

ASTR 323/423 HW #2

1. Luminosity Functions

The Schechter function gives the number of galaxies per unit luminosity, $\Phi(L)dL$. Rewrite this using the conversion between absolute magnitude and luminosity to give the number of galaxies per unit magnitude, $\Phi(M)dM$. Start with $\Phi(L)$, and replace L with M using the magnitude-luminosity equation. Then take the derivative of M with respect to L to get dM/dL so that you can replace dL with dM . Once you've worked out $\Phi(M)dM$, derive the slope of the faint end when luminosity functions are plotted logarithmically (in other words, derive $d(\log \Phi(M))/dM$) for the faint magnitudes) and show mathematically that in such a plot $\alpha = -1$ corresponds to a flat faint end slope.

Next, we'll construct a luminosity function for galaxies and actually measure this faint end slope. Go back to the Sloan Digital Sky Survey SQL [search page](#) and run the following SQL search (I will explain this syntax in more detail in class), downloading the file in CSV format to your computer:

```
SELECT TOP 100000 P.objID, P.flags_r,
  P.ra, P.dec, P.dered_g, P.dered_r, P.dered_i,
  P.err_g, P.err_r, P.err_i,
  P.petroR50_g, P.petroR90_g,
  S.z, S.zErr, S.velDisp, S.velDisperr,
  E.oh_p50, E.lgm_tot_p50, E.sfr_tot_p50
FROM Galaxy as P
  JOIN SpecObj as S on P.objID = S.BestObjID
  JOIN galSpecExtra as E on S.SpecObjID = E.SpecObjID
WHERE S.z>0.00001 AND S.z<0.3
  AND P.dered_r<17.5
  AND ((P.flags_r & 0x10000000) != 0)
  AND ((P.flags_r & 0x8100000c00a0) = 0)
  AND (((P.flags_r & 0x400000000000) = 0) or (P.err_r <= 0.3))
  AND (((P.flags_r & 0x100000000000) = 0) or (P.flags_r & 0x1000) = 0)
```

Use the redshifts to get distances using Hubble's law, and then use those distances to work out the absolute r-band magnitude M_r of each galaxy¹. Now, naively we could just make a luminosity function by making a histogram of $\log(N)$ as a function of M_r but that would be bad because this is a magnitude-limited dataset, and it is easier to see bright galaxies than faint galaxies, so we would erroneously undercount the faint galaxies. Instead, we will build a weighted histogram, where we weight the counts *inversely* by the maximum volume out to which we could find each galaxy given its absolute magnitude.

¹ For this problem ignore cosmological effects and just assume that Hubble's law and flat Euclidian geometry work. They actually do not on these scales; we are looking at galaxies that are far enough away that we should actually take into account the proper geometry of space. Ignoring this introduces errors into our calculation, but they are not so large as to render this exercise useless....

So first, for each galaxy, given its absolute magnitude M_r and the apparent magnitude limit of the catalog ($m_{r,lim} = 17.5$), write down an expression for the maximum distance d_{max} to which we could see it; in other words, at what distance would its apparent magnitude match our limiting magnitude? Then for each galaxy work out the spherical volume V_{max} corresponding to that d_{max} . Make sure your expressions give d_{max} in Mpc and V_{max} in Mpc^3 .

Now we are ready to make the luminosity function, properly weighted by $1/V_{max}$. Use numpy's histogram function to make a plot of the log of the weighted counts in bins of absolute magnitude that run from -24 to -16 in steps of 0.25 mag:

```
mbins=np.arange(-24,-16,0.25)
hist,edges=np.histogram(absmag_r,bins=mbins,weights=1/Vmax)
bincenters=0.5*(mbins[1:]+mbins[:-1])
logN=np.log10(hist)
plt.scatter(bincenters,logN)
```

Compare the shape of your LF to that of the Montero-Dorta & Prada "All Galaxies" LF shown in class. (Don't worry that the y-axis values don't match, since we haven't normalized the LF the same way. But the shape should look comparable.) On your LF, sketch where the Montero-Dorta & Prada value for M^* lies (making sure to correct their $M^* - 5 \log h$ value for the Hubble constant you've used in your analysis; describe what Hubble constant you used, and thus what you calculate for M^*). Does this M^* seem to be reasonable for your LF? In general, describe how well or how poorly your LF follows the approximate shape of a Schechter function -- where does it look reasonable, where (if anywhere) does it not?

