Chapter 2

Theoretical background

While observations can reveal properties of individual stars and stellar populations more generally, the significance of these properties can only be gained by understanding the underlying physical processes and structure of the stars themselves. This chapter details the physical and mathematical framework of the processes which motivate and underpin this project.

2.1 Stellar parameters

To understand the significance of differences between stars with respect to interstellar extinction, we must first define the fundamental features of a stellar atmosphere, which will be used in this project as the input variables on which any star-to-star variations in extinction will be modelled.

The effective temperature (T_{eff}) of a star is defined as the thermodynamic temperature of a black body which produces the same stellar flux across all wavelengths (known as the bolometric flux) as that produced by the star. A black body represents the perfect thermal emitting object. The universal equation of the radiation emitted by a black body produces the body's flux per unit wavelength per unit angular viewing area, $F_{\lambda,bb}$, known as the black body's monochromatic flux. The equation, known as the Planck Law, is as follows:

$$F_{\lambda,bb} = \frac{2hc^2}{\lambda^5 \left(\exp\left(\frac{hc}{\lambda k_B T} \right) - 1 \right)} \tag{2.1}$$

where T is the thermodynamic temperature of the black body, h is Planck's constant, c is the vacuum speed of light and k_B is Boltzmann's constant. This equation also holds if the light wave frequency is used instead of the wavelength, with the monochromatic flux $F_{\nu,bb}$ now being the black body flux per unit frequency:

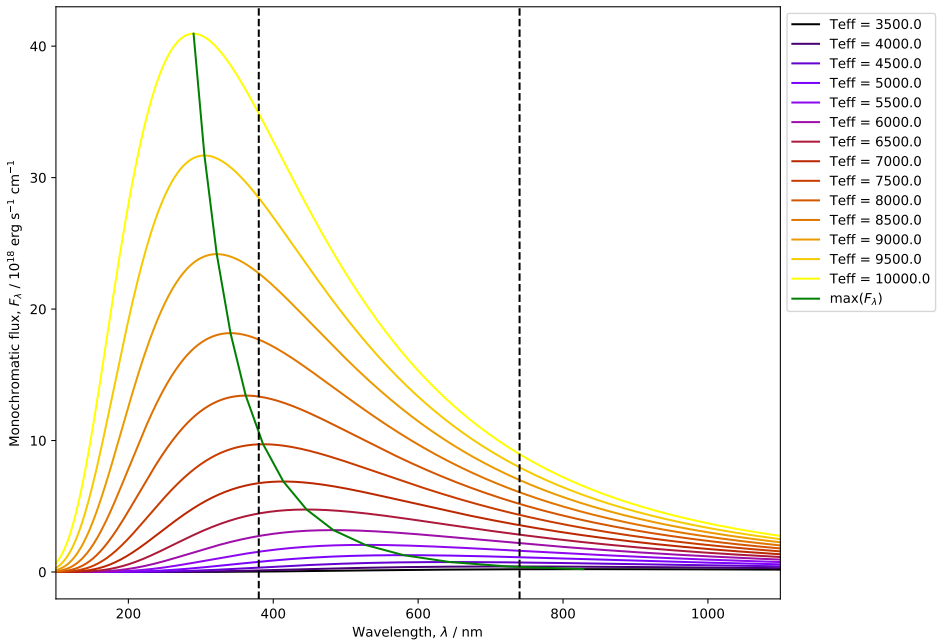

Figure 2.1: Monochromatic flux of a black body for different stellar effective temperatures. The black dashed lines mark the approximate limits of the visible part of the EM spectrum. The green curve represents the distributed of the maxima for the other curves.

$$F_{\nu,bb} = \frac{2h\nu^3}{c^2\left(\exp\left(\frac{h\nu}{k_B T}\right) - 1\right)} \tag{2.2}$$

In this project, the definition of monochromatic flux for any given object will be reserved exclusively for the flux per unit wavelength, F_λ, with any calculations involving black body fluxes using Equation 2.1.

