Project 3:
Studying Distant Galaxies with the Sloan Digital Sky Survey

Parts A&B due Nov. 15th

Parts C&D due Nov. 29th

Now we move out to more distant galaxies applying what we've learned from star clusters and populations in the nearby universe. You'll take a look at galaxies within one well-studied extragalactic field (the "Lockman Hole") selected to have small numbers of contaminating stars and foreground dust extinction. The goal of this project is for you to examine the types of information we can learn about galaxy properties, star formation histories, and star formation rates from photometry and spectra. Currently, many astronomers are using the same techniques you will learn here but applying them to ever more distant objects or ever larger samples of galaxies. The goal of this work is to try understanding how galaxies are made.

The Sloan Digital Sky Survey (SDSS) is an imaging and spectroscopic survey of about 20% of the sky (8000 deg^2) away from the plane of the Milky Way. It has measured photometry in five filters (u, g, r, i, z) for 230 million unique objects and has taken over a one million spectra. All of this was done on a dedicated 2.5-meter telescope in New Mexico.

For this project, you will be looking at galaxies in just a small section of the survey. You will each be exploring a 30' radius circle around the following positions:

Anabelle & Floria: RA=160 Dec=58

Tori & Katharine: RA=161.5 Dec=58

Perry & Nell: RA=164 Dec=58

Part A: Exploring Sloan Data

The final release of the SDSS data was done in 2008, it is known as "Data Release 7" or DR7. The main DR7 website is at:

`http://www.sdss.org/dr7/`

but all the tools you will be using are linked from the following database (a.k.a. CAS=Catalog Archive Server) page:

`http://cas.sdss.org/astrodr7/`

Open a web browser and use the "Navigate" tool on the CAS page to take a first look at the data.

1. First type in the central RA and Dec of your field and take a look. This is just a small section of the sky (the pixels are 0.4"), so you will just see a few stars and galaxies. Click on an object and a summary of that objects coordinates, type and magnitudes will be given in the top right. Find a star and one galaxy.
2. Now click the box on the left entitled "Objects with spectra". If spectra have been taken for any objects in your field, this will be shown with a small red circle which shows where the spectrum was taken.
3. Now set off! Head in one direction and keep navigating around. Your goal is to find both a clear spiral and elliptical galaxy which have a spectrum available. Explore both objects as outlined in the next step.
4. To find out more about an object, select it and then click on the "Explore" link on the right hand side of the page. This will open a new page; there are two important sections, a picture with the photometry information on it from the imaging, and then below that a spectrum and information on the redshift. Include the pictures of both galaxies and their spectrum in your report. Record the following information from this page: the redshift (as given on the plot), the u, g, and r magnitudes for the source at the top of the page (these are "model" magnitude)

as well as the magnitude of the light that fell in the spectroscopic fiber, (`fiberMag_r` in the spectrum section).

Using this data, answer the following questions:

Question A1) Calculate the distance of the two galaxies assuming a Hubble constant of 70 km/sec/Mpc. To do this, first calculate the recession velocity using the approximate formula $z \sim \frac{v}{c}$, then use the Hubble constant to find the distance in Mpc.

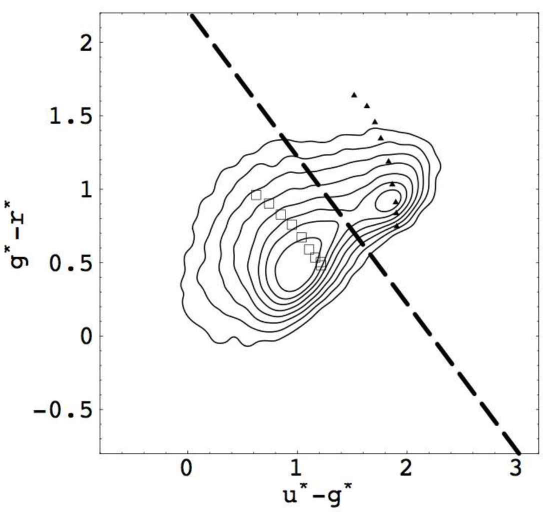
Question A2) Calculate the absolute r band magnitude for both galaxies. Then use this to determine the luminosity of the galaxies in solar luminosities. The absolute magnitude of the sun in r band is 4.7. How does this compare to the luminosity of the Milky Way ($\sim 2 \times 10^{10}$ $L_{\odot}$)?

Question A3) Compare the $g - r$ colors of the spiral and elliptical galaxy. Which one is bluer? Explain why.

Question A4) The "model" magnitude is based on a model fit to the entire galaxy, while the "fiber" magnitude is just the light that falls in the 3" diameter optical fiber used to measure the spectrum. Compare these two magnitudes in r to measure the fraction of the total galaxy light measured by the spectrum in both galaxies. How might this affect our determination of stellar ages from the spectrum?

Question A5) Compare the spectrum of the two galaxies to eachother and the spectra from HW2 (available on Sakai if needed). Why are the wavelengths in the spectrum shifted from those in the models? What absorption lines are strongest in each spectrum? What cluster spectrum matches best to each galaxy? Also comment on any emission lines and their origin.

