagram of the model is shown in Figure 14. According to the evolutionary sequence leading to this model, during its lifetime the mass fraction of hydrogen ( $X$ ) in the Sun’s center has decreased from its initial value of 0.71 to 0.34, while the central mass fraction of helium ( $Y$ ) has increased from 0.27 to 0.64. In addition, due to diffusive settling of elements heavier than hydrogen, the mass fraction of hydrogen near the surface has increased by approximately 0.03, while the mass fraction of helium has decreased by 0.03.

Because of the Sun’s past evolution, its composition is no longer homogeneous but instead shows the influence of ongoing nucleosynthesis, surface convection, and elemental diffusion (i.e., settling of heavier elements). The composition structure of the Sun is shown in Figure 14 for  ${}^1H$ ,  ${}^3He$ , and  ${}^4He$ . Since the Sun’s primary energy production mechanism is the pp chain,  ${}^3He$  is produced and then destroyed again. At the top of the hydrogen burning region where the temperature is lower,  ${}^3He$  is relatively

Temperature	$1.57 \times 10^7 \text{ K}$
Pressure	$2.342 \times 10^{16} \text{ N m}^{-2}$
Density	$1.527 \times 10^5 \text{ kg m}^{-3}$
X	0.3397
Y	0.6405

Table 3: Central conditions in the Sun. (Data from Bahcall, Pinsonneault, and Basu, Ap. J., 555, 990, 2001.) Table taken from Carroll & Ostlie (2007).

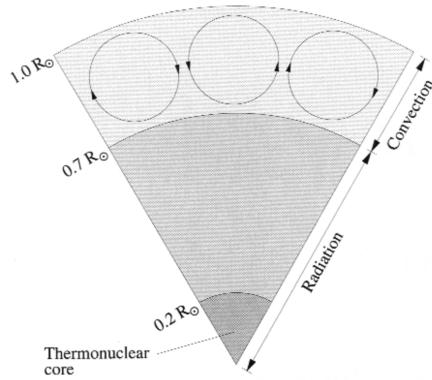


Figure 14:  
Figure taken from  
<https://web.njit.edu/gary/321/Lecture9.html>.

more abundant because it is produced more easily than it is destroyed. At greater depths, the higher temperatures allow the  $^3\text{He}-^3\text{He}$  interactions to proceed more rapidly, and the  $^3\text{He}$  abundance again decreases (the temperature profile of the Sun is shown in Figure 15). The slight ramp in the  $^1\text{H}$  and  $^4\text{He}$  curves near  $0.7 R_\odot$  reflects evolutionary changes in the position of the base of the surface convection zone, combined with the effects of elemental diffusion. Within the convection zone, turbulence results in essentially complete mixing and a homogenous composition. The base of the present-day convection zone is at  $0.714 R_\odot$ .

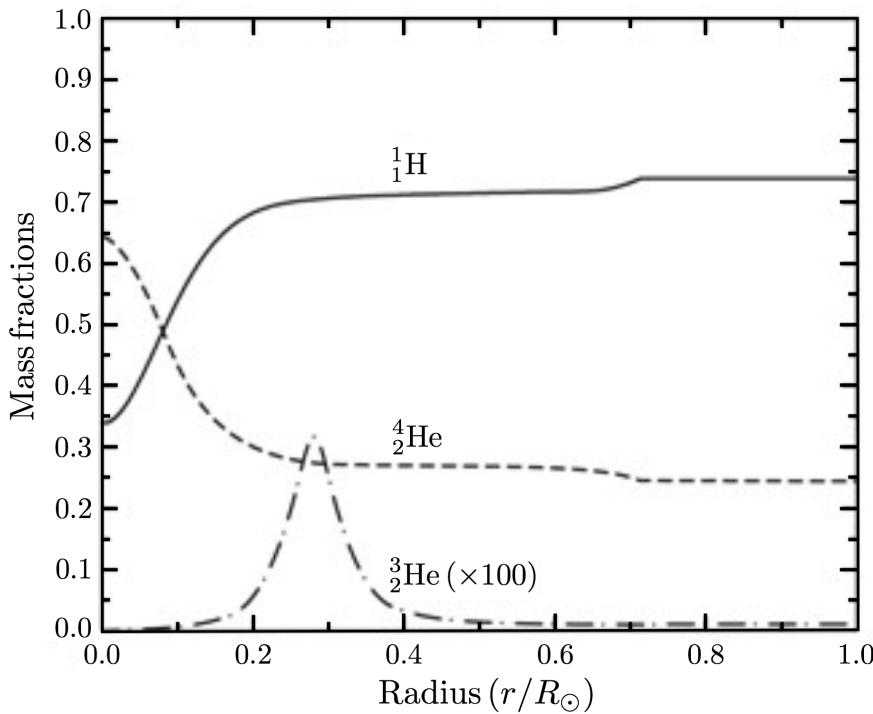


Figure 15:  
The abundance of  $^1\text{H}$ ,  $^3\text{He}$ , and  $^4\text{He}$  as a function of radius for the Sun. Note that the abundance of  $^3\text{He}$  is multiplied by a factor of 100. (Data from Bahcall, Pinsonneault, & Basu, Ap. J., 555, 990, 2001.) Figure adapted from Carroll & Ostlie (2007).

Figure 16 shows  $d \ln P / d \ln T$  versus  $r/R_\odot$ . As can be seen, the Sun is purely radiative below  $r/D_\odot = 0.714$  and becomes convective above that point. Physically this occurs because the opacity in the outer portion of the Sun becomes large enough to inhibit the transport of energy by radiation; recall that the radiative temperature gradient is proportional to the opacity:

$$\frac{dT}{dr} = -\frac{3}{4ac} \frac{\bar{\kappa} \rho}{T^3} \frac{L_r}{4\pi r^2} [\text{K m}^{-1}].$$

When the temperature gradient becomes too large, convection becomes the more efficient means of energy transport. Throughout most of the region of convective energy transport,  $d \ln P / d \ln T \approx 2.5$ , which is

characteristic of the nearly adiabatic temperature gradient of most convection zones. The rapid rise in  $d \ln P / d \ln T$  above  $0.95 R_\odot$  is due to the significant departure of the actual temperature gradient from the adiabatic one. In this case convection must be described by a more detailed treatment, such as **mixing length theory**.

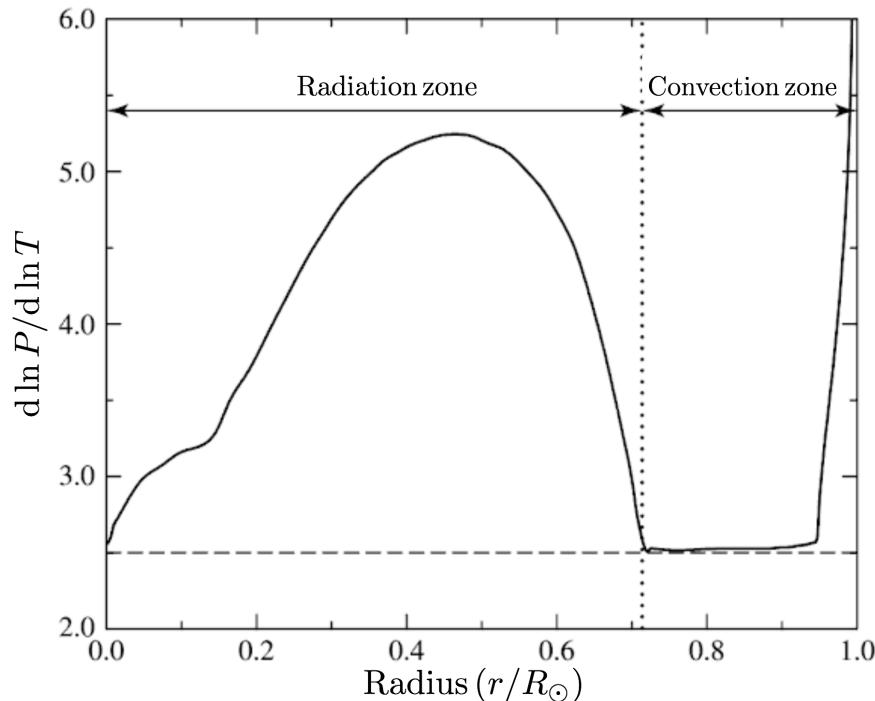


Figure 16:  
The convection condition  $d \ln P / d \ln T$  plotted versus  $r / R_\odot$ . The dashed horizontal line represents the boundary between adiabatic convection and radiation for an ideal monatomic gas. The onset of convection does not exactly agree with the ideal adiabatic case because of the incorporation of a sophisticated equation of state and a more detailed treatment of convection physics. The rapid rise in  $d \ln P / d \ln T$  near the surface is associated with a highly super-adiabatic nature of convection in that region (i.e., the adiabatic approximation that convection occurs when  $d \ln P / d \ln T < 2.5$  is invalid near the surface of the Sun.) [ $d \ln P / d \ln T$  was computed using data from Bahcall, Pinsonneault, and Basu, Ap. J., 555, 990, 2001. The data for the zones above  $0.95 R_\odot$  are from Cox, Arthur N. (editor), Allen's Astrophysical Quantities, Fourth Edition, AIP Press, New York, 2000.] Figure adapted from Carroll & Ostlie (2007).

Notice that  $d \ln P / d \ln T$  also decreases to almost 2.5 at the center of the Sun. Although the Sun remains purely radiative at the center, the large amounts of energy that must be transported outward pushes the temperature gradient in the direction of becoming super-adiabatic. Stars only slightly more massive than the Sun are convective in their centers because of the stronger temperature dependence of the CNO cycle as compared to the pp chain.

### 1.8.3 Follow-up Questions

- How do we know that the Sun's outer envelope is convective?
- How far into the surface of the Sun does the convective zone permeate? How can we measure this?

## 1.9 Question 9

What is Eddington's luminosity limit? Explain why this limit is important for the properties and lifetimes of massive stars.

