Identifying Astronomical Objects Using Spectroscopy
2/20/2022
Mikaela Larkin
mmlarkin@ucsd.edu
Group B: Nazgol Hadaegh, Megan Li, Angelo Romvos, and Joman Wong

# Abstract

Spectroscopy is a widely used form of astronomical study that depends on the interaction of light as it is absorbed and emitted by matter. A physical property of light exists in which specific wavelengths corresponding to the elements present in the source of light display drastic increases or decreases in intensity as the light interacts with matter. This property is used to astronomers' advantage to comprehend the composition of distant astronomical objects. In this lab, we examine the functionality of two single-slit spectrometers, the Ocean Optics USB 2000 handheld spectrograph and the blue arm of the Kast Spectrograph on the Shane 3-meter telescope at Lick Observatory.Both spectrographs measure the intensity of light in the visible range using charge-coupled device (CCD) detectors. Our ultimate goal is to create plots to compare images taken of the galaxy PGC 124 and the F0-type star BD+15 233 by the Kast Spectrograph. In order to do so, we create a routine to find peak intensities in helium using a gas discharge lamp with USB 2000 and then convert from pixel positions in the CCD to wavelength in nanometers using a linear least square best-fit line. We then correct frames taken on Kast for error using bias and flat frames, apply the fit routine to a lamp with a more complex composition on Kast, and then apply the wavelength solution to frames of the two desired sources with unknown composition. Our final results indicate PGC 124 to be hot and transparent, as would be expected of a galaxy with an active galactic nucleus (AGN). Our results also indicate BD+15 233 to be a hot and opaque source with interference from a cooler gas, as would be expected of a star due to the surrounding photosphere. Low standard deviation in our wavelength solution allows for precise identification of elements present in the source objects, allowing for an accurate result of elements present in 3C079 but a less precise result for BD+15 233 due to wide spectral features.

# Introduction

The study of spectroscopy depends on Kirchhoff's laws, which outline the three fundamental types of spectra, or plots of intensity versus wavelength, created by light-emitting sources. The first law states that continuous spectra result from hot, opaque objects in which all present wavelengths are visible. The second law states that hot, transparent objects produce emission line spectra in which only specific lines corresponding to particular elements present in the object appear as peaks. The third law states that hot, opaque objects with some kind of cooler gas such as a photosphere to absorb and reemit specific frequencies produce absorption line spectra in which the wavelength features appear as dips in the plot. Instruments called spectrometers (or spectrographs) are used by astronomers to collect spectra of objects in space in order to assess characteristics such as their composition. In this lab, we will use two different

single-slit spectrometers, the Ocean Optics USB 2000 and the Kast spectrograph, in order to ultimately characterize two astronomical sources.

Single-slit spectrometers such as those used in this lab take in a thin strip of light to restrict the angles at which light enters the device. The light then reaches a collimator, which redirects the incoming light rays to be parallel as they originally were as they radiated out from their source. This light then reaches a diffraction grating or prism which uses the wave nature of light to separate light of different wavelengths out at specific corresponding angles. The next object is a camera, which directs the dispersed light towards defined positions of the detector, which is why there is a direct correlation between pixel position and wavelength according to the device's particular settings. The detector captures the final spectrum for analysis, with each wavelength separated by a factor of the distance from the camera times the change in angle by the diffraction grating or prism. The relationship between the pixel position of peaks and wavelengths for a spectrometer can be found using a lamp with specific elements to produce an output spectrum, which then produces a best-fit relationship between the spectrum and wavelengths. In this lab, we use the method of linear least squares fitting to achieve final wavelength-dependent spectra.