Now choose a range of absolute magnitude over which to fit the faint end power law slope, justify that choice, and then fit a straight line to $\log N$ vs M_r over that range to derive the parameter α for the Schechter function. **Plot your fit over your data to make sure your fit looks reasonable.** Compare your value for α to values discussed in class, and discuss any systematic uncertainties that might be affecting your derived value for α .

2. Thick and thin disks (courtesy Sparke & Gallagher)

Assume the number of stars in the galaxy follows an exponential in radius and z-height:

$$n(R, z) = n_0 e^{-R/h_R} e^{-|z|/h_z}$$

Integrate that equation over all z (i.e., from $-\infty$ to $+\infty$) to show that the surface density of stars (number per unit area) at a radius R is given by $\Sigma(R) = 2n_0 h_z e^{-R/h_R}$. If the stellar population has an average luminosity per star of $\langle L \rangle$, the surface luminosity density is then $I(R) = \langle L \rangle \Sigma(R)$. Show then that the total luminosity of the disk is given by $L_{tot} = 2\pi I(0) h_R^2$.

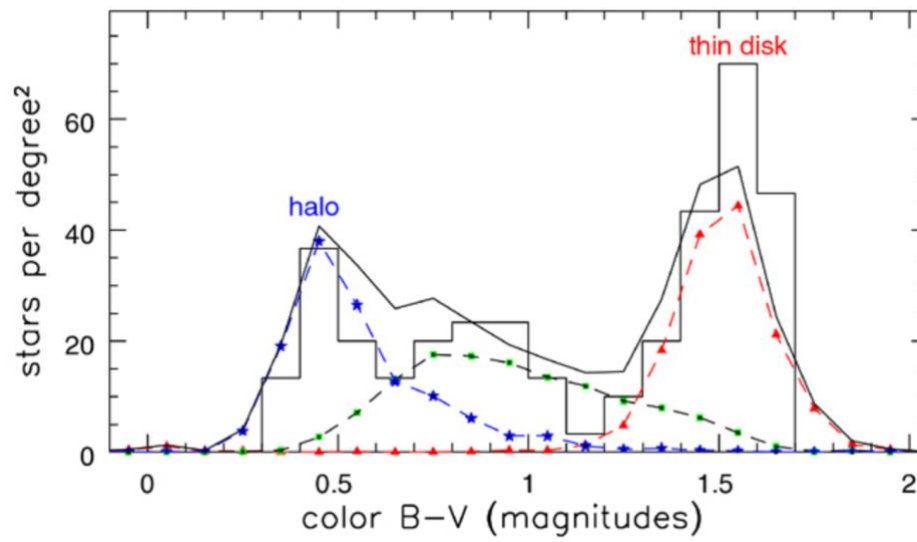
For the Milky Way, $L_{tot,V} \approx 1.5 \times 10^{10} L_{\odot,V}$, $h_R \approx 4$ kpc, and $R_{\odot} \approx 8$ kpc. Show that at the solar circle, $I \approx 20 L_{\odot,V} \text{ pc}^{-2}$. As you will work out later in this problem set, the mass surface density at the solar circle is $\approx 50 M_{\odot} \text{ pc}^{-2}$, which gives a mass-to-light ratio in solar units of

$M/L \approx 2.5$. This value is higher than what we get if we measure the value using stars within 100 parsecs of the Sun. **Why is that so?** Think about what kind of stars live only near the midplane of the disk, and what their (individual) mass-to-light ratios would be.