### 1.9.1 Short answer

Answer.

### 1.9.2 Additional context

Sir Arthur Stanley Eddington observed that radiation and gravitation both obey inverse-square laws and so there would be instances when the two forces could be in balance irrespective of distance. Thus there should exist a maximum luminosity for a star of a given mass, where the force of radiation on the surface material would exactly balance the force of gravity.

There exists a gradient of the radiation pressure

$$\frac{dP_{\text{rad}}}{dr} = \frac{a}{3} T^3 \frac{dT}{dr} [\text{P m}^{-1}],$$

which exerts, just like the gas pressure gradient, an outward acceleration ( $dP_{\text{rad}}/dr < 0$ )

$$g_{\text{rad}} = -\frac{1}{\rho} \frac{dP_{\text{rad}}}{dr} [\text{m s}^{-2}]$$

Using

$$F = -\frac{4ac}{3} \frac{T^3}{\kappa\rho} \frac{\partial T}{\partial r} [\text{erg s}^{-1} \text{cm}^{-2}]$$

we see that we can rewrite this as

$$g_{\text{rad}} = \frac{\kappa F_{\text{rad}}}{c} = \frac{\kappa L_r}{4\pi r^2 c} [\text{m s}^{-2}].$$

In case that radiation pressure completely dominates over gas pressure, a star can no longer be in hydrostatic equilibrium if  $g_{\text{rad}} > g$ . The sum of both accelerations can be written as

$$g + g_{\text{rad}} = -\frac{Gm}{r^2} \left[ 1 - \frac{\kappa L_r}{4\pi c G m} \right] = -\frac{Gm}{r^2} [1 - \Gamma_r] [\text{m s}^{-2}],$$

where  $\Gamma_r$  can be understood as the ratio of the luminosity relative to the critical luminosity at which the bracket changes sign, and thus the star becomes unbound. For  $m = M$  this critical luminosity is called the **Eddington luminosity** and is

$$L_E = \frac{4\pi c G M}{\kappa} [\text{J s}^{-1}].$$

Expressed in Solar units it is

$$\frac{L_E}{L_\odot} = 1.3 \times 10^4 \frac{1}{\kappa} \frac{M}{M_\odot}$$

and grows linearly with stellar mass. Since  $L \propto M^3$ , stars obviously reach a limit, where radiation pressure is able to drive a strong stellar wind, and which depends on the opacity.

For hot, massive stars electron scattering is the dominating opacity source, which can be approximated by

$$\begin{aligned} \kappa_\nu &= \frac{8\pi}{3} \frac{e_e^2}{\mu_e m_u} \\ &= 0.20(1 + X) [\text{cm}^2 \text{g}^{-1}], \end{aligned}$$

and is  $\kappa_{\text{sc}} = 0.20(1 + X)$ . For a mass fraction of hydrogen of 0.70 this simplifies to

$$\frac{L_E}{L_\odot} = 3.824 \times 10^4 \frac{M}{M_\odot}.$$

For  $M \approx 200 M_\odot$  the luminosity of massive main-sequence stars reach the Eddington limit and disperse. This is a rough estimate for an upper limit. In reality the instability of this mechanism occurs at lower mass. However, the Eddington limit can become quite important in other situations.

Any object that has a luminosity greater than  $L_E$  will be forced into instability by its own radiation pressure. This effectively provides a mass-luminosity relationship for super-massive stars since these

radiation-dominated configurations will radiate near their limit. It is worth noting that **magnetars**, the source for **soft gamma repeaters**, are class of gamma-ray bursts that *exceed the Eddington luminosity by far*, but are characterized by a comparably soft gamma spectrum. In a magnetar, the decaying magnetic field is the source of free energy (rather than rotation, as in pulsars).

**Limiting masses for super-massive stars:** Any star with a mass less than about half a million solar masses can come to equilibrium burning hydrogen via the CNO cycle, albeit with a short lifetime. More massive stars are destined to continue to contract. Of course, more massive stars will produce nuclear energy at an ever-increasing rate as their central temperatures rise. However, the rate of energy production cannot increase without bound. This is suggested by the declining exponents of the temperature dependence shown in Table 4. The nuclear reactions that involve  $\beta$  decay set a limit on how fast the CNO cycle can run, and  $\beta$  decay is independent of temperature. So there is a maximum rate at which energy can be produced by the CNO cycle operating in these stars.

Proton-Proton			CNO Cycle			Triple- $\alpha$ Process		
$T_6$	$\epsilon_0$ (cgs)	$v$	$\epsilon_0$ (cgs)	$v$	$T_8$	$\epsilon_0$ (cgs)	$v$	
10	$7 \times 10^{-2}$	4.60	$3 \times 10^{-4}$	22.9	0.8	$2 \times 10^{-12}$	49	
20	1	3.54	$4.5 \times 10^2$	18	1.0	$4 \times 10^{-8}$	41	
40	9	2.72	$3 \times 10^7$	14.1	2.0	15	19	
80	43	2.08	$2 \times 10^{11}$	11.1	3.0	$6 \times 10^3$	12	
100	—	—	$2 \times 10^{12}$	10.2	4.0	$10^5$	7.9	

Table 4: Table taken from Collins (2003).

If the nuclear energy produced is sufficient to bring the total energy above the binding energy curve, the star will explode. However, should the energy not be produced at a rate sufficient to catch the binding energy that is rising due to the relativistic collapse, the star will continue an unrestrained collapse to the Schwarzschild radius and become a black hole. Which scenario is played out will depend on the star's mass. For these stars, the temperature gradient will be above the adiabatic gradient, so convection will exist. However, the only energy transportable by convection is the kinetic energy of the gas, which is an insignificant fraction of the internal energy. Therefore, unlike normal main sequence stars, although it is present, convection will be a very inefficient vehicle for the transport of energy. This is why the star remains with a structure of a polytrope of index  $n = 3$  in the presence of convection. The pressure support that determines the density distribution comes entirely from radiation and is not governed by the mode of energy transport.

The star will radiate at the Eddington luminosity, and that will set the time scale for collapse. The total energy of these stars is small compared to the gravitational energy. So most of the energy derived from gravitational contraction must go into supporting the star, and very little is available to supply the Eddington luminosity. This can be seen from the relativistic Virial theorem which indicates that any change in the gravitational energy is taken up by the kinetic energy. Relativistic particles (in this case, photons) are much more difficult to bind by gravitation than ordinary matter; thus little of the gravitational energy resulting from collapse will be available to let the star shine. The collapse will proceed very quickly on a time scale that is much nearer to the dynamical time scale than the Kelvin-Helmholtz time scale. The onset of nuclear reactions will slow the collapse, but will not stop it for the massive stars.

A dynamical analysis shows that stars more massive than about  $7.5 \times 10^5 M_\odot$  will undergo collapse to a black hole. Here the collapse proceeds so quickly and the gravity is so powerful that the nuclear reactions, being limited by  $\beta$  decay at the resulting high temperatures, do not have the time to produce sufficient energy to arrest the collapse. For less massive stars, this is not the case. Stars in the narrow range of  $5 \times 10^5 M \leq M \leq 7.5 \times 10^5 M_\odot$  will undergo explosive nuclear energy generation resulting in the probable destruction of the star.

Nothing has been said about the role of chemical composition in the evolution of these stars. Clearly, if there is no carbon present, the CNO cycle is not available for the stabilization of the star. Model calculations show that the triple-alpha process cannot stop the collapse. For stars with low metal abundance, only the pp-cycle is available as an energy source. This has the effect of lowering the value of the maximum stable mass. Surprisingly, there is no range at which an explosion occurs. If the star cannot stabilize before reaching  $R_m$ , it will continue in a state of unrestrained gravitational collapse to a black hole. Thus, it seems unlikely that stars more massive than about a half million solar masses could exist.

In addition, it seems unlikely that black holes exist with masses greater than a few solar masses and less than half a million solar masses. If they do, they must form by accretion and not as a single entity.

### 1.9.3 Follow-up Questions

- Draw a force diagram of what's happening.
- What particles experience gravity the most?
- What particles experience photon pressure the most?

## 1.10 Question 10

Explain why we know what the Sun's central temperature ought to be, and how we know what it actually is.

### 1.10.1 Short answer

The hydrostatic condition

$$\frac{\partial P}{\partial r} = -\frac{Gm\rho}{r^2} \text{ [P m}^{-1}\text{]}$$

together with an equation of state for a perfect gas

$$P = \frac{\rho k_B T}{\mu m_H} \text{ [P]}$$

enables us to estimate the pressure and the temperature in the interior of a star of given mass and radius. Let us replace the left-hand side of the equation for hydrostatic equilibrium by an average pressure gradient  $(P_0 - P_c)/r$  where  $P_0$  and  $P_c$  are the pressures at the surface and at the centre, respectively. On the right-hand side, we can replace  $m$  and  $r$  by rough mean values  $M/2$  and  $R/2$ , and obtain

$$P_c \approx \frac{2GM^2}{\pi R^4} \text{ [P].}$$

From the equation of state for a perfect gas (our assumption of a perfect gas turns out to be fully justified for these values of  $P$  and  $T$ ), and with the mean density

$$\bar{\rho} = \frac{3M}{4\pi R^3} \text{ [g cm}^{-3}\text{]},$$

we find with  $P_c$  that

$$\begin{aligned} T_c &= \frac{P_c \mu m_H}{\bar{\rho} k_B} \text{ [K]} \\ &= \frac{\left(\frac{2GM^2}{\pi R^4}\right) \mu m_H}{\left(\frac{3M}{4\pi R^3}\right) k_B} \text{ [K]} \\ &= \left(\frac{2GM^2}{\pi R^4} \mu m_H\right) \cdot \left(\frac{3M}{4\pi R^3} k_B\right)^{-1} \text{ [K]} \\ &= \left(\frac{2GM^2 \mu m_H}{\pi R^4}\right) \cdot \left(\frac{4\pi R^3}{3M k_B}\right) \text{ [K]} \\ &= \frac{8G\mu m_H}{3k_B} \left(\frac{M}{R}\right) \text{ [K].} \end{aligned}$$

Therefore,

$$T_c \propto \frac{M}{R} \text{ [K].}$$

With the mass and the radius of the Sun ( $M_\odot = 1.989 \times 10^{33}$  g,  $R_\odot = 6.96 \times 10^{10}$  cm) and with  $\mu = 0.5$ , we find that

$$\begin{aligned} T_c &= \frac{8G\mu m_H}{3k_B} \left(\frac{1.989 \times 10^{33} \text{ g}}{6.96 \times 10^{10} \text{ cm}}\right) \text{ [K]} \\ &= 3.1 \times 10^7 \text{ [K].} \end{aligned}$$

Modern numerical solutions give  $T_c = 1.6 \times 10^7$  K.