The method of linear least squares fitting for data takes a model of what the desired outcome is, in this case the specific wavelength values, and compares them to what has been observed, which are the centroid values. The difference between these two data sets is then squared and minimized in order to produce a continuous best-fit function. The function maps observed data onto the desired parameters, defining a conversion between pixel number and wavelength. In order to estimate the uncertainties involved in linear least squares fitting, the process of error propagation must be accounted for and a Gaussian distribution must be assumed for the overall standard deviation. The method of linear least squares fitting and the specifics of finding the uncertainty will be further described in the Calculations and Modeling section. In this lab, final spectra of two astronomical sources, the galaxy 3C079 and the F0-type star BD+15 233, were produced according to their wavelength after completing a linear least squares fit for the Kast spectrograph in order to compare their characteristics.

## Observations and Data

In order to create a solution for converting from number of pixels to wavelength, we used data collected by the Ocean Optics USB 2000 spectrometer for various lamps with relevant chemical abundances. The USB 2000 is a handheld, single-slit spectrometer that uses a charge-coupled device (CCD) and a diffraction grating in order to disperse incoming light from the lamps. The CCD is one dimensional with a length of 2048 pixels in order to output data in the form of intensity in analog-to-digital units (ADU) across 2048 corresponding data points. The data points may then be converted to wavelengths using experimentally proven values for the chosen element's absorption lines with a best-fit comparison. The data points are applied back to the corresponding intensities using the best-fit function, creating a resultant plot showing the spectrum of the observed elements with respect to wavelength for further analysis. The USB 2000 has a limited range of wavelengths it may detect between 336 nm and 1027 nm in order to observe visible light and some lower wavelength infrared light. It is therefore an applicable basis to create a linear least squares fit for later application to observational data of visible astronomical sources taken with the Kast spectrograph.

The USB 2000 uses a standard setup for a single-slit spectrograph in a small box which connects to input lamps and to a computer with software called SpectraSuite control from Ocean Optics to set up lamps and collect desired spectra. The device continuously displays input data on the connected computer and the user is able to capture specific spectra as desired. The type of data collected for our analysis was set to be the raw intensity counts at each pixel for later application to the linear least squares fit. The x-axis can be set to wavelength automatically, but for our purposes of inspecting the accuracy of a linear least squares fit, we used pixel number-dependent data. The exposure time was set to be longer than the default, 100 ms, and the scale was set for the best visibility of the spectrum. About 100 spectra were collected for each of the four gas discharge lamp elements, hydrogen, helium, mercury, and neon, using the SpectraSuite feature to collect multiple spectra simultaneously for each element. We used helium in particular for our analysis as is explained in the Data Reduction and Methods section. SpectraSuite has processing options in order to correct for systematic error, but our data has this feature turned off in order to analyze raw data. Because of this choice, our results do not account for bias, or "dark" error from potentially oversaturated pixels of the detector. Our results also do not account for "reference" error, which is error from any component of the spectrograph found when well known spectra are produced with errors compared to what the spectrum should look like in theory. Error in our final results may be attributed to the lack of accounting for these errors in the device's functionality.

The Kast spectrograph is a device located at the Lick Observatory in Northern California on the Shane 3-meter Telescope. In particular, we used the blue arm of this device, which is calibrated to be sensitive towards blue light rather than red light. The Kast spectrograph is a single-slit spectrometer similar to USB 2000, but it makes use of a disperser that is a mix between a diffraction grating and a prism, called a "grism," rather than a simple diffraction grating. The blue arm has its own CCD detector, which was used to collect desired frames on October 26, 2013. The types of frames are in the form of bias images taken with no exposure time, dome flat field images taken of the inside of the observatory dome for uniform illumination of the detector, arc lamp images of known composition to compare to source emission lines, and images taken of several desired scientific targets. We correct for potential systematic anomalies using the various types of frames taken, so there is less opportunity for device-related error than with the USB 2000. The arc lamp used in this lab is composed of helium, mercury, argon, and neon. The two scientific targets chosen are the galaxy 3C079 and the F0-type star BD+15 233. The summary of all data used in this lab may be found in the table below.