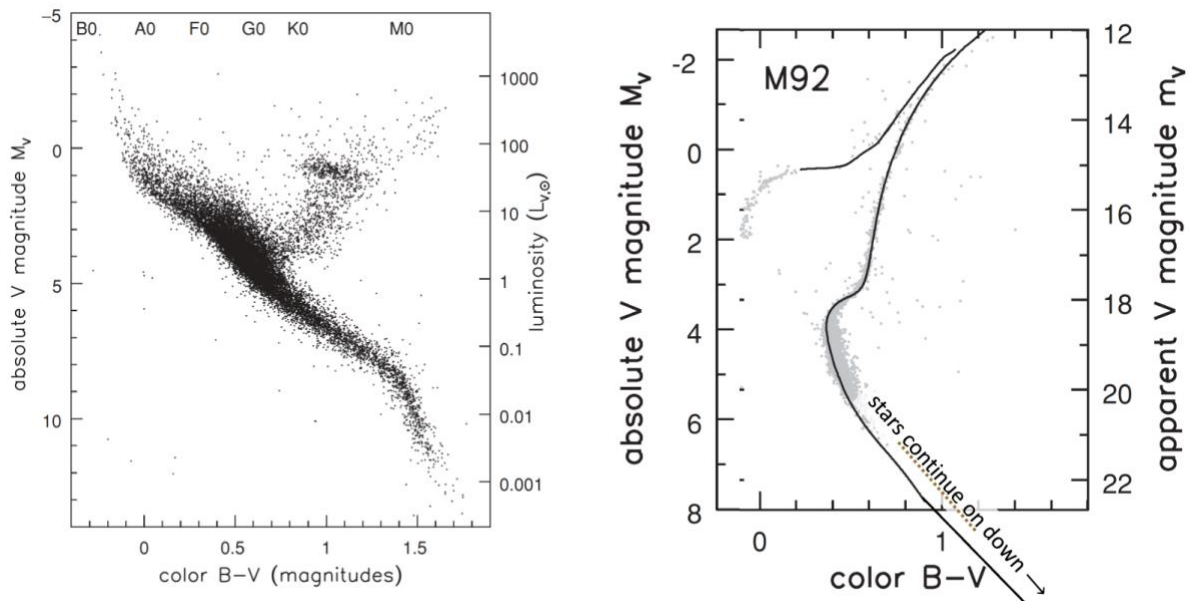
Finally, if 90% of the total stars in the midplane of the Galaxy's disk come from the thin disk, and 10% come from the thick disk, and if the thick disk has a scale height three times larger than the thin disk, **show that the surface density of stars follows the relationship $\Sigma_{thin}(R) \approx 3\Sigma_{thick}(R)$** . (For this problem, assume the thin and thick disks have the same scale length, which is not a good assumption, but it makes the problem work!)

3. Photometric Selection (courtesy Sparke & Gallagher)

The histogram below shows the density of stars near the North Galactic Pole (NGP) as a function of B–V color, in a sample of stars chosen to have apparent magnitudes in the range $m_V=19\text{--}20$. The dashed lines show the prediction of a star count model for stars in the thin disk (red), thick disk (green), halo (blue) of the Galaxy, while the black line shows the sum of those components. Note how the thin disk and halo populations separate out nicely by color in this magnitude range. In this problem you will see why this works.



Use the Hipparcos color-magnitude diagram for solar neighborhood stars (below left) as your model for thin disk stars, and the color-magnitude diagram for the metal-poor globular cluster M92 (below right) as your model for halo stars. (Note that in the M92 CMD, there **are** stars on the lower main sequence down to $B-V \sim 1.5$ and redder, it's just that they aren't showing on the plot.)



First look at blue stars, with colors $B-V \sim 0.4$. How far away must a disk star of this color be if it had $m_V=20$? In a population of stars like M92's, how far away would stars of this color be if they were at $m_V=20$?

For redder stars, what absolute magnitudes could a disk star have at $B-V=1.5$? How far would these types of stars lie if they were in our sample at $m_V=20$? In M92, the reddest stars have $B-V \sim 1.2$. How distant would these stars be in our sample?

Given those calculations, explain why the reddest stars in the first figure are likely to belong to the disk, while the bluest stars belong to the halo.

