Class 2 Notes: Binary Stars, Stellar Populations, and the HR Diagram

In the first class we focused on what we can learn by measuring light from individual stars. However, if all we ever measured were single stars, it would be very difficult to come up with a good theory for stars work. Fortunately, there are a lot of stars in the sky, and that wealth of stars provides a wealth of data we can use to build models. This leads to the topic of our second class: what we can learn from groups of stars.

A quick notation note to start the class: anything with a subscript $\odot$ refers to the Sun. Thus $M_\odot = 1.99 \times 10^{33}$ g is the mass of the Sun, $L_\odot = 3.84 \times 10^{33}$ erg s^{-1} is the luminosity of the Sun, and $R_\odot = 6.96 \times 10^{10}$ cm is the radius of the Sun. These are convenient units of measure for stars, and we'll use them throughout the class.

I. Mass measurements using binaries

Thus far we have figured out how to measure stars' luminosities, temperatures, and chemical abundances. However, we have not yet discussed how to measure perhaps the most basic quantity of all: stars' masses.

This turns out to be surprisingly difficult – how do you measure the mass of an object sitting by itself in space? The answer turns out to be that you don't, but that you can measure the mass of objects that aren't sitting by themselves. That leads to our last topic for today: binary stars and their myriad uses.

Roughly 2/3 of stars in the Milky Way appear to be single stars, but the remaining 1/3 are members of multiple star systems, meaning that two or more stars are gravitationally bound together and orbit one another. Of these, binary systems, consisting of two stars are by far the most common. Binaries are important because they provide us with a method to measure stellar masses using Newton's laws alone.

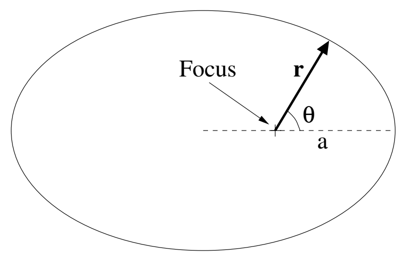
As a historical aside before diving into how we measure masses, binary stars are interesting as a topic in the history of science because they represent one of the earliest uses of statistical inference. The problem is that when we see two stars close to one another on the sky, there is no obvious way to tell if the two are physically near each other, or if it is simply a matter of two distant, unrelated stars that happen to be lie near the same line of sight. In other words, just because two stars have a small angular separation, it does not necessarily mean that they have a small physical separation.

However, in 1767 the British astronomer John Michell performed a statistical analysis of the distribution of stars on the sky, and showed that there are far more close pairs than one would expect if they were randomly distributed. Thus, while Michell could not infer that any particular pair of stars in the sky was definitely a physical binary, he showed that the majority of them must be.

A. Visual binaries

Binary star systems can be broken into two basic types, depending on how we discover them. The easier one to understand is visual binaries, which are pairs of stars which are far enough apart that we can see them as two distinct stars in a telescope.

We can measure the mass of a visual binary system using Kepler's laws. To see how this works, let's go through a brief recap of the two-body problem. Consider two stars of masses M_1 and M_2. We let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the vectors describing the positions of stars 1 and 2, and $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ be the vector distance between them. If we set up our coordinate system so that the center of mass is at the origin, then we know that $M_1\mathbf{r}_1 + M_2\mathbf{r}_2 = 0$. We define the reduced mass as $\mu = M_1 M_2/(M_1 + M_2)$, so $\mathbf{r}_1 = -(\mu/M_1)\mathbf{r}$ and $\mathbf{r}_2 = (\mu/M_2)\mathbf{r}$.



The solution to the problem is that, when the two stars are at an angle θ in their orbit, the distance between them is

$$r = \frac{a(1-e^2)}{1+e\cos\theta},$$

where the semi-major axis a and eccentricity e are determined by the stars' energy and angular momentum. Clearly the minimum separation occurs when $\theta = 0$ and the denominator has its largest value, and the maximum occurs when $\theta = \pi$ and the denominator takes its minimum value. The semi-major axis is the half the sum of this minimum and maximum:

$$\frac{1}{2}[r(0) + r(\pi)] = \frac{1}{2}\left[\frac{a(1-e^2)}{1+e} + \frac{a(1-e^2)}{1-e}\right] = a$$

The orbital period is related to a by

$$P^2 = 4\pi^2 \frac{a^3}{GM},$$

where $M = M_1 + M_2$ is the total mass of the two objects.