The general approximation of stars to black bodies (and hence the actual stellar surface temperature to T_{eff}) is valid because all stars have been observed to have spectra that closely resemble those of black bodies, with the notable exception of atmospheric absorption lines.

$$L = 4\pi R^2 \sigma_{SB} T_{\text{eff}}^4 \tag{2.3}$$

Effective temperature has an effect on interstellar extinction due to its strong effect on the stellar luminosity and, hence, the flux. For a higher flux, more photons are likely to interact with the ISM, hence a higher extinction coefficient.

The metallicity of a star is defined as the fractional abundance of heavy elements, often approximated by iron (Fe) alone, relative to the star's hydrogen (H) abundance,

compared to that of the Sun. The abundances are determined by the strength of the elements' characteristic atomic absorption lines in the stellar spectra.

$$[\mathrm{Fe/H}] = \log\left(\frac{N_{\mathrm{Fe}}}{N_{\mathrm{H}}}\right) - \log\left(\frac{N_{\mathrm{Fe},\odot}}{N_{\mathrm{H},\odot}}\right) \tag{2.4}$$

For a generic atomic species E, N_E represents its number density. For stellar observations, N_E is measured at the surface. Since the output is logarithmic, a value of $[\mathrm{Fe/H}] = 0$ indicates solar metallicity. An increase in metallicity would cause the corresponding absorption lines to be stronger, thus reducing the observable flux. An increased metallicity also implies an increase in abundance of sub-ferrous metals. The presence of more nuclear species, each with unique absorption line configurations, inevitably creates more observable lines, further increasing the apparent extinction in the spectral flux.

The definition of the stellar surface gravity g is simply the value of the standard Newtonian gravitational acceleration, applied to the stellar surface (the mass is the total stellar mass, M_*, and the distance is the stellar radius, R_*):

$$g = \frac{GM_*}{R_*^2} \tag{2.5}$$

A greater surface gravity, as can be inferred from Equation 2.5, represents a surface with a higher mass density. For stars, being self-gravitating, this infers a higher atomic number density. The effects of surface gravity upon the emitted stellar spectrum arise directly from the effect of atomic number density on the quantum properties of the interactions between the photons and atomic electrons.

Ideally, all photons would have a single exact wavelength and energy, and all electrons in a given atomic occupancy level would only be able to absorb and emit such photons. However, Heisenberg's uncertainty principle implies both a fundamental uncertainty in a particle's velocity and a direct impact of this uncertainty on the uncertainty of the particle's momentum. It can be easily shown, particularly in the case of photons, that the pairing of the particle's energy (uncertainty ΔE) and time (Δt) share these same uncertainty properties. This results in the following relation:

$$\Delta E \Delta t \geq \hbar/2 \tag{2.6}$$

When a particle absorption a photon, the absorption process is not instantaneous and therefore carries an uncertainty in the time taken for the process to be completed, with a corresponding uncertainty in energy. Across a large number of absorptions for the same initial electron state, the result is a spread in the energies of the absorbed photons. The associated emission line is therefore broadened by the multiple wavelengths of the photons. This is universal and referred to as "natural broadening".

The impact of surface gravity arises via additional broadening effects upon these same absorption lines. When broadening effects are applied to an emission spectrum, such as a spectrum from a stellar surface, the result is that fewer photons pass through the surface, thereby reducing the surface flux seen by an outside observer.

Additional broadening effects are caused by processes whose impact is dependent on the separation between interacting atoms and ions and hence on the atomic/ionic number density. The broadening is due to the perturbations induced by these processes on the Hamiltonian quantum operator for the system (Struve, 1929), which produces shifts in the associated energy eigenvalues. The broadening impact of each process obeys a distance law for a particular integer value of the index p (Peach, 1984):

$$\Delta\lambda \propto r^{-p} \tag{2.7}$$

where p is a positive integer. A local environment of higher number density, in which atoms are more closely spaced, increases the frequency of interactions between atoms/ions and the likelihood of multiple processes contributing significantly to the total broadening effect.