| Device used | File Name | Date collected | Personnel | Purpose |
|---|---|---|---|---|
| **USB 2000** | groupc-Helium02197.txt | Winter quarter 2020 | Winter 2020 Group C | Helium spectrum text file |
| **Kast (Blue)** | r101.fits | October 26, 2013 | Lick staff | Arc lamp frame |
| **Kast (Blue)** | r107.fits | October 26, 2013 | Lick staff | Bias frame |
| **Kast (Blue)** | r124.fits | October 26, 2013 | Lick staff | 3C079 frame |
| **Kast (Blue)** | r131.fits | October 26, 2013 | Lick staff | Dome flat frame |

| Device used | File Name | Date collected | Personnel | Purpose |
|---|---|---|---|---|
| **Kast (Blue)** | r133.fits | October 26, 2013 | Lick staff | BD+15 233 frame |

## Data Reduction and Methods

Data collected using the Ocean Optics USB 2000 Spectrometer provided the mathematical basis for analyzing astronomical data in this lab. The first step in the process was to produce a plot comparing the pixel numbers to their corresponding intensities in ADU using Python on JupyterHub. We chose to analyze data collected from the gas discharge lamp that we found to have the most cohesive result, which was helium. The helium spectrum, shown in **Figure 1** in the Calculations and Modeling section, displays emission lines as peaks in the plot. These peaks are the defining feature to utilize in order to achieve our ultimate goal of mapping the data from the pixel number-dependent spectrum onto a wavelength-dependent spectrum. In order to find the exact pixels at which each maximum occurs, we automated the process using centroids, which are a weighted average of the pixel values that lie within the peak. We chose an arbitrary value of 2x the mean intensity of the spectrum in order to define the minimum height requirement for peaks to be considered in our calculations. The code for automating centroids calculations may be found in the Appendix. Using nine experimentally proven emission line wavelengths of helium in the visible spectrum, we then mapped our nine calculated centroid values to these wavelengths using the method of linear least square fitting. This best-fit line is used to create the conversion between pixel values and wavelengths, which may later be applied to spectra with unknown peak wavelength values.

Data taken from Kast was first prepared for analysis by subtracting the bias frame to remove intrinsic bias from the detector, normalizing using the dome flat frame to account for differences in quantum efficiency across pixels and path differences across the detector, and removing both saturated pixels and overscan regions of the CCD. We then applied the centroid-identifying routine to the arc lamp spectrum, finding a grand total of 41 peaks. We then used researched wavelength values for helium, mercury, argon, and neon, combined into a single array in numerical order, to create a linear least squares fit for Kast. We found the results to have very low standard deviations due to the large number of input data points. To produce our final results, we applied the wavelength solution to our arc lamp and two science frames. The finalized science spectra show an emission line spectrum for the galaxy 3C079 and an absorption line spectrum for the F0-type star BD+15 233. This is the result we would expect according to Kirchhoff's second and third laws. The second law distinguishes hot transparent objects, such as galaxies with active galactic nuclei (AGN), to produce emission line spectra. The third law distinguishes hot opaque objects that are located behind cool thin gases in space, such as distant stars, to produce absorption spectra.