Part B: Lots of Galaxies!



A color-color diagram of SDSS galaxies from Strateva+ 2001. Shown are the $u - g$ and $g - r$ colors of 150,000 galaxies. These fall into two categories now often called the 'red sequence' and 'blue cloud'. The triangles and squares are models of elliptical and spiral galaxies at redshifts ranging from 0 to 0.4.

In this section, you will download data for galaxies in a 30' radius circle around your assigned central positions. To do this, go to the "Imaging Query" tool on the CAS page. Do the following:

1. Change the number at the top from 50 to 0, to get all the objects in your area.
2. Select the "CSV" output option, which will give you a comma separated table.
3. Next, select what parameters you want output. In the "Imaging" section, select *radec, model_mags* and *model_magerrs*. From the spectroscopy menu select just z (the redshift).
4. For position constraints, use the 'cone' option and input your central RA and Dec, and a radius of 30'.
5. Under imaging constraints, use a maximum r magnitude of 21 and select only "Extended Sources" (unselect Point Sources). Use a min QA of 1.

6. Now run your search and transfer the resulting file to a proj3/ subdirectory on ladyhuggins. The file should have ∼2000 lines.
7. Look at the output file and try to understand the columns. Use "Imaging" link under "Parameters to Return" for help if needed.

Now log into ladyhuggins and create an IDL program that reads in the data (use `READCOL's DELIM` keyword) and make the following plots. Please include these as specified in your report.

- Plot a color-magnitude diagram, $g - r$ vs. r. Overplot the elliptical and spiral galaxy you found in Part A using large symbols. Look at where the galaxies fall; are bright ones preferentially one color?
- Plot a color-color diagram of all your galaxies like the one shown above (but with your galaxies as points) using the same limits for the axes. Overplot the colors of the two galaxies from Part A using large symbols. *Please include in your report and answer Question B1.*
- Galaxies with spectra available have redshift values $z > 0.0$. Select out just these galaxies and create a histogram of redshifts using a binsize of 0.01 in z. *Please include in your report and answer Question B2.*
- Using your histogram of redshifts, select out the galaxies in the most significant peak in redshift. All these galaxies are at roughly the same distance and are probably part of a galaxy cluster, group or filament. Make a color-magnitude with the absolute r band magnitude estimating the distance the same way as in Part A. *Please include in your report and answer Questions B3 and B4.*
- From galaxies within your redshift bin, select out one bright blue galaxy and one bright red galaxy. Record the two galaxies RA, Dec, r magnitude, and redshift (to as many significant digits as available). You will analyze the spectrum of these in Part C.

Question B1) Do the spiral and elliptical galaxies from Part A fall where you expect in the diagram? Comparing your plot to the figure above from Strateva+ 2001, roughly estimate the fraction of red and blue galaxies in this area of the sky.
Question B2) What are the highest and lowest redshifts? Using the same method as in A1, what do these distances correspond to in Mpc and lightyears? (note: ignore cosmological distance/time effects here) Compare this to the age of the universe.
Question B3) How large a range of colors is there? The $g - r$ color roughly corresponds to the F475W-F606W color from Project 2. Compare the span of galaxy colors to the span of cluster colors there. What does this tell you about the range of ages?
Question B4) The Milky Way has an absolute magnitude in r of roughly -21. What are the relative luminosities of the brightest red galaxy and blue galaxy and which is brighter? Would the differences in mass be larger or smaller than the difference in luminosity? Explain.

Part C: Spectral Analysis

Using the coordinates of the two galaxies selected in Part B, go back to the "Navigation" tool and find these two galaxies. Your blue galaxy should look like a spiral and your red galaxy like an elliptical. In the "Explore" pane, click on the "FITS" link under "SpecObj". This will take you to a page with the spectrum above a series of links. You can right click on the right-most link on the first line to download the spectrum, a file name called `spSpec-*.fit`. This is the data of the spectrum plotted in the figure. Also record the redshift to at least three significant digits.
Now you will run a code (adapted from Tremonti+ 2004) that fits the spectrum to stellar populations models of different ages – these models are the *same ones* as in homework 2 and were created by

Bruzual & Charlot. The difference is that this time we have models of many ages, and will mix them together to best fit the spectrum of the galaxy. This gives a rough measurement of the galaxies' star formation history.
To run the code, first copy it from my directory into your proj3 subdirectory (where you spectra are): `~aseth/astro301/proj3/fit_sloan_spectrum.pro`. In IDL, type:

```
IDL> .r fit_sloan_spectrum
IDL> fit_sloan_spectrum,'spectrum.fit',redshift
```

The program will fit your spectrum to a combination of 10 spectra ranging from 10 Myr to 13 Gyr in age. It also fits for the reddening (both foreground and internal). A window shows your spectrum, with the best fit overplotted in red. Green regions are omitted from the fit (they are regions around common). Note that the spectrum has been shifted back to rest wavelength. The first line of output shows the ages of the templates, and then the next two show the fraction of light (LFRAC) and mass (MFRAC) in each age bin. After that is the best fitting dust optical depth in the r band TAU_r (defined in the next section). Finally there is the mass-to-light ration in the r band (M/L_r). Examine all these numbers and the fit on the screen and answer the follwing questions:
Question C1) Calculate the average age of stars in the galaxy weighted by the mass and luminosity. To do this, just multiply the mass and luminosity fractions times the ages and total them up. Which one gives you an older age? Why?
Question C2) Compare the typical age of stars in your blue and red galaxy. Do they correspond to what you expect?
Question C3) Calculate the mass in $M_\odot$ of the two galaxies using the r model magnitude and the mass-to-light ratio from the model. To do this, you first need to determine the luminosity in solar luminosities (see A2), and then convert it to a mass by multiplying the mass-to-light ratio. Compare the masses of your two galaxies.

