# UCF Physics: AST 5765: Astronomical Data Analysis

## Fall 2019 Final Project

## Radial Velocity and Dark Matter in a Galaxy

Measure the velocity and mass *vs.* radius in a galaxy. You will extract a spectrum of the edge-on spiral galaxy UGC 9039, calibrate it, and derive its redshift *vs.* radius. From this you will calculate the mass enclosed within a given radius.

See the file `projdesc.pdf` for general instructions.

Data for this project were obtained in April 1999 by Prof. Riccardo Giovanelli and his student Barbara Catinella, both of Cornell University. They used the red camera of the Double Spectrograph on the Hale 5-meter telescope. The grating has 1200 lines/mm. Data were acquired through the $2''$ slit, with a grating tilt angle of $43° \; 10'$. All frames were automatically bias-subtracted by the instrument electronics. The FITS files are named `bsNNN.fits`, where `NNN` is the frame number.

| Frame # | Content |
|---|---|
| 324–333 | Dome flats |
| 334–343 | Bias frames (zero integration) |
| | |
| 397 | 600 sec exposure of the galaxy UGC 9039 |
| 395 | NeAr lamp spectrum associated with UGC 9039 |

First, look at one of each kind of image. The bias and flats should look familiar. The lamp spectrum has nearly vertical lines from a glowing neon-argon gas tube, which are a wavelength reference. The slit is on the left. The horizontal axis is wavelength and the vertical is distance along the slit. The galaxy image is dominated by sky lines, which are nearly straight, just like the lamp's lines. The galaxy's own emission features are curved, due to the edge-on spiral's rotation. The amount of offset is proportional to the radial velocity of that portion of the galaxy. Once you get rid of the sky lines, you should see at least 5 lines, plus a continuum.

1. (10 points) Write interpolation and sub-pixel shifting routine. This routine will use splines to evaluate a spectrum or other linear function at locations you specify. These could be the original $x$ data shifted, or could be many points between adjacent $x$ points (i.e., interpolating to a higher-resolution grid), or both. Interpolate only once! Accept the new $x$ locations, the original $x$ locations, and the original $y$ values.

   Call your routine splinterp and run it as follows:

```
import numpy as np
import gaussian as g
mag   =  5.
lo    = -2.
hi    =  2.
shift = -0.2
xold  = np.linspace(lo, hi, np.round( hi - lo        + 1))
```

```
xnew  = np.linspace(lo, hi, np.round((hi - lo) * mag + 1)) + shift
yold  = g.gaussian(xold)
ynew  = splinterp(xnew, xold, yold)
print(ynew)
```

2. (20 points) Write a non-trivial test of your interpolation/shifting routine, describe it in your paper, and show numerically that it works. Here and below, "non-trivial" means the test will likely fail if the code is wrong. For example, interpolating and shifting a horizontal line is trivial.

3. (10 points) Write interpolated cross-correlation routine. You will do many cross correlations, and it is important to do them all exactly the same way. Using a routine will ensure this. The goal is to calculate how shifted a spectrum is compared to a reference spectrum, so the inputs will include the two spectra. Since the shifts are often less than a pixel, you will need to interpolate (using the routine above) onto a grid at least $10\times$ as fine as the pixel grid (make the amount of interpolation variable; these instructions use $10\times$ only as an example). It will also help to be able to restrict the range of lags you consider.

   Create a linear array of integers representing the pixel positions in wavelength. It should start: 0, 1, 2, ... Create another linear array that is ten times as long and has 1/10th pixel positions. It should start 0, 0.1, 0.2, ... Use the routine above to interpolate the two spectra from the first position grid to the second position grid. From each interpolated spectrum, subtract its mean value. Pad the ends of each spectrum with as many zeros as you have wavelength channels, so that no wrapping occurs. Calculate the cross correlation over the range of lags you wish to consider.

   The shift between the spectrum and the reference spectrum is the peak of the cross-correlation, which you can find by fitting a Gaussian to a small region surrounding the peak of the cross-correlation *vs.* shift. Return this shift, being careful of the meaning of the sign. Be sure to check that the Gaussian routine got a good fit and do something appropriate if it doesn't.

   As you work with the routine below, you may find that some cross correlations are too peaked for a good Gaussian fit. So, if the fit fails the first time, you might try selecting just a few shifts before and after the peak for the Gaussian fit. Or, you could make the number of points to either side of the peak be an input variable. If all else fails, just return the peak value, but warn the user somehow.

4. (20 points) Write a non-trivial test of your cross-correlation routine, describe it in your paper, and show numerically that it works.

5. (10 points) Create bias frame. These CCD spectra automatically have electronics offsets ("bias") subtracted from them. However, just to be certain, we will create bias frames. Median-combine the bias frames to form a master bias frame. Save as 32-bit FITS. Print pixel [150, 150].