The scenarios $p = 2$ and $p = 4$ are due to the linear and quadratic Stark effects, respectively. The Stark effect represents a perturbation in the quantum-mechanical potential of the electron due to the influence of an external electric field, F. The linear (first-order) Stark perturbation is proportional to F and dominates the broadening for weak fields, while the quadratic (second-order) effect is proportional to F^2 and dominates for stronger fields. For free electrons and ions with charge number Z, which can be treated as point sources of charge, $F \propto Z/r^2$, giving the stated p values for the linear and quadratic. This relation also gives the field strength dependence on number density, N: $F \propto N^{2/3}$.

For interactions between neutral atoms, the relevant processes are the standard van der Waals interaction ($p = 6$) and the so-called "resonance broadening" effect ($p = 3$), which is only valid for two atoms of the same type, with one of these atoms required to be in the ground state (Lortet & Roueff, 1969). Resonance broadening is weaker at higher densities due to the influence of multiple neighbouring atoms cancelling the individual perturbations (Kuhn, 1937).

Also, the rotation of the star has an impact on spectral line widths via a Doppler shift at each line-of-sight velocity. Together, the distribution of these velocities has the effect of broadening the absorption lines via a Gaussian distribution in wavelength, centred on the (Collins & Truax, 1995). The relationship to surface gravity arises from the gravity being the source of centripetal acceleration at the surface of a rotating star, and therefore being proportional in magnitude to the square of the surface rotational velocity.

2.2 Extinction definition

Extinction is defined using the standard astronomical system of flux magnitudes. In general, the difference between two flux measurements, f_1 and f_2, in magnitudes, is expressed as:

$$m_1 - m_2 = -2.5 \log \left(\frac{f_1}{f_2} \right) \tag{2.8}$$

where m_1 and m_2 are the magnitudes for f_1 and f_2, respectively. However, the flux of a source varies intrinsically with the distance travelled by the light to the observer (see Equation 1.1). To account for this, the distances to sufficiently close sources can be determined independently of their flux by measuring the sources' parallax from Earth.

The parallax p of an object is defined as the angular distance the object moves relative to the "fixed" background stars as the Earth moves a distance of 1AU (the average separation of the Sun and Earth during one complete orbit) perpendicular to the line-of-sight. This allows the distance, d, to be calculated using the geometry of triangles, combined with the small-angle approximation, as:

$$d/\mathrm{pc} = \frac{1}{(p/\mathrm{arcsec})} \tag{2.9}$$

For more distant stellar sources with much smaller parallaxes, a potential alternative is to view the object in a very long-wavelength filter, in which observations are likely to be impacted much less significantly by extinction, to estimate the type of star being observed, before using theoretical models to compare with observations at different wavelengths. This is more effective for brighter stars.

The role of distance gives each astronomical object two principal flux parameters. These are the apparent magnitude, m, and the absolute magnitude, M. The apparent magnitude is the flux magnitude of a source as observed by the telescope. The absolute magnitude is the predicted flux magnitude of the same source if it were to be placed at a fixed distance of 10 parsecs (pc) from the telescope with zero extinction. The absolute magnitude exists for the purpose of stellar classification, as the corresponding flux is simply a constant multiplied by the stellar luminosity (see Equation 1.1), which is an intrinsic physical property of the star.