# Calculations and Modeling

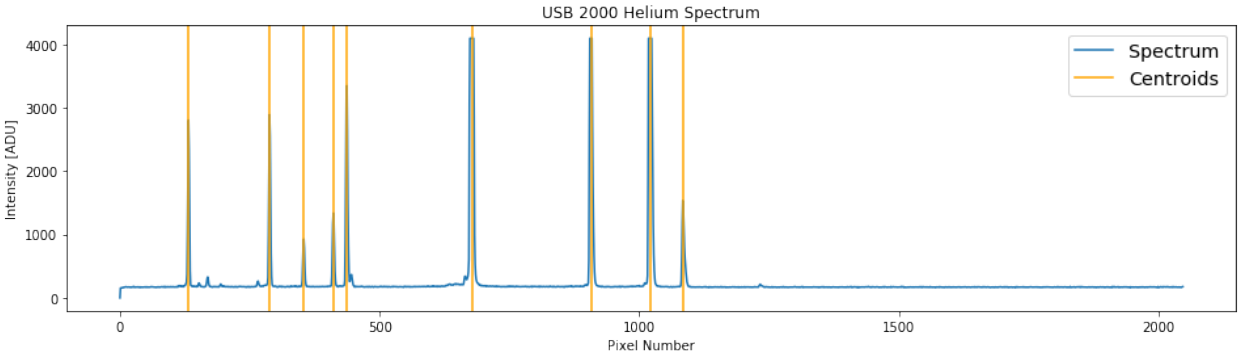



**Figure 1**: Blue spectrum of helium using intensity in ADU and pixels numbers through 2048 from CCD detector on the Ocean Optics USB 2000 Spectrometer with orange lines of centroid values over each peak.

The first stage of calculation for this lab was to take raw data of intensity in ADU at each pixel number from the Ocean Optics USB 2000 device and use it to prepare a conversion to wavelength. After assessing the results from the hydrogen, helium, mercury, and neon gas discharge lamps, we found hydrogen and helium to be the most clearly defined elements with no overlapping spectral lines. We chose helium as the basis for our wavelength solution due to the appearance of nine strong emission lines, compared to hydrogen's three emission lines, with the prediction that more input data points would produce a more precise result. The initial helium spectrum is shown by the blue line in **Figure 1**. From this spectrum, we wanted to find the centroids of each peak, meaning the exact pixel location at which the intensity peaks. Centroid values were calculated according to the equation $p_c = \frac{\sum_{i=1}^{N} p_i I_i}{\sum_{i=1}^{N} I_i}$ in which $p_c$ represents the pixel value for centroids, $p_i$ represents each pixel in the peaks, and $I_i$ represents each intensity value in the peak ADU. In order to automate the process of calculating centroid values, we created a mask to calculate intensity values above an arbitrary threshold of double the mean intensity and appended arrays of each value that lies within the peak. These values were then looped over according to the centroid equation and appended to an array of all centroid values. The code used for the centroid-finding process may be found in the Appendix. Once these centroid values were calculated, we plotted them over the initial spectrum to ensure the success of the code as shown by the orange lines in **Figure 1**. Error propagation in the form $\sigma_{p_c}^2 = \frac{\sum_j I_j (p_j - p_c)^2}{(\sum_i I_i)^2}$ was used to find standard deviation of each centroid.

Using the calculated centroid values as inputs, we then created a solution to map the spectrum onto wavelength-dependent axes using the method of linear least squares. This analytical method assumes a linear relationship between pixel number and wavelength with a best-fit line following the standard equation $y = mx + c$, with $y$ representing the wavelength in nanometers, $m$ representing slope, $x$ representing the pixel values, and $c$ representing the y-intercept. We defined the slope and intercept of the fit using the matrix equation

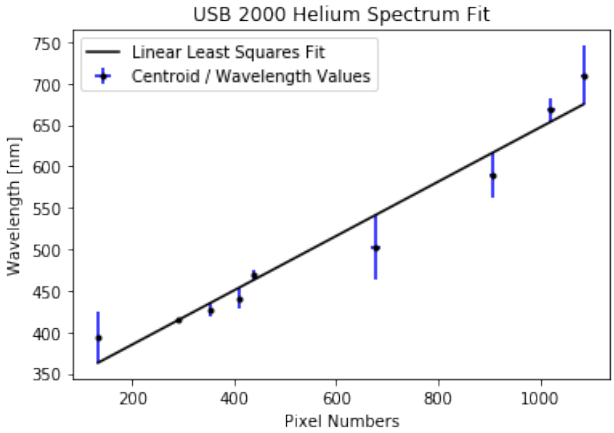



**Figure 2**: Comparison of calculated least square fit line in black to the centroid pixel position from the Ocean Optics USB 2000 Spectrometer and equivalent researched wavelengths in nanometers. Small blue error bars on the points in the x-direction indicate standard deviation in the centroid-finding routine and more significant blue error bars on the points in the y-direction indicate standard deviation of points from the fit line.

