ASTR 222 (Spring 2025)
Homework #2
Due February 19, 2025 at 5:00pm

Note: The datasite for this HW set is http://burro.case.edu/Academics/Astr222/HW/HW2/

1. Studying the Galaxy with Globular Clusters

The datafile gc.dat contains a list of Milky Way globular clusters, taken from Harris 96. It gives:

- the angular coordinates of each cluster in Galactic coordinates (l, b). Note that these coordinates are different from the coordinates discussed in class (yep, ain't life difficult?) — those coordinates (R, θ, z) measured position with respect to the center of the Galaxy, while these Galactic coordinates (l, b) are the observed angular position of the globular cluster, as viewed from the Sun (i.e., our viewing position). Here's an image showing the coordinate system. So l measures angle in the sky, with $l = 0°$ defined as pointing towards the Galactic Center, and $l = 90°$ pointing in the direction of the Galaxy's rotation, while b measures angle above or below the plane of the Milky Way. If a cluster has coordinate $(l, b) = (0°, 20°)$, it would be viewed 20 degrees above the disk plane, in the direction of the Galactic Center. Got it?
- the distance from the Sun, R, in kiloparsecs
- the metallicity of the globular cluster, [Fe/H]. (Note: if [Fe/H] $= -999$, that means there is no measurement.)
- the observed radial velocity of the cluster, V_r, in km/s. (Note: if $V_r = -999$, that means there is no measurement.)

We are going to use this data to learn about the distance to the galactic center, and the rotation speed of the Milky Way.

First, transform the observed data (l, b, R) into cartesian coordinates (X, Y, Z). $+X$ should point towards the galactic center ($l = 0°$), $+Y$ should point in the direction of rotation ($l = 90°$), and $+Z$ should point above the plane (positive b). Defined this way, we have $X = R \cos l \cos b$, $Y = R \sin l \cos b$, and $Z = R \sin b$.

Coding alert: Be careful with trig functions when coding! numpy trig functions assume angles are given in radians, not degrees. But the data file lists angles in degrees, so you must convert. So `Z=R*np.sin(b)` would be incorrect, but `Z=R*np.sin(np.radians(b))` will work.

Calculate and plot the X, Y, Z positions to get a feel for the distribution of globular clusters in the Galaxy. Plot the XY and XZ projections so that you can see how things are distributed along each axis.

- Make XY and XZ plots for the metal-poor globular clusters with [Fe/H] < -0.8.
- Make the same plot for the metal-rich globular clusters with [Fe/H] > -0.8
- Note: Be careful not to include clusters which do not have measured [Fe/H]! So make sure to "filter" your data. Try something like this:

```
MR = (feh>-0.8) & (feh!=-999) # define which clusters are metal rich
plt.scatter(x[MR],y[MR])      # only plot metal rich clusters
```

- Explain how the two distributions are different.

Coding alert: Make sure your plots have the same XYZ range, and that a circle would be a circle on your plot! Otherwise you won't be able to compare the shape and extent of the two distributions. A few tips for making these plots:

- Set the X and Y limits on your plot by hand, rather than letting python autoscale the plot.
- Tell python you want equal axis ratio on your plots. Say `plt.axis('equal')`, or if you are using subplots, say `<subplot_name>.set_aspect('equal')`.
- Finally, also use small point sizes (like, s=3 or so) or else all the points will overlap too much….

Now, in the 1910s, Harlow Shapley derived the distance to the Galactic Center by using the positions of the globular clusters. Now it's your turn:

- Calculate the (X, Y, Z) center of the metal-poor globular cluster system?
- Calculate the (X, Y, Z) center of the metal-rich globular cluster system?
- Which of these two measurements do you feel is more accurate? Why?
- Based on these calculations, give your best estimate for the distance from the Sun to the Galactic Center.

Finally, let's use the observed velocities of the globular clusters to estimate the rotation speed of the Galaxy.

Plot V_r (on the y-axis) against l (on the x-axis). Only use the metal-poor globular clusters for this exercise. Explain why we want to exclude the metal-rich clusters when estimating the rotation of the disk. You should see a lot of scatter, but a hint of a sinusoidal trend. At $l = 90°$ the radial velocities are more negative, at $l = 270°$ they are more positive. This is because the globular cluster system isn't rotating. Remember that $l = 90°$ points along the direction of rotation; clusters along $l = 90°$ are in front of us and we are moving towards them, so they have a negative (approaching) radial velocity (and vice versa for the clusters at $l = 270°$).

Note: Make sure not to include clusters which have no measured V_r in your analysis! Use the filtering technique we talked about above.

So we should be able to fit a sine function to the data: $V_r = -V_{c,proj} \sin l$, $V_{c,proj}$ is the circular velocity projected along the average line of sight to the globular clusters. To get the true circular velocity V_r, we have to correct for this projection effect. Since $V_{c,proj} = V_c \langle \cos b \rangle$ (where the angle brackets mean "average"), if we average over all b's, we get $V_{c,proj} = (2/\pi)V_c$, so the true circular velocity is given by $V_c = (\pi/2)V_{c,proj}$.