4. Gaia and stellar kinematics

I have pulled down a Gaia dataset that contains proper motion and distance data for stars with parallaxes >5 milliarcsec that also live within 10 degrees of the disk plane. It's big enough that its best to store it as a binary fits table, which you can read in python by saying:

```
from astropy.table import Table
import os
filename='Gaia_bLT10_pGT5.fits'
if not os.path.isfile(filename):
    gaia_data=Table.read('http://burro.case.edu/Academics/Astr323/HW/HW2/'+filename)
    gaia_data.write(filename)
else:
    gaia_data=Table.read(filename)
```

There are close to 300,000 stars in that database, and if you say `gaia_data.colnames`, you will see the information contained in the table. The file contains:

- ra, dec: coordinates in right ascension and declination, both in degrees.
- l,b: coordinates in Galactic longitude and latitude, both in degrees.
- parallax: parallax in milli-arcsec
- pmra,pmdec: proper motion (milli-arcsec/yr) in ra and dec coordinates
- pml,pmb: proper motion (milli-arcsec/yr) in Galactic l and b coordinates
- phot_g_mean_mag, bp_rp: apparent Gp magnitude and Bp-Rp color (these are the Gaia filter names)
- star_type: a coding I added: (1=main sequence, 2=red giant, 3=white dwarf, 0=other)

And remember you access the data by saying for example `gaia_data['parallax']`. If you wanted the parallax data only for RGB stars, you could say

```
rgb_stars = gaia_data['star_type']==2
p = gaia_data['parallax'][rgb_stars]
```

Important: Gaia provides parallax and proper motion with units of milli-arcsec and milli-arcsec/yr, respectively. Don't forget to convert these to arcsec and arcsec/yr!

First, calculate the distance to each star from the parallax, convert the apparent magnitude to absolute magnitude, and then **make a color magnitude plot** to make sure things look sensible. There are too many stars to plot each one as a dot, so [use this little code snippet to make a binned density plot](#).

Now, combine the distance and the proper motion in the b direction (Galactic latitude) to calculate each star's W velocity – the velocity up/down out of the plane. Look in the notes where we talked about proper motion and velocities to remind yourself how we do this calculation.

Do the following for each of the following three samples: **All MS stars, RGB stars, blue MS stars with Bp-Rp<0.5**:

- **Plot a histogram of the W velocities for each sample** (set your histogram to have 50 bins running from $W=-75$ km/s to $+75$ km/s)
- **Use each sample to find the Sun's W velocity and the velocity dispersion of the stars in the sample.**

Give a good astronomical discussion of why the values are different or not between the different samples, and compare your numbers to the values given in class.

5. Gaia and the vertical structure of the disk

Here we are going to compute the scale height of the thin disk using Gaia data. Before tackling the data, though, **work this out: if the density of stars follows an exponential like $\rho \sim e^{-z/h_z}$, if**

you plot $\log_{10} \rho$ vs z , how does the slope of the line relate to the scale height h_z ?

I pulled down another Gaia dataset for stars within 10 degrees of the North or South Galactic Pole (in other words, straight up or down out of the disk). For this sample, I didn't require that they be within 200 pc of the Sun, but I did make sure they had good parallaxes. So in this sample, there is no hard distance limit, but faint stars drop out of the sample because its hard to get a good parallax for faint stars.

You can read the file the same way as you did above, but this time use `filename = 'Gaia_NSGP.fits'`. Like the first dataset, there are about 300,000 stars in that database.

Like before, calculate the distance and apparent magnitude and **plot a CMD** to make sure everything looks good.

Then do the following for the following two samples: **All MS stars, red clump stars in the RGB**. You should be able to spot the red clump in your CMD; **what is its absolute magnitude?** Then select RGB stars which have an absolute magnitude that within 0.5mag of the clump center as your red clump sample.