$\begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} p_i^2 & \sum_{i=1}^{N} p_i \\ \sum_{i=1}^{N} p_i & N \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{N} p_i \lambda_i \\ \sum_{i=1}^{N} \lambda_i \end{pmatrix}$, in which $m$ and $c$ still represent the slope and y-intercept respectively , $p_i$ represents each pixel in the original peaks, $N$ represents the total number of centroid values, and $\lambda_i$ represents each expected wavelength value. The $\lambda_i$ values were found from previous experimentation finding the peak emission lines of helium to occur at nine specific wavelength values, aligning with our nine calculated centroid values. Carrying out this calculation yielded a relatively close best-fit line, as shown in **Figure 2**. The final wavelength-dependent spectrum of helium may be found in **Figure 3**.

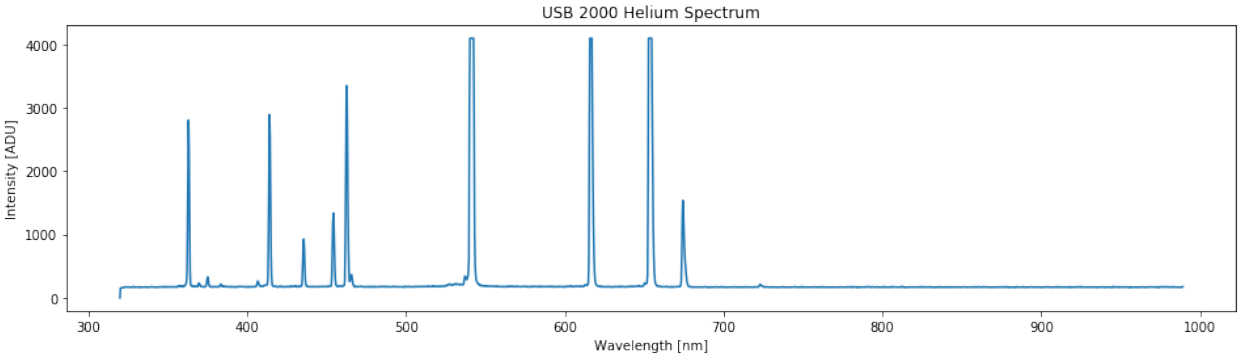
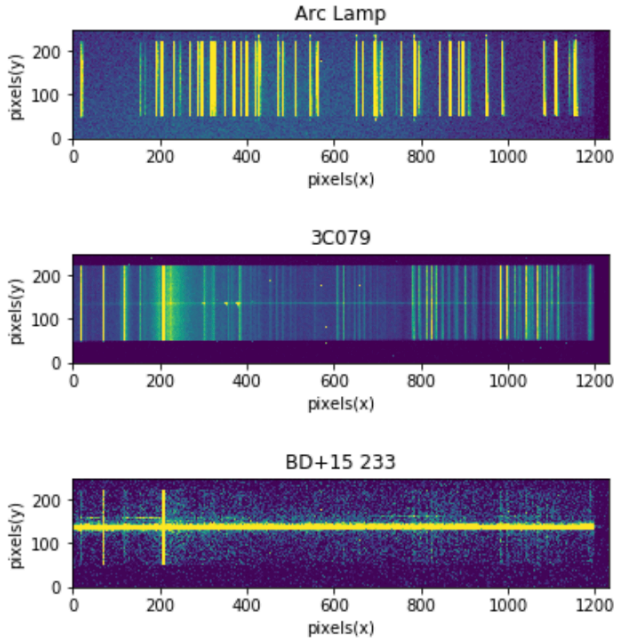



**Figure 3**: Final spectrum of helium for the USB 2000 Spectrometer using wavelength in nanometers on the x-axis and intensity in ADU on the y-axis.