### 1.10.2 Additional context

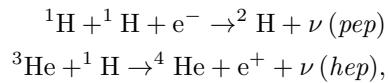
Although the Sun is a very ordinary star of average mass and in a quiet state of MS hydrogen burning, it is a unique object for stellar evolution theorists. For no other star so many quantities are known with comparable accuracy obtained by so many different and independent methods. From Kepler's laws and known distances within the solar system we can derive its mass and radius as well as the total luminosity. This yields the effective temperature by application of the Stefan-Boltzmann law. Neutrino experiments on Earth allow the determination of conditions in the innermost energy producing core. And the art

Quantity	Value	Method
Mass	$(1.9891 \pm 0.0004) \times 10^{33} \text{ g}$	Kepler's third law
Radius	$695,508 \pm 26 \text{ km}$	Angular diameter plus distance
Luminosity	$(3.846 \pm 0.01) \times 10^{33} \text{ erg s}^{-1}$	Solar constant
Effective temp.	$5,779 \pm 2 \text{ K}$	Stefan-Boltzmann law
$Z/X$	0.0245 ± 0.001 0.0165	Meteorites and solar spectrum (new determination)
Age	$4.57 \pm 0.02 \text{ Gyr}$	Radioactive decay in meteorites
Depth of conv. env.	$0.713 \pm 0.001 R_\odot$	Helioseismology
Env. helium content	0.246 ± 0.002	Helioseismology

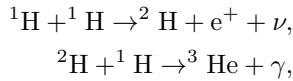
Table 5: Solar quantities and how they are derived. Table taken from Kippenhahn, Weigert & Weiss (2012).

of (helio-)seismology has returned with high accuracy the run of the sound speed throughout most of the solar interior, the helium content of the outer convective envelope, and its depth. These quantities restrict the modelling of the present Sun and allow a comparison with stellar evolution theory at a degree of precision which is almost unique in astrophysics. Table 5 summarizes the fundamental solar parameters and the method to derive them. Note that the rather large uncertainty in the solar mass is the result of the uncertainty in Newton's constant of gravity G. Kepler's third law returns their combination,  $GM_\odot$ , with a precision of  $10^{-7}$ !

**Solar neutrinos:** Some of the nuclear reactions of the pp chain, as well as of the CNO cycle, produce neutrinos. In addition, there are also neutrinos due to the very rare *pep* and *hep* reactions



the latter one being the trivial way to produce  ${}^4\text{He}$  after the **pp-chain**,



but it is occurring in only  $10^{-8}$  of all cases. However, the energy of the emitted neutrino is close to 10 MeV, and it is therefore necessary to consider this reaction. The neutrinos leave the star practically without interacting with the stellar matter. The energy spectrum of neutrinos from  $\beta$  decay is continuous, since the electrons can take part of the energy away, while neutrinos released after an inverse  $\beta$  decay are essentially monochromatic. Therefore most reactions of the pp chain have a continuous spectrum, while the pep-reaction and the electron capture on  ${}^7\text{Be}$  have a line spectrum. Since  ${}^7\text{Be}$  can decay into  ${}^7\text{Li}$  either in the ground state or in an excited state, this reaction gives two spectral lines. The neutrino spectrum of the Sun as predicted from the reactions of the pp chain, computed from our standard solar model, is given in Figure 17.

Since the solar neutrinos can leave the Sun almost unimpeded they can in principle be measured in terrestrial laboratories and thus be used to learn directly about conditions in the innermost solar core. This difficult task indeed has been undertaken since 1964, when John Bahcall and Raymond Davies began to plan for an underground neutrino detector in a mine in Homestead, North Dakota. Forty years later the experiments finally have confirmed the standard solar model, and R. Davies received the Nobel Prize for his work. The time in between, however, was characterized by the **solar neutrino problem**.

The solar neutrino problem consisted in the fact that since the first results from the so-called chlorine experiment by Davies there was a lack of neutrinos compared to solar model predictions. The chlorine experiment is sensitive to neutrinos with energies above 0.814 MeV and therefore, as can be seen in Figure 17 mainly to the  ${}^8\text{B}$  neutrinos, with some contribution from pep, hep, and  ${}^7\text{Be}$  neutrinos. The experiment is based on the reaction  ${}^{37}\text{Cl} + \nu \rightarrow {}^{37}\text{Ar}$ , where the decays of radioactive argon nuclei are counted. The rate of neutrino captures is commonly measured in solar neutrino units (SNU). One SNU corresponds to  $10^{-36}$  captures per second and per target nucleus. The predicted counts amount to 7.5 SNU for the chlorine experiment, the measurements averaged over several decades to only  $2.5 \pm 0.2$  SNU. The deficit could indicate that the solar centre is cooler than in the models.

To improve the experimental evidence, additional experiments were started. First, another kind of radiochemical detector using gallium in the detector fluid measured, due to a much lower energy threshold, the majority of neutrinos, including those from the pp-reaction. Later, electron-scattering detectors were developed, which are sensitive to the highest energies only, but which provide directional information about the neutrino source (for these detectors the *hep*-neutrinos have to be taken into account.). All

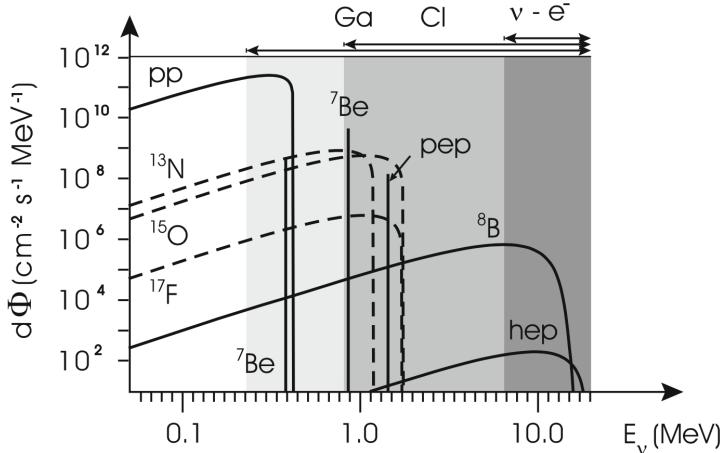


Figure 17: The neutrino spectrum of the Sun as predicted from the theoretical standard solar model. The solid lines belong to reactions of the pp chain while the broken lines are due to reactions of the CNO cycle. The neutrinos from most of the reactions have continuous spectra, while mono-energetic neutrinos come from  ${}^7\text{Be}$  and from the pep-reaction. The flux for the continuum sources is given in  $\text{cm}^2 \text{s}^{-1} \text{MeV}^{-1}$  and for the line sources in  $\text{cm}^2 \text{s}^{-1}$ . The sensitivity of the three types of neutrino experiments is indicated above the figure and by the shaded regions. Figure taken from Kippenhahn, Weigert & Weiss (2012).

experiments confirmed that the solar neutrino flux was of the right order of magnitude, and therefore that indeed the Sun shines by the nuclear fusion of hydrogen, but they also consistently measured a deficit of neutrinos. This deficit, however, varied between different kinds of detectors.

With more and more experimental data it became evident that even hypothetical changes to the solar centre cannot solve the problem and that the solution is most likely to be found in the properties of neutrinos. All nuclear reactions emit electron neutrinos, and these are the only ones that were measured in terrestrial experiment, with the exception of the electron-scattering Sudbury Neutrino Observatory (SNO) experiment in Canada, where heavy water (with a high percentage of deuterium isotopes) was used as the detector. Here also reactions with the two other types (flavours) of neutrinos, muon and tau neutrinos can be detected. Summing these and the electron neutrinos up, the total number of detections is completely consistent with the solar model prediction, within a few percent. What created the apparent solar neutrino deficit is the fact that neutrinos can change their flavour, both while travelling through vacuum and more efficiently in the presence of electrons in the solar interior. A similar effect was also confirmed for muon neutrinos arising in the Earth's upper atmosphere from high-energy cosmic radiation, when measured before or after they have travelled through the Earth's interior. The modelling of the solar interior, together with sophisticated experiments, has therefore resulted in new knowledge about fundamental properties of neutrinos. In particular, these so-called **neutrino oscillations** are possible only if neutrinos have mass.

## 1.11 Question 11

Which have higher central pressure, high-mass or low-mass main-sequence stars? Roughly, what is their mass-radius relation? Derive this.

### 1.11.1 Short answer

Answer.

### 1.11.2 Additional context

Additional context.

### 1.11.3 Follow-up Questions

- How would we actually know the central pressure?
- What properties can we measure to test models of stellar structure?

## 1.12 Question 12

Sketch the SED of an O, A, G, M, and T star. Give defining spectral characteristics, such as the Balmer lines and Balmer jump and Calcium doublets, and describe physically.

### 1.12.1 Short answer

Figure 18 shows the spectral sequence from the hottest **early-type stars**, the O-type stars to spectral type G2. Figure 19 shows the spectral sequence from the coolest **late-type stars**, the G2-type stars to spectral type M4.5. Figure 20 shows the spectrum of a typical mid-T dwarf from the red to the mid-infrared. In the red, the SED is shaped by the very extensive wings of the KI resonance line; in the far red and infrared, the most prominent features are due to CH<sub>4</sub> and H<sub>2</sub>O. Indeed, T dwarfs are distinguished from the L dwarfs by the presence of CH<sub>4</sub> absorption in their near-infrared and infrared spectra. All T dwarfs are brown dwarfs.