For each of those two samples, count the number of stars in bins of distance, and convert that into a density of stars as function of distance from the plane. Use this code to do that:

```
N,edges=np.histogram(distance,bins=100,range=[0,4000])
vol=(4.*np.pi/3)*edges**3
dvol=np.diff(vol)
logdens=np.log10(N/dvol)
bincent=0.5*(edges[1:]+edges[:-1])
plt.scatter(bincent,logdens)
```

In your writeup, **explain line-by-line what that code snippet is doing**. I want you to understand the code, not just run it!

Fit a straight line to the data for bins ≤ 1 kpc distant, where the thin disk dominates. Give the parameters of the fit, overplot it on your data, and use it to derive the thin disk scale height.

After you've done that for each sample, compare the two numbers (the scale heights for the "All MS" sample and the "red clump" sample). How well do they compare to each other, and how well do they compare to values given in class? The red clump stars should give a better estimate for the scale height than the All MS sample. Explain why. Think about the luminosity spread of the two samples, and how that connects to the numbers of stars you observe as a function of distance.

How well does your fit work over the whole range of distances on your plot (not just where you fit)? Explain any differences you see between your fit and the behavior of the data points.

6. The mass density of the disk

The scale height and vertical velocity dispersion are tracing the surface mass density of the disk. A rough approximation for how they are related was given in class: $\sigma_W^2 \approx 2\pi G \Sigma h_z$.

Using values that you calculated in the previous two problems to solve for the surface mass density (in $M_\odot \text{ pc}^{-2}$) of the Galactic disk. *PLEASE DON'T WORK IN SI UNITS!!*

Remember, if you measure distance in parsecs, mass in solar masses, and velocity in km/s, a reasonable approximation for the gravitational constant is $G = 4.3 \times 10^{-3}$.

Now take your surface density and divide by the disk thickness (roughly twice your measured disk scale height) to get a rough estimate of the local volume density of the disk: $\rho \approx \Sigma / (2h_z)$ (in $M_\odot \text{ pc}^{-3}$).

Compare your values of Σ and/or ρ to estimates published in the refereed literature for the mass density of the Milky Way's disk in the solar neighborhood. Use the [ADS abstract service](#) to look up a few of the following papers and see what they find: Chakrabarti et al 2021; McGaugh 2016; Bienayme et al 2005; Holmberg & Flynn 2004; Siebert et al 2003. Also compare your estimate to the *observed* density of stars and gas described in McKee et al 2015.

Grad Student problem #1 (courtesy Heather Morrison):

Assume that you are observing stars in a square degree straight up out of the disk plane, that the thick disk has a scale height of 1 kpc and the thin disk 300 pc, and the halo has an r^{-3} density distribution. Furthermore, assume that the ratio of thin to thick to halo stars at the Sun is 1000:100:1 and that the local density of halo giants is 35 per kpc^3 . Draw plots of the number of stars expected as a function of z distance (i.e., the number per square degree per parsec of sightline) from $z=0$ to 15 kpc. At what distance will the counts of thin disk stars peak? thick disk stars? halo stars? (You will need to account for the variation in volume element along the line of sight as well as the variation in star density.)

Grad Student problem #2:

Let's go back to the solar motion problem and work out the U and V solar motion as well.

First, if we restrict our analysis to stars in the disk plane (small b), the observed velocity of stars projected onto Galactic longitude is a simple linear combination of trig functions that looks like this:

$$v_l = \pm U_\odot \times \text{trig}_U(l) \pm V_\odot \times \text{trig}_V(l)$$

Work out what the trig functions are and make the plus/minus decisions. It's easiest just to do this with geometry and common sense: make a sketch, draw vectors, think. Explain how you arrived at your answer.

Now make plots of v_l vs l for two samples (separately): **RGB stars, blue MS stars with $B_p - R_p < 0.5$** . For each sample, fit that function to the data (make sure you also overplot your fit) to derive U_\odot and V_\odot for your two samples. This is a nonlinear fit you are making, which can be tricky. Use `scipy.optimize.curve_fit`, and as an input guess give it the values for U_\odot and V_\odot that I gave in class. Discuss your fitted values for U_\odot and V_\odot , and how well each one matches the solar motion values given in class. Describe differences between the values derived for the two samples.