However, to calculate the absolute magnitude of a source, we first require its distance and extinction coefficient, A. To account for extinction, it is necessary to define a new quantity, known as the intrinsic apparent magnitude, m_0. This is defined as the flux magnitude of a source corrected for losses due to extinction but not due to distance. In practice, it represents the apparent magnitude for light passing through a fully-transparent medium. The relation between m_0 and M, can be found by combining Equations 1.1 and 2.8:

$$m_0 - M = -2.5 \log \left(\left(\frac{10\text{pc}}{d} \right)^2 \right) = 5 \log \left(\frac{d}{\text{pc}} \right) - 5 \tag{2.10}$$

This quantity is known as the distance modulus. As seen in Equation 2.10, it varies only with distance. The quantity $m - M$, linking the initial observed data with the final theoretical data, is known as the apparent distance modulus and varies with both distance and extinction.

Therefore, the extinction coefficient A, defined as the flux lost due to the effects of the intervening line-of-sight medium, can be defined as:

$$A = m - m_0 \tag{2.11}$$

This fits with the physical definition of interstellar extinction given earlier, i.e. as the flux lost solely due to scattering and absorption in the interstellar medium.

The object of this project is to compare different extinction treatments and the subsequent effect on the interpretation of observational cluster datasets. For any observational set of stars, the stars' extinction coefficients will be completely unknown from the data alone. In order to compare observational and theoretical data, the most convenient approach is to add the (theoretical) extinction coefficient(s) to the theoretical dataset magnitudes (i.e., absolute magnitudes), before comparing to the distance-corrected observational data. As a result, the quantity from each dataset that is being compared is the absolute magnitude plus the extinction coefficient. If we label this quantity $M_{\text{ext},X}$ for a given filter, we can define it as:

$$\begin{aligned} M_{\text{ext},X} &= M_X + A_X \text{ (theoretical data)} \\ &= m - (m - M)_{X,0} \text{ (observational data)} \end{aligned} \tag{2.12}$$

However, as noted earlier, in astronomy it is not feasible to attempt observations by a single instrument at all wavelengths. Telescopes instead are purpose-built to study a single wavelength range within the full EM spectrum. Within this range, telescope observation ranges are further divided by filters or passbands, one of which is placed on their aperture at any given observation time. As shown in Figures 1.4-1.1, any single filter X has a limited range of wavelengths for which it is able to detect flux. It can also be seen in these figures that the transmittance of the filter changes as a function of wavelength. These instrumental factors must be considered, in addition to the observational challenges of distance and interstellar extinction, in order to correctly analyse observation telescope data.

If we compare the individual black body spectra in Figure 2.1, it can be seen that the maximum monochromatic flux of the black body occurs at an increasingly shorter wavelength for objects with increasingly higher temperatures. This makes the object

appear bluer to an observer. The relationship between the wavelength at which the monochromatic flux is maximal (λ_{max}) and the black body temperature is quantified by Wien's displacement law:

$$\lambda_{max} T = 2.898 \times 10^6 \text{ nm K} \tag{2.13}$$

More importantly, for two wavelength regions which are sufficiently far apart, the change in flux between the regions is always greater for stars with higher effective temperatures. Therefore, to measure a star's effective temperature, observers compare the star's observed flux in two filters operating at different wavelengths within the UV-IR wavelength range. The difference between the star's flux magnitudes in each of the two filters is then taken, with the flux in the redder filter being deducted from that of the bluer filter. This quantity is known as the colour index. For two filters X and Y, with X being bluer than Y, the colour index of observations made using those filters, $(X-Y)$, is defined as:

$$\begin{aligned}(X-Y) &= m_X - m_Y \\ &= (m_{X,0} - m_{Y,0}) + (A_X - A_Y) \\ &= (X-Y)_0 + E(X-Y)\end{aligned} \tag{2.14}$$

where $(X-Y)_0$ is the true or intrinsic colour index of the object and $E(X-Y) = A_X - A_Y$ is known as the colour excess, but can also be denoted in literature using the term "reddening". A major advantage of using the intrinsic colour index over absolute magnitudes is that it is completely independent of distance. The colour excess represents the effect of extinction on the observed colour index. Its importance arises from the prominence of the intrinsic colour index in determining effective temperature. Higher values of $(X-Y)$ indicate redder stars, with lower effective temperatures.