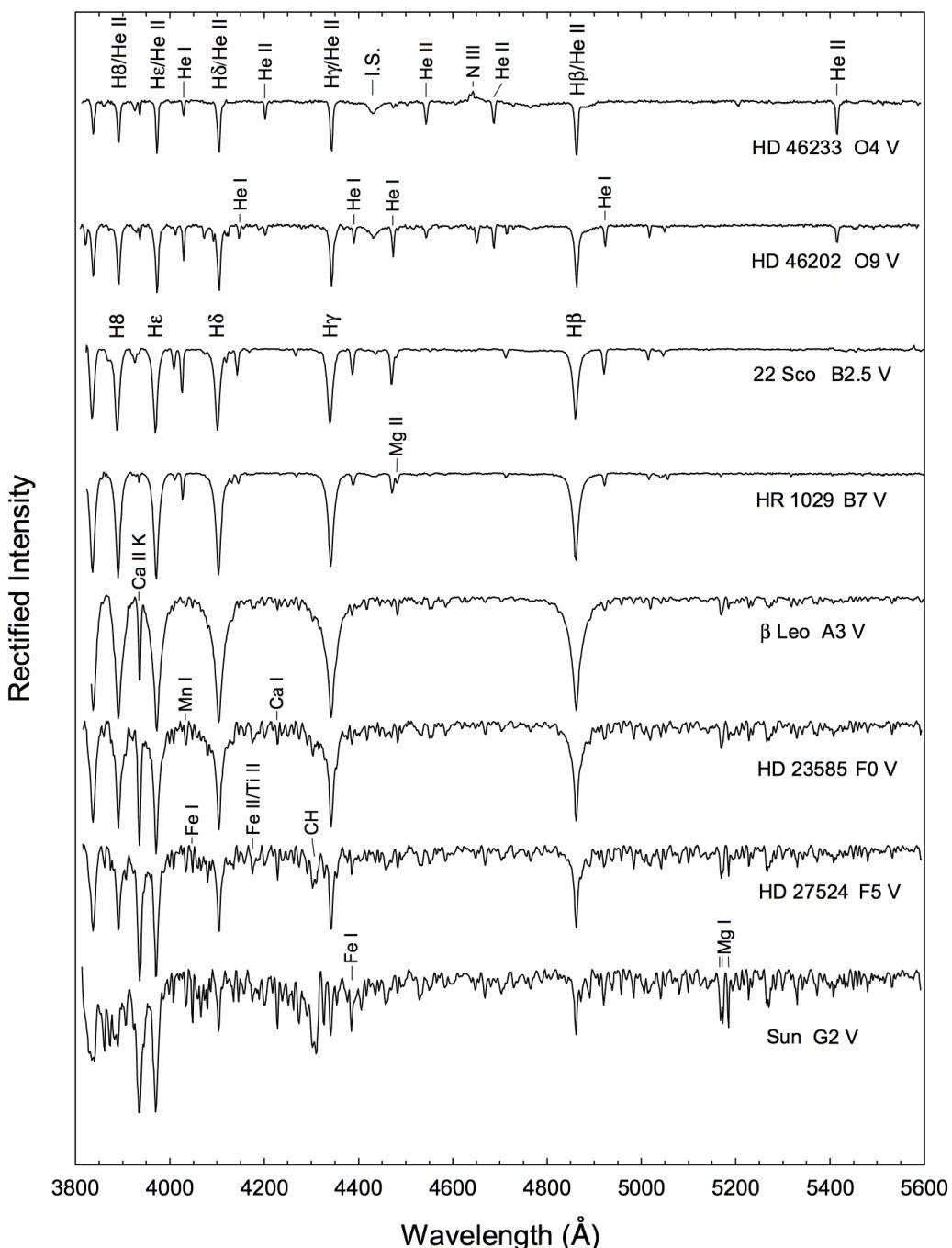


Figure 18: The Main Sequence from O4 to G2, the “early-type” stars. Figure taken from Gray (2009).

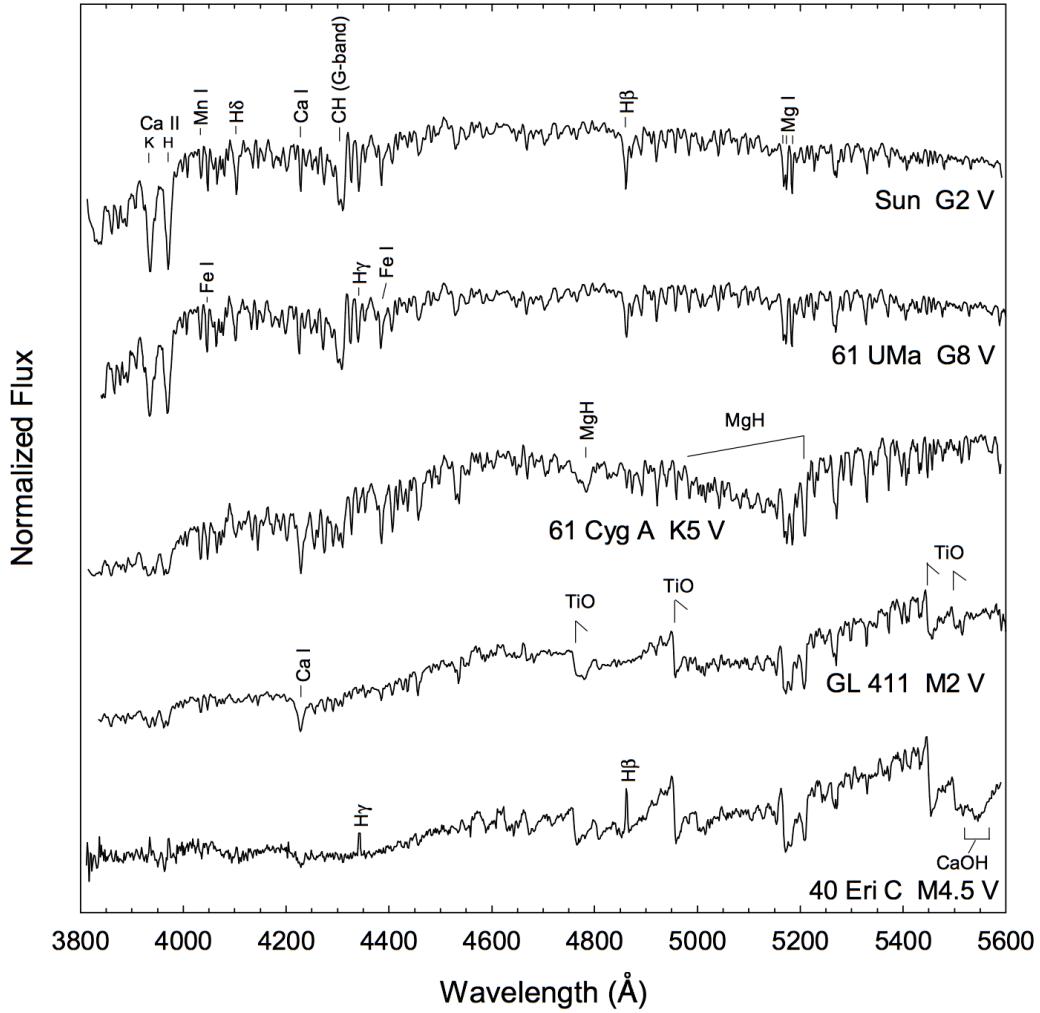


Figure 19: The Main Sequence from G2 to M4.5, the “late-type” stars. Figure taken from Gray (2009).

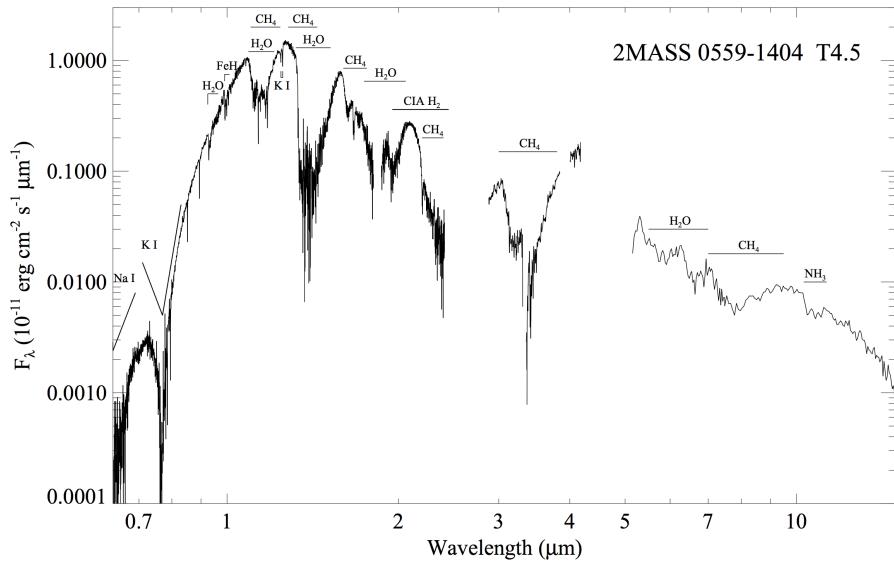


Figure 20: The spectrum of a mid-T dwarf spanning the wavelengths  $0.63 - 15 \mu\text{m}$ . Prominent atomic and molecular features are labeled. Reproduced from Stellar Spectral Classification, courtesy Adam Burgasser. Figure taken from Gray (2009).