This describes how the separation between the two stars changes, but we instead want to look at how the two stars themselves move. The distance from each of the two stars to the center of mass is given by

$$r_1 = \frac{\mu}{M_1} r = \frac{\mu}{M_1}\left[\frac{a(1-e^2)}{1+e\cos\theta}\right] \qquad r_2 = \frac{\mu}{M_2} r = \frac{\mu}{M_2}\left[\frac{a(1-e^2)}{1+e\cos\theta}\right].$$

Again, these clearly reach minimum and maximum values at $\theta = 0$ and $\theta = \pi$, and the semi-major axes of the two ellipses describing the orbits of each star are given by half the sum of the minimum and maximum:

$$a_1 = \frac{1}{2}[r_1(0) + r_1(\pi)] = \frac{\mu}{M_1} a \qquad\qquad a_2 = \frac{1}{2}[r_2(0) + r_2(\pi)] = \frac{\mu}{M_2} a.$$

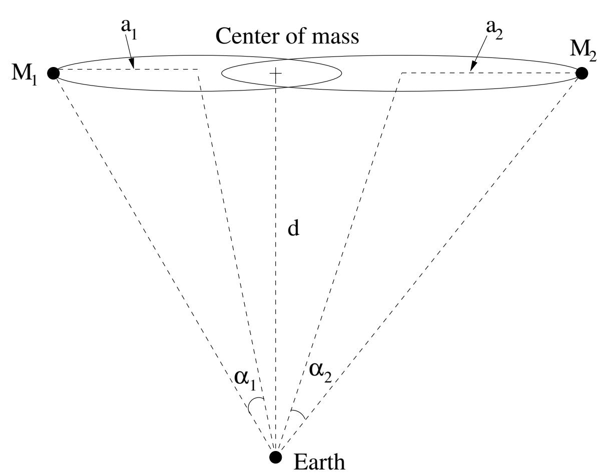
Note that it immediately follows that $a = a_1 + a_2$, since $\mu/M_1 + \mu/M_2 = 1$.

We can measure the mass of a visual binary using Kepler's laws. To remind you, there are three laws: first, orbits are ellipses with the center of mass of the system at one focus. Second, as the bodies orbit, the line connecting them sweeps out equal areas in equal times – this is equivalent to conservation of angular momentum. Third, the period P of the orbit is related to its semi-major axis a by

$$P^2 = 4\pi^2 \frac{a^3}{GM},$$

where M is the total mass of the two objects.

With that background out of the way, let's think about what we can actually observe. We'll start with the simplest case, where the orbits of the binary lie in the plane of the sky, the system is close enough that we can use parallax to measure its distance, and the orbital period is short enough that we can watch the system go through a complete orbit. In this case we can directly measure four quantities, which in turn tell us everything we want to know: the orbital period P, the angles subtended by the semi-major axes of the two stars orbits, α_1, and α_2, and the distance of the system, d.



The first thing to notice is that we can immediately infer the two stars' mass ratio just from the sizes of their orbits. The semi-major axes of the orbits are $a_1 = \alpha_1 d$ and $a_2 = \alpha_2 d$. We know that $M_1 r_1 \propto M_2 r_2$, and since $a_1 \propto r_1$ and $a_2 \propto r_2$ by the argument we just went through, we also know that $M_1 a_1 \propto M_2 a_2$. Thus it immediately follows that

$$\frac{M_1}{M_2} = \frac{a_2}{a_1} = \frac{\alpha_2}{\alpha_1}.$$

Note that this means we can get the mass ratio even if we don't know the distance, just from the ratio of the angular sizes of the orbits.

Similarly, we can infer the total mass from the observed semi-major axes and period using Kepler's 3rd law:

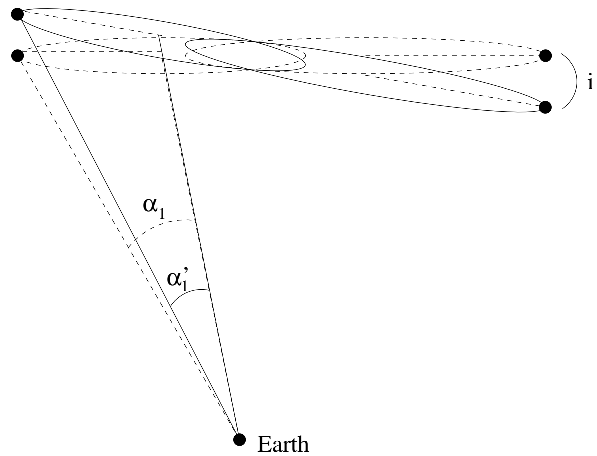
$$M = 4\pi^2 \frac{(a_1 + a_2)^3}{GP^2} = 4\pi^2 \frac{(\alpha_1 d + \alpha_2 d)^3}{GP^2},$$