The most commonly-used colour index, employed as a reference for most optical observations, is the Johnson $(B-V)$ index. This is due to these filters being the among most long-lived, well-used and best-studied available, allowing for better comparisons of different data, including data from older archives.

Due to the potential confusion due to the use of the term "reddening" for both A_X and $E(X-Y)$ in literature, A_X will be referred to here as the "extinction" or "extinction coefficient" and $E(X-Y)$ as the "colour excess".

2.3 Bolometric corrections

All the equations in Section 2.2, including those for extinction, are not useful when applied to telescopes, as any filter will only detect a small fraction of the bolometric stellar flux that reaches the telescope. The missing information resulting from this observational constraint renders it difficult to determine the interstellar extinction. These

constraints much be mitigated before an accurate value of the extinction coefficient can be determined. This mitigation is carried out by employing bolometric corrections.

The use of bolometric corrections requires the detailed knowledge of stellar spectra least susceptible to significant extinction, i.e., nearby stars with high apparent fluxes. Only with complete knowledge of the spectrum from a reference star can the true spectrum of a distant star with unknown extinction be calculated. The spectra of these stars can be computed by using a grid of predicted fluxes from a stellar atmosphere model, the grid being composed of the stellar parameters known to change emission in stellar atmospheres. These are effective temperature, surface gravity and metallicity. For all filter systems studied in this project, the nearby bright star Vega (α Lyr) was used as the reference object. Using Vega as the reference star is the most well-known approach to photometric calibration (Casagrande & VandenBerg, 2014).

After accounting for a general extinction effect on an object's emission, its apparent magnitude in the wavelength range of a given filter X, defined as increasing from the shortest (λ_1) to the longest (λ_2) wavelength for which its response function is non-zero, can be calculated as:

$$m_X = -2.5 \log_{10} \left(\frac{\int_{\lambda_1}^{\lambda_2} f_\lambda \left(10^{-0.4 A_{X,\lambda}}\right) S_\lambda d\lambda}{\int_{\lambda_1}^{\lambda_2} f_\lambda^0 S_\lambda d\lambda} \right) + m_X^0 \tag{2.15}$$

where f_λ represents the (theoretical) monochromatic flux at a given wavelength λ at the observer distance from the source, $A_{X,\lambda}$ is the extinction coefficient in X as a function of wavelength and S_λ represents the filter response function of X. f_λ^0 and m_X^0 represent the monochromatic flux and apparent magnitude, respectively, in $X$of a known reference object, which is Vega in the case of this project.

To derive the equation linking a bolometric correction with the extinction parameter, we start with the definition of a bolometric correction in a filter X, which is denoted by BC_X:

$$BC_X \equiv M_{\text{bol}} - M_X \tag{2.16}$$

where M_X is the absolute magnitude of the object in X and M_{bol} is its (predicted) absolute bolometric magnitude, defined relative to the Sun using:

$$M_{\text{bol}} = M_{\text{bol},\odot} - 2.5 \log_{10} \left(\frac{4\pi R^2 F_{\text{bol}}}{L_\odot} \right) \tag{2.17}$$

where F_{bol} is the bolometric stellar flux at its surface, R is the stellar radius, $M_{\text{bol},\odot}$ is the solar absolute bolometric magnitude, which is assumed in this work to have a value of 4.75 and $L_\odot$ is the solar luminosity, for which a value of 3.844×10^{33} erg s^{-1} is used. Bolometric corrections can be expressed as a function of extinction using the universal definition of M_X in terms of m_X and the distance d to the source:

$$M_X = m_X - 2.5 \log_{10} \left(\left(\frac{d}{10\text{pc}} \right)^2 \right), \tag{2.18}$$

together with the equation $f_\lambda d^2 = F_\lambda R^2$, where F_λ is the monochromatic flux at λ at the stellar surface. This gives the final function for a bolometric correction for filter X:

$$\begin{aligned} BC_X = M_{\text{bol},\odot} - m_X^0 - 2.5 \log_{10} \left(\frac{4\pi R^2 F_{\text{bol}}}{L_\odot} \right) \\ + 2.5 \log_{10} \left(\frac{\int_{\lambda_1}^{\lambda_2} F_\lambda \left(10^{-0.4 A_{X,\lambda}} \right) S_\lambda d\lambda}{\int_{\lambda_1}^{\lambda_2} f_\lambda^0 S_\lambda d\lambda} \right) \end{aligned} \tag{2.19}$$

For a filter X, the extinction parameter $A_X = A_{X,\lambda}$ must be calibrated relative to a known value. In this work we will input a value of the extinction in the well-studied Johnson-V filter, A_V. To extract A_X, we use the simple relation:

$$A_{X,\lambda} = \left(\frac{A_{X,\lambda}}{A_V} \right) A_V \tag{2.20}$$

together with the chosen value of A_V (for this project the values were $A_V = 0$, 1 - note that $BC_X(A_V = 0)$ essentially assumes no extinction in any filter), before taking the difference between the two $BC_X(A_V)$ outputs, giving the following equation:

$$\begin{aligned} BC_X(0) - BC_X(A_V) &= 2.5 \log_{10} \left(\frac{\int_{\lambda_1}^{\lambda_2} F_\lambda S_\lambda d\lambda}{\int_{\lambda_1}^{\lambda_2} F_\lambda \left(10^{-0.4 \left(A_{X,\lambda}/A_V \right) A_V} \right) S_\lambda d\lambda} \right) \\ &\approx (A_X/A_V) \end{aligned} \tag{2.21}$$

As demonstrated in the equation above, any dependence of the A_X/A_V data on the Vega measurements or (as yet unknown or uncertain) bolometric quantities from Equation 2.19 is eliminated during the subtraction.

Most papers employing filters in the relevant spectral regions use the model of A_λ/A_V as described by Cardelli et al. (1989) or a modified version (O'Donnell (1994) and Fitzpatrick (1999) are among the most frequently used). This project uses the original model. The assumptions used to equate A_X/A_V with the result of Equation 2.21 are valid because the observational flux data used by Cardelli et al. (1989) when deriving their monochromatic extinction model were obtained using broad-band filters, and so are subject to the effects of variations inherent in both the real stellar spectra and the filter response functions, making the data heterochromatic (Casagrande & VandenBerg, 2014), in the same manner as the output data.

2.4 Forbes effect

The Forbes effect occurs as a broadband beam of light, such as that passes through an extended partially-transparent medium, such as a series of glass plates, the Earth's atmosphere or an interstellar gas cloud. It states that the greater the distance travelled by a light beam through the medium, the more penetrating the beam becomes (Forbes, 1842). If we use the case of the glass plates as an example, this means that the fraction of the light incident on the nth plate in the series which is dissipated by that plate from the original path is always greater than the corresponding fraction at the $(n+1)$th plate (Grebel & Roberts, 1995). The physical basis for the Forbes effect is that those photons in the original beam with wavelengths that make them the most likely to be absorbed or refracted are separated from the beam earlier. Therefore, as the beam travels through the medium, its constituent photons are progressively less likely to be separated. Since a higher fraction of its photons are retained as the distance through the medium increases, the beam is more penetrating (Ohvril et al., 1999).

The Forbes effect thus has an impact on the non-zero A_V value used in Equation 2.21 because if a uniform medium is assumed, as here, where R_V is held constant at the standard diffuse ISM value of 3.1 (Cardelli et al., 1989), a larger A_V value implies a longer path through the ISM, and thus a stronger Forbes effect. According to Girardi et al. (2008), any significant impact from the Forbes effect on the values of A_X/A_V occurs for a chosen $A_V \gtrsim 4$. They found that the effect was particularly strong for stars with $T_{\rm eff} \lesssim 3000$K and that, unsurprising, it became greater as the wavelength range covered by the filter response function increased. The choice of $A_V = 1$ made for this project should avoid serious problems from the Forbes effect, even for the widest filters.

