8.902 Fall 2023 - Problem Set #1

Due Tuesday, September 19

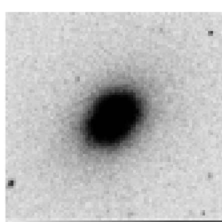
This problem set builds literacy with image analysis software and scientific programming techniques. There are many packages available for data analysis; you should feel free to use whatever you find most comfortable and useful. You do not need to turn in your code, but please write enough work (equations and explanations where needed) to demonstrate what you've done.

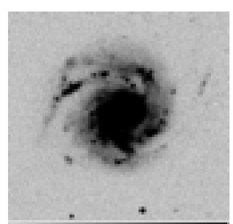
One way to investigate the empirical properties of galaxies is to "observe" a few representative examples. We will use the archives of the Sloan Digital Sky Survey to study two randomly selected nearby galaxies in detail. Optical (G band) images and spectra of the two galaxies have been posted in the Pset1 folder of the course webpage with filenames given in the table below.

These files are in FITS format, the de facto standard for astronomical data transfer. Spend some time learning to read these files into software arrays, displaying them, and playing with the color stretch (e.g. using the excellent and freely-available program ds9 or Astropy). The approximate pixel positions of the science objects are given in the table below.

The data value of each pixel in the file is expressed as a number (usually called DN or digital number) that is linearly proportional to the flux of photons upon that pixel. For now, don't worry about the proportionality constant; we will calibrate it below.

The spectral files contain 4 x N arrays of pixel values, the first row of which comprises a 1D array of pixel values. Each 1-D array corresponds to the intensity of flux from the object at a given wavelength. The wavelength of each pixel (in Angstroms) can be calculated with the formula: $\lambda_i = 10^{w+0.0001i}$. Here, w is a constant and i runs from 0 to N-1. Values are given in the table below.





Object	Image File	Spectrum File	Image X	Image Y	w	N
Galaxy 1	galaxy1.fit	galaxy1spec.fit	1831.5	636.5	3.5823	3818
Galaxy 2	galaxy2.fit	galaxy2spec.fit	1346.0	945.0	3.5825	3815

1 Galaxy Profiles

A) Plot the surface brightness profiles of galaxies 1 and 2. For this part, you should plot in units of the log of the pixel counts (DN) as a function of linear radial distance from the galaxy center in arcseconds and use a bin width of one pixel. Be sure to subtract off an estimate of the underlying contribution from the sky (which does not have zero flux)! You may also need to refine the center pixel positions to improve your profile. The instrument used to take the image has a scale on the sky of 0.394 arcseconds per pixel.

B) To improve the signal-to-noise ratio, group the data points into radial bins of width 2 arcseconds, and plot the median surface brightness in each bin as a function of galactocentric bin radius for each object.

C) The image for Galaxy 1 (Galaxy 2) has been calibrated so that a 20th magnitude object contributes a *total* of 2328 (2496) counts above the sky background level, when integrated over its entire profile. Re-cast the y-axis of your radial profiles from (B) into the customary units for surface brightness — magnitudes per square arcsecond — and re-plot the profile.

D) Most galaxies have surface brightness profiles that fit into one of two families: the exponential profile or the so-called de Vaucouleurs profile (sometimes called the $R^{1/4}$ profile):

$$I(r) = I_0 \exp[-r/r_s] \qquad \text{Exponential} \qquad (1)$$

$$I(r) = I_0 \exp[-7.67(r/r_{eff})^{1/4}] \qquad \text{de Vaucouleurs} \qquad (2)$$

Compare your measured surface brightness data with these two profile models. Numerically fit the two profiles to each galaxy and estimate the "scale length" r_s for the exponential profile and the "effective radius" r_{eff} for the de Vaucouleurs profile. Overplot your estimate of the model onto the data. Is Galaxy 1 better described by an exponential or de Vaucouleurs profile? What about Galaxy 2?

E) Does a single-component model (i.e. single de Vaucouleurs or exponential function) adequately fit the entire profiles of both galaxies? If not, then indicate where the assumption breaks down, and suggest a physical explanation.

2 Galaxy Spectra – Redshifts

A) Examine the spectra of the two galaxies. In each spectrum, the underlying continuum flux represents the integrated spectra of the galaxies' stars. Assuming that stellar spectra are perfect blackbodies, use Wien's displacement law to estimate the temperature of the "typical" star in Galaxy 1 and Galaxy 2. Roughly what spectral type do these represent? Do not trouble with more complicated stellar models, as the blackbody assumption is not particularly good for real stars. Just get the rough idea here.

B) The most prominent feature in the spectrum of Galaxy 2 is an emission spike located at 6679 Angstroms. This is an Hα line associated with hydrogen recombination radiation from the galaxy's HII regions. The rest wavelength of Hα is 6563 Angstroms, but the line is redshifted due to cosmic expansion. Find the centroid of the line profile, and use this to calculate the recession velocity of Galaxy 2 from the earth in km/s. Galaxy 1 has no emission spike, but there are several strong absorption features. The absorption line at 8741 Angstroms corresponds to the redshifted CaII line (rest frame 8545 Angstroms). Find and use the centroid of this line to calculate the recession velocity of Galaxy 1.

C) Edwin Hubble tells us that a galaxy's recession velocity is proportional to its distance from Earth: $v = H_0 d$. Here, H_0, the local "Hubble constant", is 71 km/s/Mpc. Calculate the physical distance to each of the two galaxies in Mpc. Ignoring cosmological corrections, use your estimates of the characteristic radii found in (1D) to estimate the physical extent (i.e. radius) of the luminous matter in each galaxy (in kpc).

D) By integrating the model profiles determined in (1D) and combining with the distance estimates from (2C), calculate the total G band luminosity of each galaxy, in solar units. For the purposes of the problem, assume that 90% of the Sun's total luminosity is emitted in the G band.

3 Galaxy Spectra – Lines

A) Examine the CaII absorption line in the spectrum of Galaxy 1 near 8740 Å. Fit a Gaussian profile to this absorption line to estimate the width σ of the feature. Express your answer in both Angstroms and in km/s. It is helpful to remember the very-useful approximation (for $\sigma_v << c$) that $\sigma_\lambda/\lambda = \sigma_v/c$.

B) The broadening of this line is caused by Doppler motions of the stars orbiting in the galaxy's potential. However, there is also a contribution to line broadening caused by motions of atoms within the stars themselves, as well as a contribution from the response function of the instrument used to make the measurements. Observations of a control sample of similar stars in the Solar neighborhood made with the same spectrograph exhibit sodium line widths of $\sigma_v = 95$ km/s. Assuming that instrumental effects, atomic motions in the stars, and broadening from stellar orbits all behave like Gaussian distributions (i.e. their σ's add in quadrature), estimate the velocity dispersion of the stars in Galaxy 1 by subtracting off the other effects. Express your answer in units of km/s.

C) Suppose the dispersion measured in (3B) is characteristic of the stellar orbital velocities at the characteristic radius measured in (1D). What is the mass of Galaxy 1 interior to this radius, in solar mass units? Compare this result with your estimate of the number of stars in the galaxy determined in (2D).