where everything on the right hand side is something we can observe. Given the mass ratio and the total mass, it is of course easy to figure out the masses of the individual stars. If we substitute in and write everything in terms of observables, we end up with

$$M_1 = 4\pi^2 d^3 \frac{(\alpha_1 + \alpha_2)^2}{GP^2} \alpha_2 \qquad\qquad M_2 = 4\pi^2 d^3 \frac{(\alpha_1 + \alpha_2)^2}{GP^2} \alpha_1.$$

This is the simplest case where we see a full orbit, but in fact we don't have to wait that long – which is a good thing, because for many visual binaries the orbital period is much longer than a human lifetime! Even if we see only part of an orbit, we can make a very similar argument. All we need is to see enough of the orbit that we can draw an ellipse through it, and then we measure α_1 and α_2 for the inferred ellipse. Similarly, Kepler's second law tells us that the line connecting the two stars sweeps out equal areas in equal times, so we can infer the full orbital period just by measuring what fraction of the orbit's area has been swept out during the time we have observed the system.

The final complication to worry about is that we don't know that the orbital plane lies entirely perpendicular to our line of sight. In fact, we would have to be pretty lucky for this to be the case. In general we do not know the inclination of the orbital plane relative to our line of sight. For this reason, we do not know the angular sizes α_1' and α_2' that we measure for the orbits are different than what we would measure if the system were perfectly in the plane of the sky. A little geometry should immediately convince you that that $\alpha_1' = \alpha_1 \cos i$, where i is the inclination, and by convention $i = 0$ corresponds to an orbit that is exactly face-on and $i = 90°$ to one that is exactly edge-on. The same goes for α_2. (I have simplified a bit here and assumed that the tilt is along the minor axes of the orbits, but the same general principles work for any orientation of the tilt.)



This doesn't affect our estimate of the mass ratios, since $\alpha_1/\alpha_2 = \alpha_1'/\alpha_2'$, but it does affect our estimate of the total mass, because $a \propto \alpha_1$. Thus if we want to write out the total mass of a system with an inclined orbit, we have

$$M = 4\pi^2 \frac{(\alpha_1' d + \alpha_2' d)^3}{G P^2 \cos^3 i}.$$

We get stuck with a factor $\cos^3 i$ in the denominator, which means that instead of measuring the mass, we only measure a lower limit on it.

Physically, this is easy to understand: if we hold the orbital period fixed, since we can measure that regardless of the angle, there is a very simple relationship between the stars' total mass and the size of their orbit: a bigger orbit corresponds to more massive stars. However, because we might be seeing the ellipses at an angle, we might have underestimated their sizes, which corresponds to having underestimated their masses.

B. Spectroscopic binaries

A second type of binary is called a spectroscopic binary. As we mentioned earlier, by measuring the spectrum of a star, we learn a great deal about it. One thing we learn is its velocity along our line of sight – that is because motion along the line of sight produces a Doppler shift, which displaces the spectrum toward the red or the blue, depending on whether the star is moving away from us or toward us.

However, we know the absolute wavelengths that certain lines have based on laboratory experiments on Earth – for example the Hα line, which is produced by hydrogen atoms jumping between the 2nd and 3rd energy states, has a wavelength of exactly 6562.8 Å. (I could add several more significant digits to that figure.) If we see the Hα line at 6700 Å instead, we know that the star must be moving away from us.

The upshot of this is that you can use a spectrum to measure a star's velocity. In a binary system, you will see the velocities change over time as the two stars orbit one another. On your homework, you will show how it is possible to use these observations to measure the masses of the two stars.

The best case of all is when a pair of stars is observed as both a spectroscopic and a visual binary, because in that case you can figure out the masses and inclinations without needing to know the distance. In fact, it's even better than that: you can actually calculate the distance from Newton's laws!

Unfortunately, very few star systems are both spectroscopic and visual binaries. That is because the two stars have to be pretty far apart for it to be possible to see both of them with a telescope, rather than seeing them as one point of light. However, if the two stars are far apart, they will also have relatively slow orbits with relatively low velocities. These tend to produce Doppler shifts that are too small to measure. Only for a few systems where the geometry is favorable and where the system is fairly close by can we detect a binary both spectroscopically and visually. These systems are very precious, however, because then we can measure everything about them. In particular, for our purposes, we can measure their masses absolutely, with no uncertainties due to inclination or distance.