### 1.12.2 Additional context

The spectra of early-type stars are dominated by the Balmer lines of hydrogen. Note how they increase in strength and reach a maximum at A2, and then fade thereafter. The O- and B-type stars are characterized by lines of helium; the A-, F-, and G-type stars are characterized by lines of metals. The O-type stars constitute the hottest normal stars, and are characterized by moderately weak hydrogen lines, lines of neutral helium (HeI) and by lines of singly ionized helium (HeII). The spectral type can be judged easily by the ratio of the strengths of lines of HeI to HeII; HeI tends to increase in strength with decreasing temperature while HeII decreases in strength. The ratio HeI 4471Å to HeII 4542Å shows this trend clearly. As we move toward later (cooler) types, the helium lines continue to fade until they essentially disappear in spectra of this resolution (1.8Å) at a spectral type of about A0.

The definition of the break between the O-type stars and the B-type stars is the absence of lines of ionized helium (HeII) in the spectra of B-type stars. The lines of HeI pass through a maximum at approximately B2, and then decrease in strength towards later (cooler) types. Were it not for the presence of helium peculiarities in both the early and late B-type stars, the behavior of HeI would be sufficient to accurately estimate the spectral type in the B-type stars. However, because of these peculiarities, the spectral type in the early B-type stars is, instead, estimated on the basis of ratios of lines of silicon ions. For instance, between B0 and B1 (an interval which can be divided into B0, B0.2, B0.5, B0.7 and B1 classes), the SiIV 4089Å/SiIII 4552Å ratio may be used. Between B1 and B3 the SiIII 4552Å/SiII 4128.32Å ratio may be used.

The **Be stars** are B-type stars that are characterized (or have been characterized in the past) by emission in one or more of the Balmer lines of hydrogen, sometimes accompanied by emission in lines of singly ionized metals, most commonly FeII. The Be class excludes B-type supergiants, as well as pre main-sequence stars such as the Herbig Ae/Be stars. Many Be stars may be classified without difficulty on the MK system, although the more extreme present significant difficulties. An extension to the MK system was introduced for the Be stars comprised of an **e index** ( $e_1 \rightarrow e_4$ , in direction of increasing emission-line strength). Some typical Be spectra are illustrated in Figure 21.

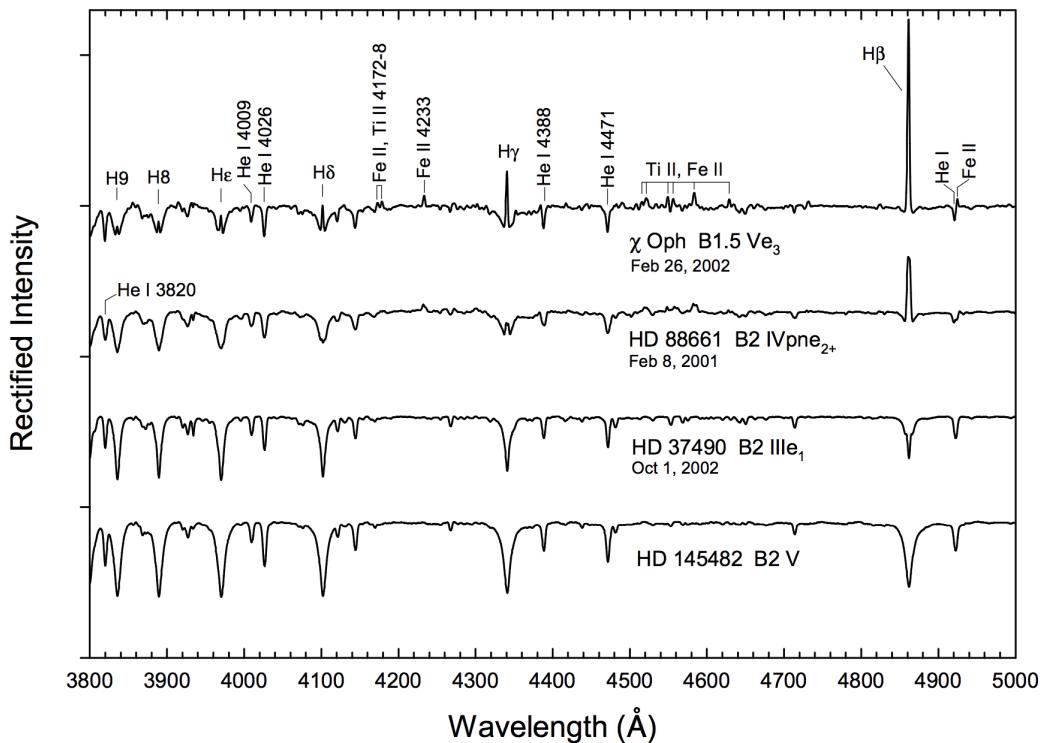


Figure 21: A selection of Be stars compared with a B2V secondary standard, HD 145482. The spectral types are on the Lesh (1968) system. Spectra from the Paranal Observatory spectral library. Figure taken from Gray (2009).

Near a spectral type of A0, the primary luminosity criterion is the progressive widening and strengthening of the hydrogen lines with decreasing luminosity. The hydrogen lines reach a maximum in the A-type stars. The CaII K-line increases significantly in strength through the A-type stars, and its absolute strength, or, more usefully, its ratio with He $\epsilon$  or H $\delta$  is a sensitive indicator of the temperature type, although one that is dependent on the metallicity. The general metallic-line spectrum also increases in strength through the A-type stars. These three criteria give a consistent temperature type in the normal

A-type stars. Any disagreement between these three criteria indicates the star is peculiar in some way. **Ap stars** are A-type (actually more commonly late B-type) stars that show significant enhancements of certain elements. **Herbig Ae stars** generally show emission in the H $\alpha$  line (outside of the spectral range of most of the spectra used in this atlas), and quite often emission in H $\beta$  and even H $\gamma$ . Many Herbig Ae stars are still contracting to the main sequence, and are thus either still surrounded by remnants of their stellar cocoons, or have developed massive stellar winds.

Metallic-line A-type stars, or, for short, **Am stars**, are defined as stars that show a difference between the K-line type and the metallic-line type. Many Am stars show the **anomalous luminosity effect** in which the luminosity criteria in certain spectral regions (in particular, the region from 4077Å–4200Å) may indicate a giant or even a supergiant luminosity, whereas other regions indicate a dwarf luminosity or even lower.

The hydrogen lines continue to weaken through the F-type stars, and the CaII K-line strengthens, although it becomes essentially saturated by the late F-type stars. The general strength of the metallic-line spectrum grows dramatically. Around F2, depending upon the resolution of the spectrum, the G-band makes its first appearance. The G-band is a molecular band, composed of thousands of closely spaced lines due to the diatomic molecule CH. By F0, the hydrogen lines have lost most of their sensitivity to luminosity. Note, however, that they can still be used to distinguish the supergiant classes from lower luminosities. Near F0, the luminosity class is estimated from the strength of lines due to ionized iron and titanium. By F8, the hydrogen lines have lost all sensitivity to luminosity, and it is now necessary to rely solely on lines and blends of ionized species.

Later than G0 along the main sequence, the hydrogen lines continue to fade, while the strength of the general metallic-line spectrum continues to increase. The G-band also increases in strength until the early K-type stars (about K3), and then begins to fade.

Molecular features first appear in the spectra of F-type main sequence stars and grow to dominate the spectra of the late-type stars, especially those of types K and M. The main sequence does not end at M4.5, but continues through the late M-, the L- and the T-type dwarfs, a sequence that spans the transition from hydrogen-burning dwarf stars to brown dwarfs. Visible in a number of stars is a broad, shallow depression, centered near 4430Å. This is an example of a diffuse interstellar band, probably due to molecules in the ISM between us and the star. A number of G and K-type giants show either strong or weak molecular bands involving carbon.

Almost all barium-rich stars are members of binary systems in which the companion is a white dwarf. It is believed that the enhanced abundances of s-process elements, such as barium, were gained from the companion star when it was on the AGB. The mass transfer came about either through **Roche-lobe overflow** or wind accretion. The **s-process elements** are the result of slow neutron capture, which occurs only at advanced stages of nuclear burning.

Notice that H $\beta$  is in emission in the M4.5 star; many M dwarfs have active chromospheres and exhibit strong flares many times more energetic than solar flares. One of the manifestations of this is emission in the hydrogen lines.

### 1.12.3 Follow-up Questions

- Are there emission lines?
- What molecular lines are in the Sun?
- What important lines are there are much longer wavelengths than those in the optical?
- What if the A star had a protoplanetary disk?
- What is the significance of a  $\lambda F_\lambda$  (or  $\nu F_\nu$ ) spectrum?
- How can the relative height of the stellar vs disk bumps change? (i.e., total energy of the system cannot change; dust bump can't be higher than star bump without extinction.)

## 1.13 Question 13

What can be learned about young stars (T Tauri and pre-main-sequence stars) from an analysis of their spectral features?

### 1.13.1 Short answer

Answer.

### 1.13.2 Additional context

Since we think star formation begins with a core that is purely gas, we expect to begin with a cloud that is cold and lacks a central point source. Once a protostar forms, it will begin gradually heating up the cloud, while the gas in the cloud collapses onto the protostar, reducing the opacity. Eventually enough material accretes from the envelope to render it transparent in the near infrared and eventually the optical, and we begin to be able to see the star directly for the first time. The star is left with an accretion disk, which gradually accretes and is then dispersed. Eventually the star contracts onto the main sequence.

This theoretical cartoon has been formalized into a system of classification of young stars based on observational diagnostics. At one end of this sequence lies purely gaseous sources where there is no evidence at all for the presence of a star, and at the other end lies ordinary main sequence stars. In between, objects are classified based on their emission in the infrared and sub-mm parts of the spectrum. These classifications probably give more of an impression of discrete evolutionary stages than is really warranted, but they nonetheless serve as a useful rough guide to the evolutionary state of a forming star. Consider a core of mass  $\sim M$ , seen in dust or molecular line emission. When a star first forms at its center, the star will be very low mass and very low luminosity, and will heat up only the dust nearest to it, and only by a very small amount. Thus the total light output will still be dominated by the thermal emission of the dust at its equilibrium temperature. The spectral energy distribution of the source will therefore look just like that which prevailed before the star formed.