2.5 Basics of stellar evolution

It can be seen from Equation 2.3 that the luminosity (and therefore flux) of a star is dependent on radius as well as effective temperature. If a plot is made of luminosity against effective temperature (known as a Hertzsprung-Russell diagram), it can be seen that all stars in a given population follow a single track. Because the stars are approximately the same age, this track is know as an isochrone. Isochrones for different population ages and metallicities are calculated using theoretical stellar models for the largest possible spread of initial stellar masses. An isochrone in the HR diagram has a number of distinct features:

- Most stars, including the coolest and least-luminous objects, follow a tight pattern of luminosity increasing with temperature. This is known as the main sequence (MS).
- This pattern stops as the luminosity continues to increase slowly but with decreasing temperature. The upper end of the MS is called the main-sequence

turn-off (MSTO), which is followed by the sub-giant branch (SGB).

- After the SGB, the gradient becomes much steeper, with temperate decreasing slowly and luminosity increasing rapidly. This is the red-giant branch (RGB).
- At the tip of the RGB, stars start becoming fainter and their temperatures increase. Eventually, their is a sequence of stars with near-constant luminosity but a range of effective temperatures. This is the horizontal branch (HB).
- After the horizontal branch, there is again a rapid increase in luminosity accompanied by a slow decrease in temperature. This is the asymptotic giant branch (AGB).

Stars with higher initial mass have shorter lifetimes for each evolutionary stage, due to their correspondingly higher temperatures, which increases the nuclear fusion efficiency. This causes higher surface luminosities. On the main sequence, therefore, more massive stars appear in the upper-left part and leave the MS earlier, which is the physical cause of the MSTO. As a population ages, the MSTO moves down the MS as progressively less-massive stars leave the MS.

On the main sequence, nuclear fusion occurs in the core and any products must be subjected to mixing effects to reach the atmosphere. Since stellar interiors are physically fluid, heavier nuclei, which are more dense, preferentially gather in regions close to the star's centre of gravity. Therefore, processes such as convection, thermohaline mixing and radiative levitation are required to induce a noticeable change in surface composition. However, these processes require certain physical conditions on local scales, if they are to be sustained for long enough to induce a visible change in the stellar spectrum.

On the main-sequence, only low-mass stars ($M_* \lesssim 0.38 M_\odot$) (Baraffe & Chabrier, 2018) have fully-convective interiors. Massive ($M_* \gtrsim 1.5 M_\odot$) stars have convective cores, but radiative envelopes. Intermediate-mass stars, including the Sun, have radiative cores and convective envelopes. Since regions in which energy transport is dominated by radiative effects are stable against convection (Salaris & Cassisi, 2017), stars with masses greater than $0.38 M_\odot$ are highly unlikely to show large changes in atmospheric composition along the main sequence. However, the fully-convective lower-mass stars are able to do so. Therefore, the faintest and reddest parts of the main sequence are most likely to have atmospheres enriched with metals, which would be confined in the cores of more massive stars. These enriched atmospheres are therefore highly sensitive to the overall metallicity of the original gas cloud in which the star, and its parent cluster if applicable, formed.