II. The HR Diagram

We've already learned a tremendous amount just by looking at pairs of stars, but we can learn even more by looking at larger populations. The most basic and important tool we have for studying stellar populations is the Hertzsprung-Russell diagram, or HR diagram for short. Not surprisingly, this diagram was first made by Hertzspring and Russell, who actually created it independently, in 1911 for Ejnar Hertzsprung and in 1913 for Henry Norris Russell.

A. The Observer's HR Diagram

The HR diagram is an extremely simple plot. We simply find a bunch of stars, measure their luminosities and surface temperatures, and make a scatter plot of one against the other. Since surface temperatures and absolute luminosities are often expensive to measure in practice, more often we plot close proxies to them. In place of total luminosity we put the luminosity (or magnitude) as seen in some particular range of wavelengths, and in place of surface temperature we plot the ratio of the brightness seen through two different filters – this is a proxy for color, and thus for surface temperature. For this reason, we sometimes also refer to these diagrams as color-magnitude diagrams, or CMDs for short.

An important point to make is that it is only possible to make an HR diagram for stars whose distances are known, since otherwise we don't have a way of measuring their luminosities. The largest collection of stars ever placed on an HR diagram comes from the Hipparcos catalog, which we discussed last time – a collection of stars near the Sun whose distances have been measured by parallax.

[Slide 1 – Hipparcos HR diagram]

This particular HR diagram plots visual magnitude against color. A note on interpreting the axes: this HR diagram was made using two different filters: B and V. Here B stands for blue, and it is a filter that allows preferentially blue light to pass. V stands for visible, and it is a filter that allows essentially all visible light to pass. On the y axis in this plot is absolute visual magnitude is basically the logarithm of the luminosity in the V band, which is close to the total luminosity because stars put out most of their power at visible wavelengths. The approximate conversion between mass and luminosity is shown on the right y axis. Note that, since the magnitude scale is backward and higher magnitude corresponds to lower luminosity, the scale on the left is reversed – magnitude decreases upward, so that brighter stars are near the top, as you would intuitively expect.

On the x axis is B magnitude minus V magnitude. Since magnitudes are a logarithmic scale, $B - V$ is a measure of the ratio of the star's luminosity in blue light to its total luminosity. Since magnitudes go in the opposite of the sensible direction (bigger numbers are dimmer), a high value of $B - V$ corresponds to a small ratio of blue luminosity to total luminosity. A low value of $B - V$ is the opposite. Thus moving to the right on this diagram corresponds to getting redder, and moving left corresponds to getting bluer. The value of $B - V$ corresponds to an approximate surface temperature, which is indicated on the top x axis.

The first thing you notice about the diagram is that the stars don't fall anything like randomly on it. The great majority of them fall along a single fat line, which we call the main sequence. The Sun sits right in the middle of it. The main sequence extends from stars that are bright and blue to stars that are dim and red. It covers an enormous range of absolute luminosities, from 10^{-3} $L_\odot$ to 10^3 $L_\odot$ – and that's just for nearby stars. The range is larger if we include more stars, because very bright stars are rare, and this particular HR diagram doesn't have any of the brightest ones on it.

You can also see a second prominent population, extending like a branch of the main sequence. These are stars that are red but bright. For reasons we'll see in a moment, this means they must have very large radii, and so they are called red giants.

The HR diagram we've been looking at is limited to relatively bright and nearby stars. One can extend it by adding in some observations to measure dimmer stars, as well as a selection of other more exotic stars. This HR diagram includes 22,000 stars from Hipparcos supplemented by 1,000 stars from the Gliese catalog.

[Slide 2 – extended Hipparcos HR diagram]

As before, we see that the most prominent feature is the main sequence, and the second most prominent is a branch extending out of it consistent of bright, red stars. However, we can also see some other populations start to emerge. First, there is a collection of stars that run from medium color to blue, but that are very dim – a factor of ~ 1000 dimmer than the Sun. These too fall along a rough line.

As we will see in a moment, the combination of high surface temperature and low luminosity implies that these stars must have very small radii. For this reason, we call them dwarfs. Because these stars have fairly high surface temperatures, their colors are whitish-blue. Thus, these stars are called white dwarfs.

One can also get glimpses of other types of stars in other parts of the diagram, which don't fall on either the main sequence, the red giant branch, or the white dwarf sequence. Almost all of these stars are very bright, and lie above the main sequence. These other types of stars must be very rare, since even a catalog containing 23,000 stars includes only a handful of them. We'll discuss these more exotic types of stars when we get to stellar evolution in the second half of the course.