However, there might be other indicators that a star has formed. For example, the density distribution might show a very sharp, unresolved peak. Another sign that a star has formed might be the presence of an outflow, which all protostars seem to generate. Outflows coming from the center of a core can be detected in a few ways. Most directly, one can see bipolar, high velocity structures in molecular emission (Figure 22).

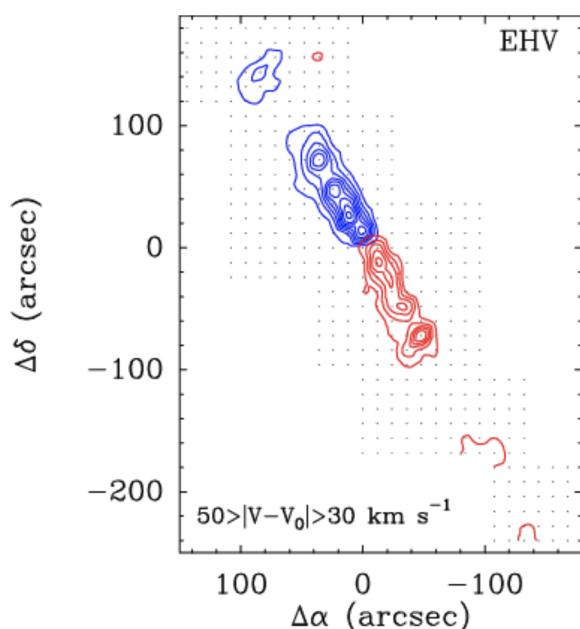


Figure 22:  
An integrated intensity map in  $\text{CO}(2 \rightarrow 1)$ , showing material at velocities between  $\pm 30 - 50 \text{ km s}^{-1}$  (blue and red contours, respectively) relative to the mean (Tafalla et al., 2004). Contours are spaced at intensities of  $1 K \text{ km s}^{-1}$ . The outflow shown is in the Taurus star-forming region. Figure taken from Krumholz (2015).

One can also detect indirect evidence of an **outflow**, from the presence of highly excited molecular line emission that is produced by shocks at hundreds of  $\text{km s}^{-1}$ . One example of such a line is  $\text{SiO}(2 \rightarrow 1)$  line, which is generally seen in gas moving at several tens of  $\text{km s}^{-1}$  with temperatures of several hundred K – this is taken to be indication that emission in this line is produced in warm shocks. Since we know of no processes other than formation of a compact object with a  $\gtrsim 100 \text{ km s}^{-1}$  escape velocity that can

accelerate gas in molecular clouds to such speeds, the presence of such an outflow is taken to indicate that a compact object has formed.

These are the earliest indications of star formation we have available to us. We call objects that show one of these signs, and do not fall into one of the other categories, **class 0 sources**. The dividing line between class 0 and class 1 is that the star begins to produce enough heating that there is non-trivial infrared emission. Before the advent of Spitzer and Herschel, the dividing line between class 0 and 1 was taken to be a non-detection in the IR, but as more sensitive IR telescopes became available, the detection limit went down of course. Today, the dividing line is taken to be a luminosity cut. A source is said to be class 0 if more than 0.5% of its total bolometric output emerges at wavelengths longer than  $350\text{ }\mu\text{m}$ , i.e., if  $L_{\text{smm}}/L_{\text{bol}} > 0.5\%$ , where  $L_{\text{smm}}$  is defined as the luminosity considering only wavelengths of  $350\text{ }\mu\text{m}$  and longer (see Figure 23).

In practice, measuring  $L_{\text{smm}}$  can be tricky because it can be hard to get absolute luminosities (as opposed to relative ones) correct in the sub-mm, so it is also common to define the class 0 – 1 divide in terms of another quantity: the bolometric temperature  $T_{\text{bol}}$ . This is defined as the temperature of a blackbody that has the same flux-weighted mean frequency as the observed SED. That is, if  $F_\nu$  is the flux as a function of frequency from the observed source, then we define  $T_{\text{bol}}$  by the implicit equation

$$\frac{\int \nu F_\nu d\nu}{\int F_\nu d\nu} = \frac{\int \nu B_\nu(T_{\text{bol}}) d\nu}{\int B_\nu(T_{\text{bol}}) d\nu}.$$

The class 0 – 1 dividing line is also sometimes taken to be  $T_{\text{bol}} = 70\text{ K}$ . In cases where  $L_{\text{smm}}$  is accurately measured,  $T_{\text{bol}}$  is observed to be a reasonably good proxy for  $L_{\text{smm}}/L_{\text{bol}}$ .

Once protostars reach **class I**, their evolution into further classes is defined in terms of the infrared SED. The motivating cartoon is as follows. At early times, the envelope of dust around the protostar is very optically thick at visible and even near infrared wavelengths. As a result, we don't get to see the stellar photosphere at all. All the radiation is absorbed by the envelope. The dust is in thermal equilibrium, so it re-radiates that energy. Since the radius of the sphere of dust is much larger than that of the star, and the luminosity radiated by the dust must ultimately be equal to that of the star, this emission must be at lower temperature and thus longer wavelengths. Thus as the radiation propagates outward through the dust it is shifted to longer and longer wavelengths. However, dust opacity decreases with wavelength, and thus eventually the radiation is shifted to wavelengths where the remaining dust is optically thin, and it escapes. What we observe is therefore not a stellar photosphere, but a “dust photosphere”.

Given this picture, the greater the column density of the dust around the star, the further it will have to diffuse in wavelength in order to escape. Thus the wavelength at which the emission peaks, or, roughly equivalently, the slope of the spectrum at a fixed wavelength, is a good diagnostic for the amount of circumstellar dust. Objects whose SEDs peak closer to the visible are presumed to be more evolved, because they have lost most of their envelopes.

More formally, this classification scheme was based on fluxes as measured by the IRAS satellite. We define

$$\alpha_{\text{IR}} = \frac{d \log \lambda F_\lambda}{d \log \lambda} \text{ [dimensionless]},$$

as the **IR spectral index**. In practice we measure  $\alpha_{\text{IR}}$  using two points from the IRAS SED:  $2.2\text{ }\mu\text{m}$  and  $10 - 25\text{ }\mu\text{m}$ . More positive values of  $\alpha_{\text{IR}}$  indicate SEDs that peak at longer wavelengths, further into the IR, while more negative values indicate SEDs that peak closer to visible. We define sources with  $\alpha_{\text{IR}} \geq 0.0$  (i.e., rising at longer wavelengths from  $2 - 25\text{ }\mu\text{m}$ , as class I sources. Alternately, in terms of bolometric temperature, the class I to class II transition is generally taken to be at  $650\text{ K}$ .

As more of the envelope accretes, it eventually becomes optically thin at the peak emitting wavelengths of the stellar photosphere. In this case we see the stellar blackbody spectrum, but there is also **excess infrared emission** coming from the disk of warm, dusty gas that till surrounds the star. Thus the SED looks like a stellar blackbody plus some extra emission at near- or mid-IR wavelengths. Stars in this class are also known as **classical T Tauri stars**, named for the first object of the class, although the observational definition of a T Tauri star is somewhat different than the IR classification scheme, so the alignment may not be perfect.

In terms of  $\alpha_{\text{IR}}$ , these stars have indices in the range  $-1.6 < \alpha_{\text{IR}} < 0$ . (Depending on the author, the break-point may be placed at  $-1.5$  instead of  $-1.6$ . Some authors also introduce an intermediate classification between 0 and I, which they take to be  $-0.3 < \alpha_{\text{IR}} < 0.3$ .) A slope of around  $-1.6$  is what we expect for a bare stellar photosphere without any excess infrared emission coming from circumstellar material. Since the class II phase is the last one during which there is a disk of any significant mass, this is also presumably the phase where planet formation must occur.

The final class is **class III**, which in terms of SED have  $\alpha_{\text{IR}} < -1.6$ . Stars in this class correspond to **weak line T Tauri stars**. The SEDs of these stars look like bare stellar photospheres in the optical