6. (10 points) Create flat field. Subtract bias from the dome flats and use normalized median combination to create a master dome flat. This is your flat field image, already normalized to unity. Save as 32-bit FITS. Print pixel [150, 150].

7. (10 points) Correct images. Subtract the bias frame from the target galaxy and the lamp frame. Divide the results by the flat field. Save as 32-bit FITS. Print pixel [150, 150].

8. (10 points) Find the $y$ location of the galaxy's brightness peak. Take the median of the image along the horizontal axis, zero the messy edge regions, plot it, and find the row number of the peak. This roughly defines zero velocity. Use the plot to choose the upper and lower limits of the region with galaxy flux. Print these limits, and the $y$ index of the row with the galaxy center.

9. (10 points) Define spectrum "stripes" and extract spectra. You want a series of spectra taken at different distances from the galaxy's center. The image is largely full of galaxy flux. To extract spectra while eliminating bad pixels, divide the galaxy image into a series of horizontal stripes, each 5 pixels high and running the whole spectral range of the image. The middle of the galaxy must be in the middle vertical pixel of a stripe. Eliminate the messy edges of the frame, but keep as much sky as you can, for background subtraction.

   Along each stripe, take the median of each 5-pixel stack, yielding for each stripe a single, 1D spectrum with no bad pixels. You can do this without a loop. Plot this clean, smaller image. Find the row number of the galaxy center in this image. This is your reference row, the index of the reference spectrum in the new galaxy image. Print the number of rows in this reduced image and the index of the reference spectrum (which should be at about half the number of rows). Plot the reference spectrum.

10. (10 points) Extract galaxy background spectrum. Taking advantage of the long slit, create a spectrum of the sky background in the galaxy image. Locate a region free of target flux and containing sky flux near the top of the image, using the plot from step 8 and looking at the image. Make a new image containing just these background rows. Calculate the median of each column (i.e., each wavelength); this is your sky spectrum. Ensure that the result is not contaminated by bad pixels. Print the low and high row numbers of the regions and plot the sky spectrum. Make A COPY of the galaxy frame, and subtract the sky from each row. The sky lines should almost go away, but you will still see some residual sky lines. You'll need to correct the sky lines better, below (which is why you made a copy here). Plot the sky-subtracted galaxy.

11. (10 points) Extract lamp spectra for intrinsic shifts. The Double Spectrograph optics distort spectral lines slightly to produce an intrinsic shift. That is, constant spectral lines curve very slightly in the image, resulting in the imperfect sky removal, above, if ignored. The lamp spectrum for the galaxy has constant-wavelength, high S/N lines that you can measure. Start by dividing the image into stripes and extracting spectra. Use the same striping as before. Plot this clean, smaller image. Plot the striped lamp spectrum at the index of the galaxy center.

12. (10 points) Calculate intrinsic shifts. For each lamp spectrum, calculate the shift relative to the lamp reference spectrum (which came from the same row as the galaxy reference spectrum). Cross-correlate the two spectra for a range of shifts that is substantially larger than the plausible range (say, -10 to +10 *original* pixels). These shifts are a systematic error you must remove from your galaxy's shifts and from the background spectra. Plot the shifts.

Now, do the same calculation, but between the galaxy's reference spectrum and each stripe in the galaxy. Plot the result.

Do the shifts resemble the intrinsic shifts or the S-curve of the galaxy's Doppler-shifted lines that you are really after? Depending on how you chose the background, they may resemble the intrinsic shifts, which is why you need to redo the background subtraction in the next step, shifting the background spectrum by the amount of the intrinsic shift for each stripe, first.

13. (10 points) Shift and remove background spectra. Find the average shift in your background region by averaging the shifts of lamp lines in the region you used to create the background spectrum. For each galaxy spectrum, find the difference between that spectrum's shift and the shift of the background. Shift the background spectrum by this amount and subtract it from the galaxy spectrum. The sky lines should disappear, leaving a series of at least five warped galaxy spectral lines. If you have an error, your sky-line removal will be poor and the cross correlation may line up the residual background rather than the galaxy. Be sure to check for this. Plot the sky-subtracted galaxy. Print the value of pixel [150, 150].

14. (10 points) Find galactic-rotation Doppler shifts. Cross-correlate each galaxy spectrum with the reference spectrum for a range of shifts that is substantially larger than the plausible range. Plot the Doppler shifts. The shifts near the frame edges will not be meaningful where there is no galaxy signal. Identify the rows with meaningful shifts, trim the rest, and plot. From now on, calculate only with those good shifts.

15. (10 points) Remove systematic shifts. Correct each good galactic Doppler shift for the intrinsic shift at that location, relative to the galactic reference spectrum. Overplot on the final plot from the last step.