The calculation of how much standard deviation there was associated with this best fit calculation had to take into account the process of error propagation since we were solving for two unknown variables, slope and y-intercept. With the understanding that there is no covariance, meaning that measurements slope and y-intercept are not dependent on measurements of wavelength, the equations for each of their standard deviations are $\sigma_m^2 = \sum_j^N (\frac{\partial m}{\partial \lambda_j})^2 * \sigma_j^2$ for slope and $\sigma_c^2 = \sum_j^N (\frac{\partial c}{\partial \lambda_j})^2 * \sigma_j^2$ for y-intercept. In these equations, $\lambda_j$ represents each wavelength and $\sigma_j$ represents the overall standard deviation of the fit. We expect this overall standard deviation to be Gaussian, meaning it follows the equation

$\sigma^2 = \sum_j \sigma_j^2 = \frac{1}{N-2} \sum_i (\sigma_i - (mp_i + c))^2$ with each variable having the same meaning as the previously defined functions. The factor of $N-2$ in this equation comes from the fact there are two unknown values and therefore two degrees of freedom. The value we calculated for overall standard deviation for the USB 2000 helium spectrum was approximately 715.7 nm. From the original matrix equation, the derivatives can be expanded to give final equations for each standard deviation: $\sigma_m^2 = \frac{N\sigma^2}{N\sum_i^N p_i^2 - (\sum_i^N p_i^2)^2}$ and $\sigma_c^2 = \frac{\sigma^2 \sum_i^N p_i^2}{N\sum_i^N p_i^2 - (\sum_i^N p_i^2)^2}$. For the USB 2000 Helium spectrum, the standard deviation of the slope was calculated to be approximately $7.5*10^{-4}$ nm, and the standard deviation of the intercept was found to be approximately 341.6 nm.




**(A)**

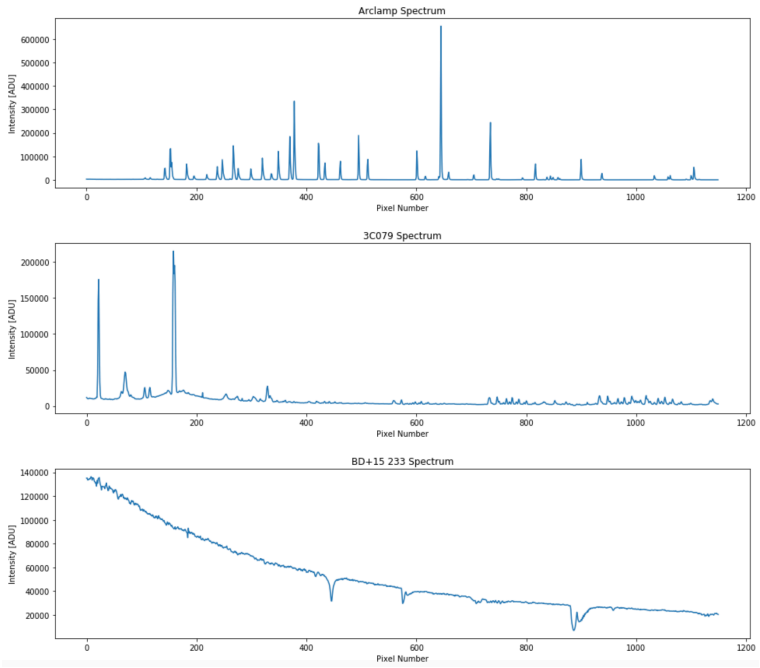
**(B)**

**(C)**

**Figure 4**: Images of the three desired frames for analysis from the Kast Spectrograph after bias subtraction and dome flat normalization. Image **(A)** displays the arc lamp with spectral lines for helium, mercury, argon, and neon. Image **(B)** displays the spectral lines for the galaxy 3C079. Image **(C)** displays the spectral lines for the star BD+15 233. The resolution of each image was specified through trial-and-error for the most clear final data for analysis. The colorization of each image highlights the horizontal slit of the spectrograph and spectral lines in a yellow/ green color.