$$\nabla_{\mathrm{rad}} = \frac{3\kappa L P}{16\pi a c G m T^4} \tag{2.22}$$

To illustrate, let us consider a bubble of gaseous material in pressure-equilibrium with its surroundings and represent mixing as a significant change in the bubble's (radial) position on a significant time-scale, arising from small differences in the remaining 3 thermodynamic quantities between the bubble and its surroundings. For a non-rotating star, using a simple linear approach, together with the Archimedes principle, gives a set of 4 homogeneous differential equations for the (small) differences in P, T, μ and r (Equations (3.1)-(3.4) in Salaris & Cassisi (2017)). If Δx_i are the differences in the 4 parameters, taking the ansatz form $\Delta x_i = B_i e^{nt}$ allows for a solution as a 3rd-order polynomial in n (Equation (3.5) in Salaris & Cassisi (2017)), if the determinant of the relevant matrix (dependent of the values of the B_i) is zero. The Routh-Hurwitz stability criterion can then be applied to this polynomial to give a general solution for n. For a physically-unstable solution, the exponent in the Δx_i equation must be positive, i.e. n must satisfy the condition $\mathrm{Re}(n) > 0$. Hence, the subsequent constraints on the polynomial coefficients form all the possible conditions for instability, of which at least one must be satisfied. These constraints take the following form:

$$\nabla_\mu < 0 \tag{2.23}$$

$$\nabla_{\mathrm{rad}} > \nabla_{\mathrm{ad}} \tag{2.24}$$

$$\nabla_{\mathrm{rad}} > \nabla_{\mathrm{ad}} + \left(\frac{\phi}{\delta}\right) \nabla_\mu \tag{2.25}$$

where $\nabla_\mu = d\ln\mu / d\ln P$, $\nabla_{\mathrm{rad}} = (\partial \ln T / \partial \ln P)_{\mathrm{rad}}$ and $\nabla_{\mathrm{ad}} = (\partial \ln T / \partial \ln P)_{\mathrm{ad}}$ are the temperature-pressure gradients for the local environment (dominated by radiation pressure) and the bubble (treated as an adiabatic ideal gas), respectively, $\phi = (\partial \ln \rho / \partial \ln \mu)_{P,T}$ and $\delta = -(\partial \ln \rho / \partial \ln T)_{P,\mu}$ (Kippenhahn et al., 1980).

In stellar evolution, the stages described above represent extended periods of time in which the stellar interior, described in theoretical models as a series of (quasi-)spherical shells, can be described as being in hydrostatic equilibrium, i.e., the gravitational pressure inwards on material in a given shell, due to the mass enclosed by the shell, is equal to the radiation pressure outwards generated by nuclear fusion occurring within the area enclosed by the shell, either in the core or in an inner shell. Giant stars are objects with luminosities significantly higher and surface gravities lower than stars on the main sequence (which generally have gravities in the range $4 \lesssim \log(g) \lesssim 5.5$).

2.6 Colour-magnitude diagrams

By examining the flux-magnitude equations from Section 2.2, it becomes clear that both the absolute filter magnitudes and the intrinsic color indices can be used, together with bolometric corrections, to calculate the axis parameters of the HR diagram from observational data. To determine the detailed properties of stellar populations,

all stars in an observational sample or star cluster are plotted together on a pair of axes known as a colour-magnitude diagram (CMD), which represents an observational analogue of the HR diagram. The absolute magnitude of stars in a given filter Z, M_Z, is on the y-axis, with the flux increasing (and the magnitude value decreasing) upwards. The intrinsic colour index of the stars in two filters X and Y, $(X-Y)_0$, is on the x-axis, with the values increasing (and stars becoming redder) to the right. In practice, due to the unknown extinction coefficients of the individual stars and the cluster as a whole, the axial parameters are $M_{\mathrm{ext},Z}$ and $(X-Y)$. Note that filter Z may be the same as either X or Y.

In practice, the universal general shape and position of stellar populations in the HR diagram and each observational CMD, particularly the position and shape of the main sequence, provides a highly useful tool for comparing stellar populations with unknown distances and extinction coefficients to known examples and to theoretical models. This is done by alignment of the respective main sequences in CMDs, particularly the upper main sequence, which contains the most luminous MS stars and is less sensitive to the (initially unknown) value of the cluster metallicity than the lower MS.