16. (10 points) Get a line atlas. This step's goal is to determine the relationship between a pixel in the spectral ($x$) direction and the physical units of wavelength. This is called the *dispersion*, and its units are [Å/pix]. The dispersion will convert your pixel shifts into wavelength shifts.

    The lamp image in the set is the spectrum of a neon-argon (NeAr) lamp, taken at the same grating settings as your galaxy image. It shows atomic emission lines. You will compare your lamp reference spectrum to a spectrum with identified lines called a line atlas. Put a plot of the lamp spectrum *vs.* pixel number on screen.

    Next, generate the atlas spectrum. If you search the web for "HeNeAr line atlas", you will find a site that makes synthetic lamp spectra and line lists to order. The wavelengths in the galaxy image and lamp spectrum run from about 6500 Å on the left to about 7200 Å on the right. This is good to perhaps 1% in wavelength, so you'll need to identify some lines in the lamp spectrum to get a precise value. Enter these limits (or a little wider range) into the site, choose the option to view a plot, and download both. A viewer or image manipulation program will let you expand and contract this plot to match the spectrum you extracted from the lamp frame.

    Save any plots and files from the web site in the project directory. Be sure to save the URL and to cite it properly in your paper according to *ApJ* style.

17. (10 points) Match lines between atlas and lamp reference spectrum. Shift and stretch the reference spectrum with respect to the lamp spectrum until the *positions* of most lines match pretty precisely between the two. The heights of lines from any one gas (Ne or Ar) should be in rough relative agreement with one another between the lamp and atlas (website) spectra, but *this is a secondary consideration* as the sensitivity of the optics varies over the span of the spectrum. Further, there is no reason why the Ar and Ne lines should be at the same relative heights: the height ratio of a Ne to an Ar line might be *dramatically* different between the two datasets. There are a few unidentified lines in both the lamp and the web site spectra. If you have a close match with an identified line, you can use the line's wavelength in the solution. If you are missing an identified line or if something is not quite in vertical proportion, don't worry about it. Do expect a rather precise matchup in the horizontal separations, however.

    If you were calibrating the wavelengths of the whole array, you would use a process similar to the demo from lecture, where you would precisely identify the center of each line, fit a low-order polynomial to the positions *vs.* identified wavelength, and evaluate it at all the $x$ positions.

    For our purposes, however, just take the two most widely separated lines in the lamp reference spectrum, extract a small section that spans the line center and a few pixels on either side of the peak, fit a Gaussian to it, and identify the pixel position of the center of the line precisely. Calculate the ratio of the two lines' separations in wavelength and pixels to get the dispersion in Å/pix. Print the dispersion.

18. (5 points) Calculate velocity. Check that the shift at the galaxy's center is zero (if not, you have an error). Convert the shifts to $\Delta\lambda$, and from that derive the velocity for each spectrum. Plot the velocities.

19. (10 points) Calculate velocity residuals. Fit a polynomial (not too high-order) to the velocities. Calculate and plot the residuals. Find and print their standard deviation. This is your velocity uncertainty.

20. (5 points) Calculate galactic distance. Using the red shift of the galaxy ($\sim$10,100 km/s), and the Hubble Constant ($\sim$67 km/s/Mpc), find and print the distance to the galaxy in Mpc. Is this a reasonable distance to a low-redshift galaxy?

21. (10 points) Calculate galactic radius at each pixel. Find the spatial scale of the Double Spectrograph in arcsec/pix in the FITS header. Using this, the distance to the galaxy, and the small-angle approximation, convert the $y$-pixel locations of your shifts to distances from the galaxy center in kpc. Is this a reasonable scale for a spiral galaxy? Using 0.5 as the radius uncertainty in pixels, derive an expression for the radius uncertainty in kpc, and calculate for each spectrum with a good Doppler shift. Plot the radial distance from the center of the galaxy to the center of each spectrum, with uncertainties.

22. (5 points) Calculate enclosed mass. Recalling that material in spiral galaxies is on roughly circular orbits, estimate the mass of the galaxy interior to each of your tabulated radii. Plot this.

23. (10 points) Calculate uncertainty on enclosed mass as a function of radius. Be sure to use the full error propagation equation applied to a single equation that has all the quantities with uncertainty in it. Remember to derive this equation in your paper! Overplot this on the prior plot.

24. (20 points) Prove dark matter exists. Figure out and carry out a meaningful calculation that establishes the presence of dark matter in this galaxy. It can be done with the data products already calculated, but it requires careful thought due to the nature of the relevant quantities in their current forms. Make plot(s) or print value(s) as appropriate.

25. (10 points) Include a copy of your class log file in your handin. Print the Git log for your main project file into a text file named `project-<username>-prob25-gitlog.txt`. Include the PDF file for the paper in your handin.