Fit this sine model to the data, and derive V_c, the circular velocity of the Milky Way. How does your value compare to the "standard" value of 230-235 km/s?

Coding alert: remember you dont actually have to do a sine fit of V_r versus l, which would be a complicated non-linear fit. Instead, fit V_r versus $\sin l$, which is a simple linear fit! *(But remember to convert l to radians when using the sine function!)*

Now, what is the velocity dispersion of the globular cluster system? That is, if the globular clusters weren't moving at all, the sine fit would work perfectly and go through all the data points (projection effects aside). It clearly doesn't -- there is a lot of scatter around that line. That tells us that the clusters themselves are moving as well, but in a random sense, rather than as a rotating system. The scatter of the velocities around the sine fit gives us an estimate of how fast the clusters are moving in this random motion, which is referred to as their velocity dispersion.

You can estimate the scatter around a fit by calculating the the standard deviation of the residuals. In other words, for each data point, the residual (i.e., the difference) is given by $V_r - V_{r,fit} = V_r + V_{c,proj} \sin l$. You can then calculate the standard deviation of the residuals, and that's your estimate for how much the individual velocities are scattering around the fit, something called the dispersion (σ_r).

From your plot and fit, estimate the velocity dispersion around the fit, σ_r.

Finally to get the true velocity dispersion from the scatter, we again have to correct for projection effects. Since the radial velocity that we see is only one component of the 3D velocity vector of each globular, and velocities add in quadrature, we have

$$\sigma_{3D} = \sqrt{\sigma_r^2 + \sigma_\theta^2 + \sigma_\varphi^2} = \sqrt{3} \times \sigma_r$$

if we assume the velocity scatter in each direction is the same. That 3D velocity dispersion σ_{3D} is essentially just the characteristic speed of a typical globular cluster orbiting in the Milky Way's halo.

From your calculations, estimate the true 3D velocity dispersion σ_{3D} of the globular cluster system. Compare it to the rotation speed you calculated for the Milky Way's disk. They ought to be roughly similar – explain why?

2. A simple Oort limit

The datafile flynn.dat (taken from Flynn & Freeman 1993) contains apparent magnitude ('V'), radial velocity ('vrad', in km/s), and absolute magnitude ('MV') for a sample of red giant stars along a line of sight at Galactic latitude of $b \approx -90°$ (in other words looking directly "down" out of the plane of the Milky Way's disk). We are going to use this dataset to measure the mass density of the Milky Way's disk in the solar neighborhood. Remember how we do this, by using the Oort limit, which means we need the velocity dispersion and scale height of the stars.

Deriving the Scale Height of the Red Giants

Calculate the distance (d) to each star from its apparent and absolute magnitude. Because these stars are all at $b \approx -90°$, their Z distance from the plane is pretty much given by the distance of the star, once we correct for the distance of the Sun from the galactic plane. Since the Sun is "above" the disk plane (at $Z = +30$ pc) and these stars are seen "below" the disk plane, their distance from the disk plane is simply given by $Z = d - 30$ pc.

Then find the number of stars (N_*) as a function of Z by making bins of, say, 200 pc size and counting the number of stars in each bin. Make a plot of $ln(N_*)$ (y axis) vs Z (x axis) and find the scale height (z_0) of the stars by fitting a straight line to the data and deriving the slope. Mathematically, show how the slope of the line relates to the scale height of the stars. When fitting, remember to make smart choices about which bins to include in your fit; including all your bins in the fit is probably not a good choice!

Coding alert: First, take a look at the binning and fitting example code shown in class and given on the Python help page to figure out how to code up this exercise. Also, be careful with logs when coding. In numpy, `np.log(x)` means "natural log of x" and `np.log10(x)` means "base 10 log of x". Make sure you use the correct one when doing calculations! While usually in astronomy we use base-10 logs, in ***this*** case we want the natural log, since we are plotting an exponential function.

Deriving the Velocity Dispersion of the Red Giants

Just like we said a star's distance from the disk plane was roughly given by its Z coordinate (once we corrected for the Sun's Z coordinate), we can also say a star's radial velocity is roughly its Z velocity once we correct for the Sun's Z velocity.

From the distribution of velocities, calculate the Sun's Z velocity, and describe how you did it. Then subtract the Sun's velocity from all the velocities to get the Z velocity of each star. Then bin up the stars in bins of Z, calculate the velocity dispersion of the stars within each of the bins, and then make a plot of velocity dispersion as a function of Z height. Describe how and why the dispersion changes with Z. If you had to pick one number to best describe the velocity dispersion, what would it be, and why?

Calculate the Oort limit estimate for the mass density of the Milky Way's disk

Given the scale height you calculated and your estimate of the velocity dispersion, calculate the mass density of the disk using the method shown in class.

Describe what you think some of the sources of error might be in this calculation, and describe how they may have affected your result.