Next, the program will measure the emission line strengths of the emission lines for you. These lines include Hα, Hβ, [OII]/3727Å, [OIII]/5003Å, the [NII] lines around H$alpha$, and the [SII] lines at 6730Å. You will use some of these to derive the star formation rate of the galaxy in the next section. The program uses the best fit stellar population model from the first part of the program and subtracts it off before fitting the emission lines. The fits to all the emission lines are shown in graphs and the Hα and Hβ fluxes are outputted to the screen. Record these fluxes for use in Part D.
Question C4) Why do we subtract off the model for the galaxy's starlight before calculating the emission line fluxes?

Part D: Deriving Star Formation Rates

In this final section you will calculate the star formation rates for both of your galaxies using both optical emission lines and the mid-infrared flux. For this section, show all your work deriving the star formation rates and answer the questions.
You will use two relations for the star formation rate for the Hα luminosity and the 24μm infrared luminosity compiled by Calzetti+ 2009:
SFR [$M_\odot$/year]= $5.45 \times 10^{-42}\, L_{H\alpha}$ [ergs/sec]
SFR [$M_\odot$/year]= $2.46 \times 10^{-43}\, L_{24\mu m}$ [ergs/sec]

1) The Hα SFR:

First, for both galaxies (if emission is detected in both), calculate the luminosity of Hα from the flux

using the inverse square law for light. Then plug this into the equation uncorrected for dust extinction; this is a lower limit on the SFRs.
You will now correct for the dust extinction in two ways. In addition to the A_V in magnitudes we met earlier in class, sometime dust extinction is characterized using the optical depth τ:
(Measured Luminosity) = (True Luminosity) $\times e^{-\tau(\lambda)}$

First, you will use the best-fit TAU_r value for the starlight determined in the previous section. Since the r band is at nearly the same wavelength as Hα, we can use it to correct our Hα flux. Calculate the true luminosity and convert this to a star formation rate.
Next you will use the emission lines themselves to estimate the dust extinction. For a typical HII region, the intrinsic ratio of the flux ratio Hα/Hβ=2.85 (Osterbrock 1989). The dependence of the optical depth, τ on wavelength is roughly described by (Charlot & Fall 2000):
$\tau(\lambda) \propto \lambda^{-0.7}$
First show that $\tau_{H\beta} = 1.236 \times \tau_{H\alpha}$. Then use your knowledge of the true/intrinsic Hα/Hβ ratio to calculate the dust optical depth τ at Hα and use this to determine the luminosity and star formation rate.
Question D1) For just your blue galaxy, compare the τ derived from the starlight fit to the one derived from the HII regions. Think about the different sources of light being used to make this correction, which would you think would give a larger dust extinction and why? Which is more accurate for measuring the Hα flux?

2) The Infrared SFR:
The field has been imaged as part of the Spitzer Wide-Area Infrared Extragalactic Survey (SWIRE). The 24μm band in Spitzer detects warm dust associated with star formation and is known to be a good star formation indicator.
Determining the IR SFR is relatively simple. Copy the catalog of SWIRE 24μm fluxes
`~aseth/astro301/proj/swire_24um_fluxes.dat`
into your directory. The file contains four columns with the coordinates, flux and error for all the sources. The fluxes are in micro-Janskys (μJy). Match your two sources to the closest SWIRE sources (they should be within a few arcseconds), and then convert the fluxes to luminosities. This is a bit tricky. Some help:
1 Jy = 10^{-23} ergs/sec/cm^2/Hz.
To convert to a luminosity in ergs/sec you need to use the inverse square law and multiply by the frequency of light at 24μm. Then compute the star formation rate.
Question D2) How do your SFR estimates agree? Consider the uncertainties in each measurement and argue which you think is more reliable.
Question D3) In Part C, you calculated the total mass of the galaxy and also learned the mass fraction of young stars with ages of ~10 Myr. Use this information to calculate a total mass in stars of this age, and compare it to your derived SFR, assuming it has been constant for the last 10 Myr.
Question D4) Measuring the cosmic star formation rate density as a function of redshift (often referred to as the Madau plot) is an important way to look at when galaxies evolved. You have calculated the SFRs for just a couple galaxies. If you wanted to measure the star formation rate density (e.g. the number of stars forming per cubic Mpc) in the universe 1 billion years ago, could you do it with the data that you have used in this project? If not, what additional data would you need?