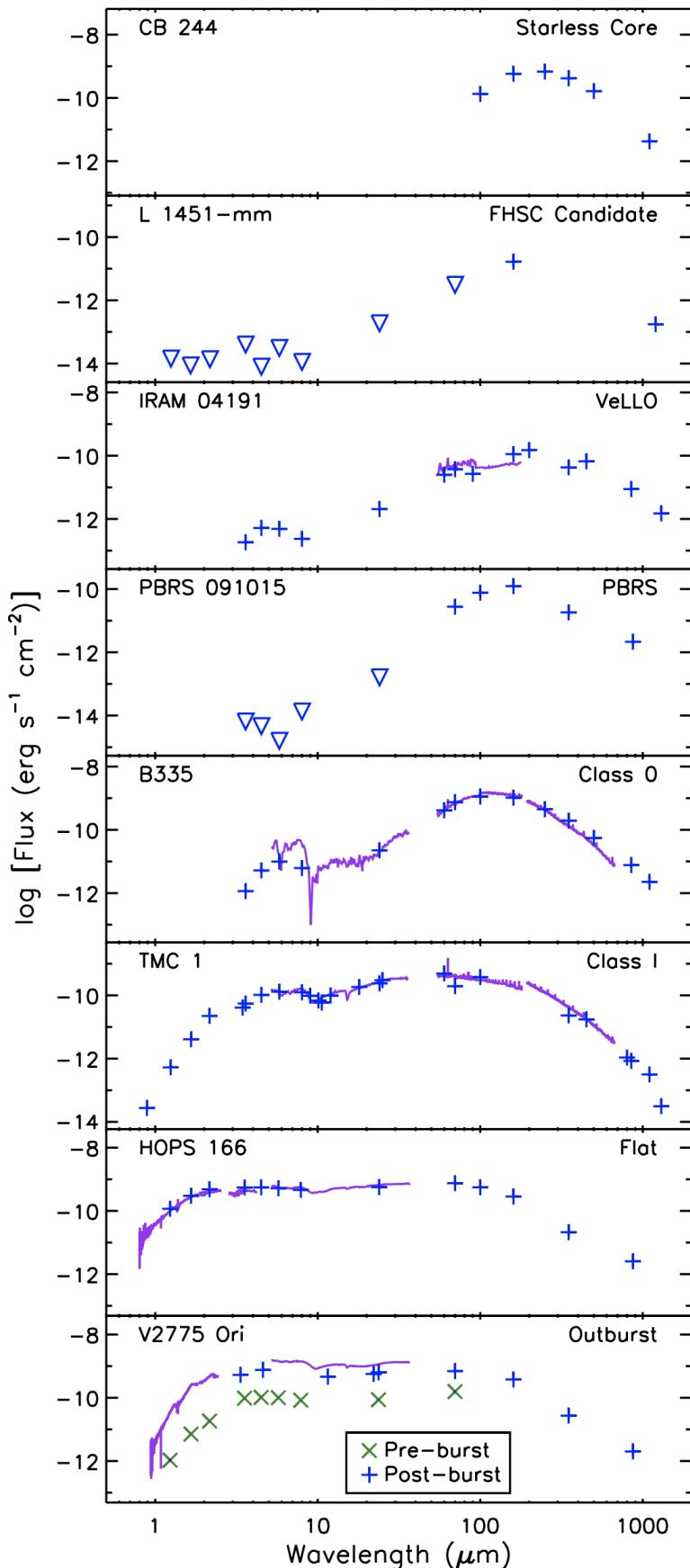


Figure 23: Sample SEDs of protostellar cores, together with the assigned class, as collected by Dunham et al. (2014). Figure taken from Krumholz (2015).

through the mid-infrared. If there is any IR excess at all, it is in the very far IR, indicating that the

emitting circumstellar material is cool and located far from the star. The idea here is that the disk around them has begun to dissipate, and is either now optically thin at IR wavelengths or completely dissipated, so there is no strong IR excess.

However, these stars are still not mature MS stars. First of all, their temperatures and luminosities do not correspond to those of MS stars. Instead, they are still puffed up to larger radii, so they tend to have either lower effective temperatures or higher bolometric luminosities (or both) than MS stars of the same mass. Second, they show extremely high levels of magnetic activity compared to MS stars, producing high levels of x-ray emission. Third, they show lithium absorption lines in their atmospheres. This is significant because lithium is easily destroyed by nuclear reactions at high temperatures, and no main sequence stars with convective photospheres show Li absorption. Young stars show it only because there has not yet been time for all the Li to burn.

To summarize,

- **Class 0** protostars are objects that are so heavily obscured that their spectra peak at  $\lambda > 100 \mu\text{m}$ . For these sources,  $\alpha_{\text{IR}}$  is not a useful characteristic, because the source may be invisible at  $2.2 \mu\text{m}$ , and at  $10 \mu\text{m}$  there may be deep absorption by cold foreground silicate dust. Inward motions of the gas are sometimes revealed by asymmetric profiles of molecular emission lines. The lifetime of a Class 0 object is short,  $\sim 1 - 3 \times 10^4 \text{ yr}$ .
- **Class I** protostars have  $\alpha_{\text{IR}} > 0$ : there is more power being radiated near  $10 \mu\text{m}$  than near  $2 \mu\text{m}$ . Blackbodies with  $T < 870 \text{ K}$  have  $\alpha_{\text{IR}} > 0$ . Class I protostars are thought to have typical ages  $\sim 1 - 2 \times 10^5 \text{ yr}$ .
- **Class II** YSOs have  $-1.5 < \alpha_{\text{IR}} < 0$ . Blackbodies with  $870 < T < 1540 \text{ K}$  have such IR spectral indices. Class II YSOs correspond to classical T Tauri stars, which are pre-main-sequence stars, still undergoing gravitational contraction, with substantial accretion disks and accretion rates  $\sim 10^{-6} M_\odot \text{ yr}^{-1}$ .
- **Class III** protostars have  $\alpha_{\text{IR}} < -1.6$ . Blackbodies with  $T > 1540 \text{ K}$  have such spectral indices. Class III YSOs correspond to **weak-lined T Tauri stars**, which are pre-main-sequence stars still undergoing gravitational contraction, but where the accretion disk is either weak or perhaps entirely absent.

The observed flux ratios are determined by the temperature of the emitting regions (star and inner disk) and by wavelength-dependent extinction. Classification of a given object as either Class I or Class II may depend on the source orientation. If we are viewing the object face-on it, may be classified as Class II, but an identical disk viewed edge-on could heavily redden the light reaching the observer, so that the object could be classified as Class I or even Class 0.

**Infall signatures:** The H $\alpha$  is particularly interesting, because it tells us something about the star's immediate environment. For MS stars, the H $\alpha$  line profile is a result of absorption at the stellar photosphere and emission from the chromosphere. At the photosphere there is a population of neutral hydrogen atoms in the  $n = 2$  level that absorbs photons at H $\alpha$  frequencies, producing absorption. Above that in the chromosphere is an optically thin, hot gas, which contains atoms in the  $n = 3$  level. Some of these emit H $\alpha$  photons, partially filling in the absorption trough, but leaving the line overall in absorption.

Producing H $\alpha$  in emission is tricky, however. The emitting material must be above the stellar photosphere, so it can fill in the absorption trough created there. This gas must be at temperatures of  $5,000 - 10,000 \text{ K}$  to significantly populate the  $n = 3$  level. However, in order to produce enough H $\alpha$  photons to fill in the trough and produce net emission, this gas must also be dense enough for the collision rate to be high enough to force the  $n = 3$  level close to LTE.

Ordinary stellar chromospheres have densities that are much too low to meet this requirement. This H $\alpha$  emission implies the presence of material around the star at temperatures of  $5,000 - 10,000 \text{ K}$ , but at densities much higher than found in an ordinary stellar chromosphere. Moreover, the width of the H $\alpha$  emission requires that this material be moving at velocities of hundreds of  $\text{km s}^{-1}$  relative to the stellar surface (i.e., comparable to the free-fall velocity). This cannot be thermal broadening, because this would require temperatures of  $\sim 10^6 \text{ K}$  high enough to completely ionize hydrogen. It must therefore be bulk motion.

The standard inference is that this indicates the presence of gas infalling onto the stellar surface. Such gas would provide the high densities required to produce H $\alpha$  in emission. The infall of this material would provide the requisite bulk motion. Internal shocks and the shocking of this gas against the stellar surface could easily heat gas to the required temperatures. Finally, this hot material would also produce continuum emission, explaining the continuum veiling.

Quantitative radiative transfer calculations that attempt to fit the observed veiling and line emission can be used to infer the densities and velocities of the circumstellar gas, thereby constraining the accretion

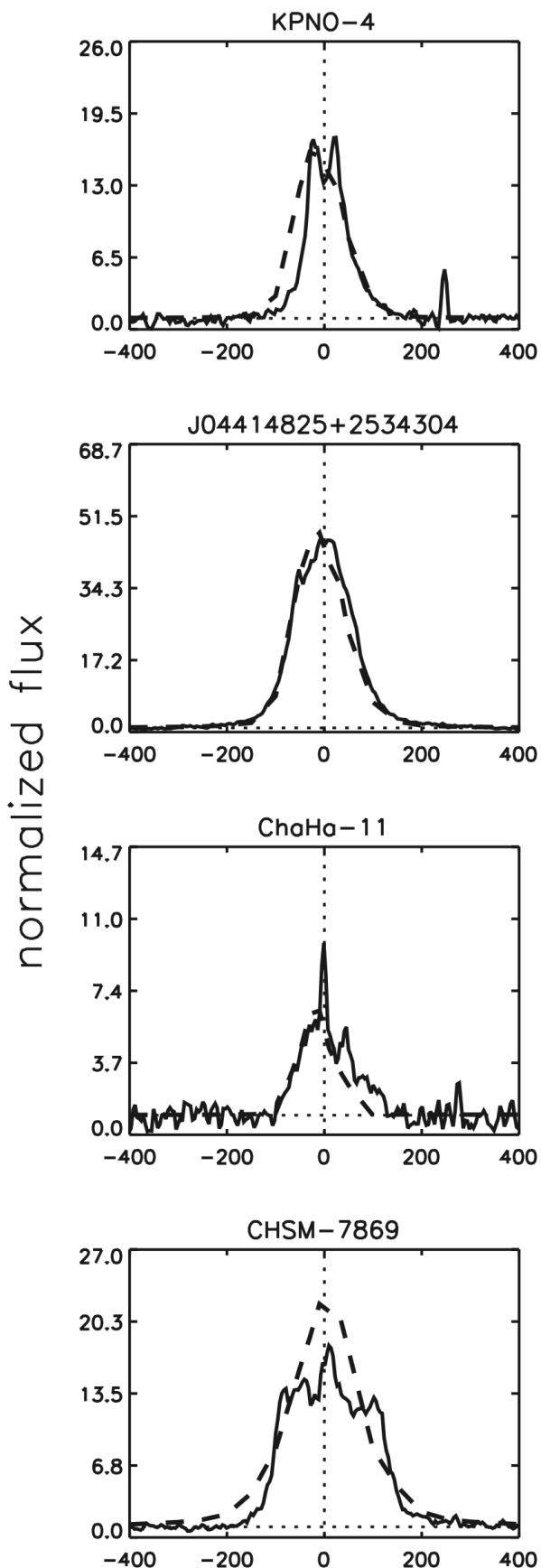


Figure 24: Comparisons between observed (solid) and model (dashed)  $\text{H}\alpha$  line profiles for a sample of T Tauri stars (Muzerolle et al., 2005). The  $x$  axis shows velocity in  $\text{km s}^{-1}$ . Each model curve is a fit in which the accretion rate is one of the free parameters. Figure taken from Krumholz (2015).