It's worth stopping for a moment to realize that the HR diagram is rather surprising. Why should it be that stars do not occupy the full range of luminosities and temperatures continuously, and instead seem to cluster into distinct groups? Explaining the existence of these distinct groups where stars live, and their relationship to one another, is the single big theoretical problem from stellar physics. Our goal at the end of this class is to understand why the HR diagram looks like this, and what it means.

Similarly, the fact that the main sequence is a single curve means that all stars of a given luminosity have about the same color, and vice-versa. We would like to understand why that is.

B. HR Diagrams of Clusters

So far we've been looking at HR diagrams of stars that happen to be near the Sun, which are a hodgepodge of stars of many different masses, ages, and abundances of heavy elements. However, it is very instructive to instead look at the HR diagram for more homogenous populations. How do we pick a homogenous population? Fortunately, nature has provided for us. Some stars are found in clusters that are held together by the stars mutual gravity.

[Slide 3 – globular cluster M80]

The slide shows an example of a star cluster: a globular cluster known as M80, which contains several hundred thousand stars. The stars in a cluster like M80 generally formed in a single burst of star formation, and as a result they are very close to one another in age and chemical abundance. The stars in a cluster are also of course all at about the same distance from us, which means that we can compare their relative brightnesses even if we don't know exactly how far away the cluster is.

[Slide 4 – HR diagram for NGC 6397]

The slide shows the HR diagram for the star clusters NGC 6397. The axis labels here correspond to the filters used on the Hubble telescope, but the idea is the same as before: the y axis shows the magnitude in a single filter using a reverse axis scale, so it measures luminosity, with up corresponding to higher luminosity. The x axis shows the ratio of the luminosities in two different color filters, with the orientation chosen so that red is to the right and blue is to the left.

So what's different about this HR diagram as opposed to the ones we looked at earlier for nearby stars? Just like in the previous picture we see that most stars fall along a single curve – the main sequence – and that there is a secondary curve at lower luminosity and bluer color – the white dwarf sequence. In comparison to the other HR diagram, however, this main sequence is much narrower. Instead of a fat line we have a very thin one. In fact, it's even thinner than it appears in the diagram – many of the points that lie off the main sequence turn out to be binary stars that are so close to one another that the telescope couldn't separate them, and thus treated them as a single star.

The thinness of this main sequence suggests that the spread we saw in the solar neighborhood main sequence must be due to factors that are absent in the star cluster: a spread in stellar ages and a spread in chemical composition. The fact that all the stars fall along a single thin line is compelling evidence that there is a single intrinsic property of stars that varies as we move along the main sequence, and is responsible for determining a star's luminosity and temperature. The age and chemical composition can alter this slightly, causing the thin sequence to puff up a little, but basically there's one number that is going to determine everything about a star's properties. The natural candidate, of course, is the star's mass.

If you're very sharp you may have noticed something else slightly different about this HR diagram, compared to the one we looked at for the solar neighborhood. You may have noticed that, in the solar neighborhood, the white dwarf sequence does not go as far toward the blue as the bluest part of the main sequence. In this HR diagram, on the other hand, the white dwarfs go further toward the blue. This is not an accident, as becomes clear when we compare HR diagrams for different star clusters.

[Slide 5 – HR diagrams for M 67 and NCG 188]

The slide shows HR diagrams for two different clusters: M 67 and NGC 188. These are both a type of cluster called on open cluster. The scatter is mostly an observational artifact: it's due to stars that happen to be along the same line of sight as the star cluster, but aren't really members of it, and are at quite different distances. Those have been cleaned out of the HR diagram for NGC 6397, but not for these clusters.

The interesting thing to notice is the difference between the two main sequences. It seems that the dim, red end of the main sequence is about the same from one cluster to another, but the bright, blue side ends at different points in different clusters. Where the main sequence ends, it turns upward and becomes the red giant branch, and the red giant branch is at two different places in the two clusters. The place where the main sequence ends and bends upward into the red giant branch is called the main sequence turn-off.

What is going on here? What's the difference between these two clusters? The answer turns out to be their ages. M 67 formed somewhat more recently than NGC 188. If we repeat this exercise for clusters of different ages, we see that this is a general trend. In younger clusters the main sequence turn off is more toward the bright, blue side of the sequence, and it vanishes completely in the youngest clusters. In older clusters it moves to lower luminosity and redder color. That's the reason that the white dwarf sequence extended further to the blue than the main sequence in the globular cluster NGC 6397 – globular clusters are very, very old.