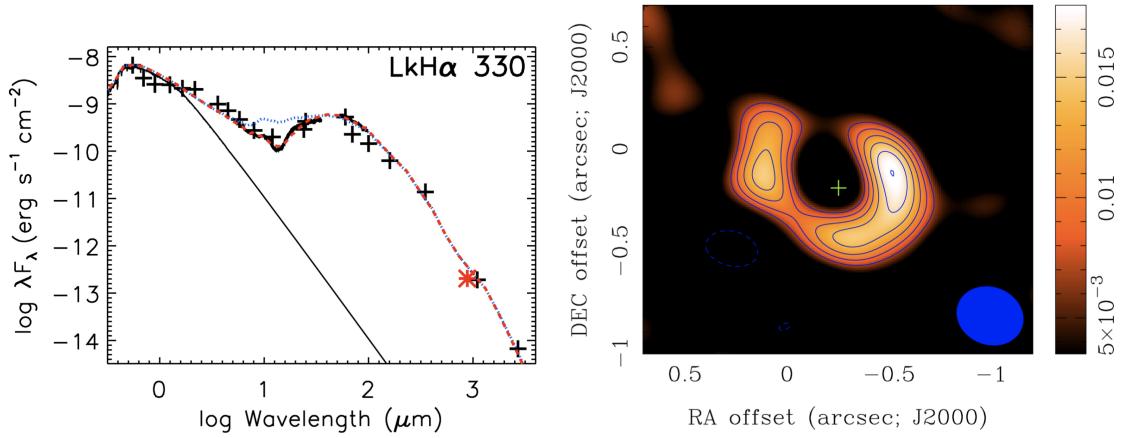


Figure 25: (**Left:**) The SED of the star LkH $\alpha$  330 (Brown et al., 2008). Plus signs indicate measurements. The black line is a model for a stellar photosphere. The blue line is a model for a star with a disk going all the way to the central star, while the red line is a model for disk with a 40 AU hole in its center. Figure taken from Krumholz (2015). (**Right:**) Dust continuum image of the disk around LkH $\alpha$  330, taken at 340 GHz by the SMA (Brown et al., 2008). Colors show the detected signal, and contours show the signal to noise ratio, starting from S/N of 3 and increasing by 1 thereafter. The green plus marks the location of the star. The blue circle is the SMA beam. Figure taken from Krumholz (2015).

rate (Figure 24). The inferred accretion rates depend on the strength of the H $\alpha$  emission, and are typically  $10^8 \text{ Myr}^{-1}$ . There is a broad range, however, running from  $10^{-11} - 10^{-6} \text{ Myr}^{-1}$ , with a very rough correlation  $\dot{M}_* \propto M_*^2$ .

These accretion rates are generally low enough so that accretion luminosity does not dominate over stellar surface emission. However, the estimated accretion rates are extremely uncertain, and the models used to make these estimates are very primitive. In general they simply assume that a uniform density slab of material arrives at the free-fall velocity, and covers some fraction of the stellar surface, and the accretion rate is inferred by determining the density of this material required to produce the observed spectral characteristics.

Despite this caveat, though, the H $\alpha$  line and other optical properties do seem to indicate that there must be some dense infalling material even around these stars that lack obvious envelopes. This in turn requires a reservoir of circumstellar material not in the form of an envelope, which is most naturally provided by a disk. Indeed, before the advent of space-based IR observatories, optical indicators like this were the only real evidence we had for disks around T Tauri and Herbig Ae/Be stars.

**Disk lifetimes:** We now turn to the disks that surround T Tauri and similar stars. Prior to the 2000s, we had very little direct information about such objects, since they are not visible in the optical. That changed dramatically with the launch of space-based IR observatories, and the development of ground-based millimeter interferometers. These new techniques made it possible to observe disks directly for the first time.

One of the most interesting properties of disks for those who are interested in planets is their lifetimes. This sets the limit on how long planets have to form in a disk before it is dispersed. In discussing disk lifetimes, it is important to be clear on how the presence or absence of a disk is to be inferred, since different techniques probe different parts and types of disks. In general what all these techniques have in common is that one uses some technique to survey young star clusters for disks. The clusters can be age-dated using pre-main sequence or main sequence HR diagrams. One then plots the disk fraction against age.

One signature of disks we have already discussed: optical line emission associated with accretion in T Tauri stars, particularly H $\alpha$ . Surveys of nearby groups find that H $\alpha$  line emission usually disappears at times between 1 and 10 Myr (Figure 19.4). This tells us that the inner parts of disks ( $\lesssim 1 \text{ AU}$ ) which feed stars disappear over this time scale. In contrast, ground-based near IR observations tell us about somewhat more distant parts of the disk, out to a few AU. The timescales implied by these results are very similar to those obtained from the H $\alpha$ : roughly half the systems lose their disks within  $\sim 3 \text{ Myr}$ .

These observations are sensitive primarily to the inner disk, and the IR techniques are generally sensitive only in cases where the dust in these regions is optically thick. Some optically thin material could still be present and would not have been detected. Observations at longer wavelengths, such as Spitzer's 24  $\mu\text{m}$  band and in the mm regime from ground-based radio telescopes, probe further out in disks, at distances of  $\sim 10 - 100 \text{ AU}$ . They are also sensitive to much lower amounts of material. Interestingly, unlike the shorter wavelength observations, these measurements indicate that a small but non-zero fraction of systems retain some disks out to times of  $\sim 10^8 \text{ yr}$ . The amounts of mass needed to explain the long

wavelength excess are only  $\sim 10^{-5} M_{\odot}$  in dust. Thus the older systems are likely looking at an even later evolutionary phase than T Tauri disks, one in which almost all the gas and inner disk material is gone. These are **debris disks**, which are thought to originate from collisions between larger bodies than to be made up of dust from interstellar gas.

**Transition disks:** The observation that accreting disks and inner optically thick disks disappear on a few Myr timescales, but that some fraction leave behind very small amounts of mass in the outer disk is a very interesting one. Observationally, we would like to get some constraints on how the process of a gaseous, accreting T Tauri disk transitioning to a low-mass debris disk occurs. For this purpose, there is an intriguing class of objects known as **transition disks**. Spectrally, these are defined as objects that have a significant  $24 \mu\text{m}$  excess (or excess at even longer wavelengths), but little or no IR excess (Figure 25, left). This SED suggests a natural physical picture: a disk with a hole in its center. The short wavelength emission normally comes from near the star, and the absence of material there produces the lack of short wavelength excess. Indeed, it is possible to fit the SEDs of some stars with models with holes.

In the last decade it has become possible to confirm the presence of inner holes in transition disks directly, at least cases where the inferred hole is sufficiently large (Figure 25, right). The sizes of the holes inferred by the observations are generally very good matches to the values inferred by modeling the SEDs. The holes are remarkably devoid of dust: upper limits on the masses of small dust grains within the hole are often at the level of  $\sim 10^{-6} M_{\odot}$ . The sharp edges of the holes indicate that the effect driving them isn't simply the growth of dust grains to larger sizes, which should produce a more gradual transition. Instead, something else is at work. However, in some transition disks gas is still seen within the gap in molecular line emission, which also suggests that whatever mechanism is removing the dust does not necessarily get rid of all the gas as well.

### 1.13.3 Follow-up Questions

- How does the spectrum change as planets start to form?

## **1.14 Question 14**

Sketch the spectral energy distribution (SED) of a T Tauri star surrounded by a protoplanetary disk. How would the SED change: (a) if the disk develops a large inner hole, (b) if the dust grains in the disk grow in size by agglomeration (with the same total mass)?

### **1.14.1 Short answer**

Answer.

### **1.14.2 Additional context**

Additional context.

## **1.15 Question 15**

What are the primary origins of the heat lost to space by infrared luminosity of Jupiter, Earth, and Io?

### **1.15.1 Short answer**

Answer.

### **1.15.2 Additional context**

Additional context.

## **1.16 Question 16**

Explain the observational problem of radius inflation for hot Jupiters and describe two possible solutions.

### **1.16.1 Short answer**

Answer.

### **1.16.2 Additional context**

Additional context.

## **1.17 Question 17**

Explain the effects of an atmosphere on a planet's surface temperature and the position of the habitable zone. What special considerations must one make for habitability around M-type stars?

### **1.17.1 Short answer**

Answer.

### **1.17.2 Additional context**

Additional context.

## **1.18 Question 18**

Explain the process of nuclear fusion and give two examples of important fusion processes that affect the lives of stars.

### **1.18.1 Short answer**

Answer.

### **1.18.2 Additional context**

Additional context.

## **1.19 Question 19**

What is Fermi's Paradox? Explain its logic and assess the current state of the Paradox in light of modern knowledge.

### **1.19.1 Short answer**

Answer.

### **1.19.2 Additional context**

Additional context.

### **1.19.3 Follow-up Questions**

- What percentage of stars have planets?

## **1.20 Question 20**

The so-called r- and s- processes are mechanisms that produce elements heavier than iron. Describe these mechanisms and evidence for them from abundance patterns. Where is the r-process thought to act?

### **1.20.1 Short answer**

Answer.

### **1.20.2 Additional context**

Additional context.

## 1.21 Resources

- Stellar Populations; Greggio & Renzini (2011)
- The Fundamentals of Stellar Astrophysics, Collins (2003)
- Stellar Structure and Evolution, Kippenhahn, Weigert & Weiss (2012)
- Evolution of Stars and Stellar Populations, Salaris & Cassisi (2005)
- The Astronomical Reach of Fundamental Physics, Burrows & Ostriker (2014)
- Opacity, Huebner & Barfield (2014)
- Radiative Processes in Astrophysics, Rybicky & Lightman (1979)
- Understanding Variable Stars, Percy (2007)
- A Library of Stellar Spectra; Jacoby & Hunter (1984)
- A Digital Spectral Classification Atlas; Gray (2009)
- Notes on Star Formation; Krumholz (2015)
- Physics of the Interstellar and Intergalactic Medium; Draine (2011)