Thus, we have another clue: bright, blue main sequence stars disappear at a certain age. The brighter and bluer the star, the shorter the time for which it can be found. This too is something that our theory needs to be able to explain.

C. Stellar Radii on the HR Diagram

So far we have talked only about quantities we can observe directly – the observational HR diagram. Now let's see what we can infer based on our knowledge of physics. The first and most obvious thing to do with an HR diagram is to see what it tells us about stars' radii.

To remind you, radius, luminosity, and temperature are all related by the blackbody formula we wrote down last time:

$$L = 4\pi R^2 \sigma T^4$$

Since the HR diagram is a plot of temperature versus luminosity, at every point in the plot we can use the temperature and luminosity to solve for the corresponding radius:

$$R = \sqrt{\frac{L}{4\pi\sigma T^4}}$$

Putting this in terms of some useful units, this is

$$\frac{R}{R_\odot} = 1.33 \left(\frac{L}{L_\odot}\right)^{1/2} \left(\frac{T}{5000\text{ K}}\right)^{-2} \qquad (1)$$

Just from this formula you can figure out what a line of constant radius will look like. We've been plotting things on a logarithmic scale, so our HR diagrams have $\log L$ on the y axis and $\log T$ on the x axis. If we take the logarithm of this equation, we get

$$\log L = 4 \log T + \log(4\pi\sigma) + 2 \log R.$$

If R is constant, then in the $(\log T, \log L)$ plane this is just a line with a slope of 4.

[Slide 6 – theorist's HR diagram]

Note that the line does have a slope of 4, but because of the annoying astronomy convention that we plot red, lower temperature, to the right, the lines appear to have a negative slope. Comparing these lines to the main sequence, you can see that the main sequence has a somewhat steeper slope than 4 in the $(\log T, \log L)$ plane. This means that stars at the low temperature, low luminosity end have smaller radii than stars at the high luminosity, high temperature end. The range in radii is a factor of ~ 100.

Although they're not shown on this particular diagram, you can immediately see why white dwarfs and red giants have the names they do. The red giant branch extends above and to the right of the main sequence, so it goes up to several hundred $R_\odot$. The white dwarf sequence is below the main sequence, so that it hovers around 0.01 $R_\odot$. To put these numbers in perspective, $100 R_\odot = 0.47$ AU, so a star with a radius just above $200 R_\odot$ would encompass the Earth's orbit. A radius of 0.01 $R_\odot$ corresponds to 10% more than the radius of the Earth. Thus the largest red giants would swallow the Earth, while the smallest white dwarfs are about the size of the Earth.

D. Stellar Masses on the HR Diagram

The radii of stars we can measure directly off the HR diagram, but the masses are a bit trickier, since we need additional information to obtain those. As we discussed earlier, we can only get independent measurements for the masses of stars if they are members of binary systems. Fortunately, astronomers have now compiled a fairly large list of binary systems within which we can measure masses, so we can plot mass against luminosity and color.

[Slide 7 – mass-luminosity and mass-effective temperature relations]

Note that this slide only contains main sequence stars, not red giants, white dwarfs, etc. Also note that, on the plot on the right, the axes are reverse, so more massive, brighter stars are to the left.

Here's one interesting thing to take away from this plot: the luminosity-mass relation is a line just like the main sequence. In other words, all main sequence stars of a given mass have the same luminosity, and, with some error, the same radius and surface temperature. Since we already saw that the main sequence is a single curve on the HR diagram, this plot tells us something critical: *for main sequence stars, the stellar mass determines where the star falls on the main sequence.*

This is a profound statement. It means that, for a main sequence star, if you know it's mass, then you know pretty much everything about it. the properties of a star are, to very good approximation, dictated solely by its mass. Nothing else matters much. In this class we will attempt to understand why this is.

Another interesting point is that over significant ranges in mass, the mass-luminosity relation is a straight line on a log-log plot. What sort of function produces a straight line on a log-log plot? The answer is a powerlaw. To see why, write down the equation of a line in the log-log plot:

$$\log L = p \log M + c,$$

where p is the slope of the line and c is the y-intercept. It immediately follows that

$$L = cM^p.$$

Thus the luminosity is proportional to some power of the mass, and the power is equal to the slope of the line on the log-log plot. If you actually measure the slope of the data, you find that for masses in the vicinity of 1 $M_\odot$, a slope of 3.5 is a reasonable fit.