Once the linear least squares method was defined and shown to be accurate for USB 2000, we read in five data files collected using the Kast Spectrograph. The frames taken were one bias frame, one dome flat frame, one arc lamp frame, and two science frames. The two science frames were one of the galaxy 3C079 and one the F0-type star BD+15 233. The data was prepared for reduction by first subtracting the bias frame from each of the other images to remove the detector's intrinsic bias. We set the science and arc frame intensities that were above a threshold of 65000ADU to equal 1ADU in order to account for saturated pixels in the detector. The next step was normalizing dome flat field frames to create a normalization factor array according to the equation $F_{norm} = \frac{flat - bias}{median(flat - bias)}$. We used another correction in the normalization factor array by setting all values equal to 0 to equal 1 in order to avoid dividing the science and arc frames by zero in the next step. The arc lamp and science frames were then divided by the corrected normalization factor in order to account for the varying quantum efficiency of each pixel and for path differences across the detector. The finalized images of the

arc lamp and two science spectra after correction are shown in **Figure 4**. We tested these images a few times in order to find proper color normalization to resolve each individual image using code shown in the Appendix. Examining the images in **Figure 4**, we found the pixels in the columns below the value 50 and above 1200 to be extraneous as the CCD's overscan region, so we excluded from the range used in our arc lamp and science frame analysis. We also narrowed the y-axis down to a small, central range of 30 rows of pixels between 135 and 165 where the slit of Kast occurs.




**(A)**

**(B)**

**(C)**

**Figure 5**: Spectra using pixel values from the Kast Spectrograph on the x-axis before applying wavelength solution and intensity in ADU on the y-axis. Plot **(A)** displays the spectrum of the chosen arc lamp. Plot **(B)** displays the spectrum of the galaxy 3C079. Plot **(C)** displays the spectrum of the star BD+15 233.

The final portion of our analysis was to apply our centroid identifying and linear least squares routines developed with USB 2000 to the Kast arc lamp spectrum for application to the two science frames. The spectra of the arc lamp, 3C079, and BD+15 233 compared to their original pixel values may be found in **Figure 5**. The chosen arc lamp from Kast consists of helium, mercury, argon, and neon, which have 9, 6, 11, and 15 emission lines respectively for a grand total of 41 centroids. Some of these emission lines are faint in comparison to others, so our

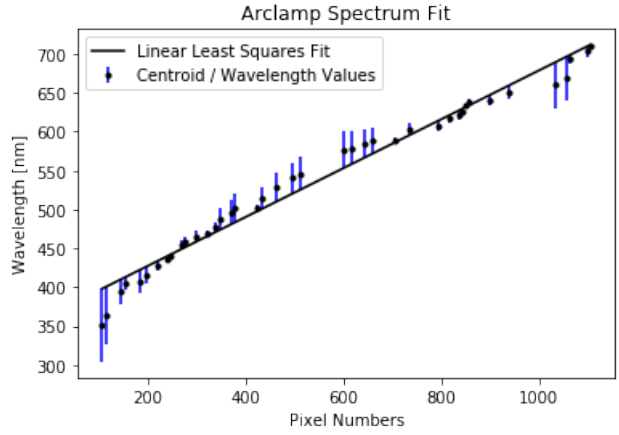



**Figure 6**: Comparison plot between linear least squares fit in black for the centroid pixel position from the Kast spectrograph and equivalent researched wavelengths of helium, mercury, argon, and neon emission lines in nanometers. Small blue error bars on the points in the x-direction indicate standard deviation in the centroid-finding routine and more significant blue error bars on the points in the y-direction indicate standard deviation

centroid identifying routine was adapted for a minimum peak height requirement to simply be above the mean intensity rather than double the mean. The 41 researched emission line wavelengths were then combined in numerical order and run through our linear least squares routine to produce the best-fit shown in **Figure 6**. The standard deviations calculated were less than those for helium from USB 2000, with final values of approximately 263.7 nm for overall Gaussian standard deviation, $6.7*10^{-5}$ nm for the slope, and 26.9 nm for the y-intercept. The finalized, wavelength-dependent spectra of the arc lamp, 3C079, and BD+15 233 may be found in **Figure 7**. The centroid routine was then applied to the final wavelength-dependent spectrum of 3C079 in order to find spectral lines and characterize the galaxy.

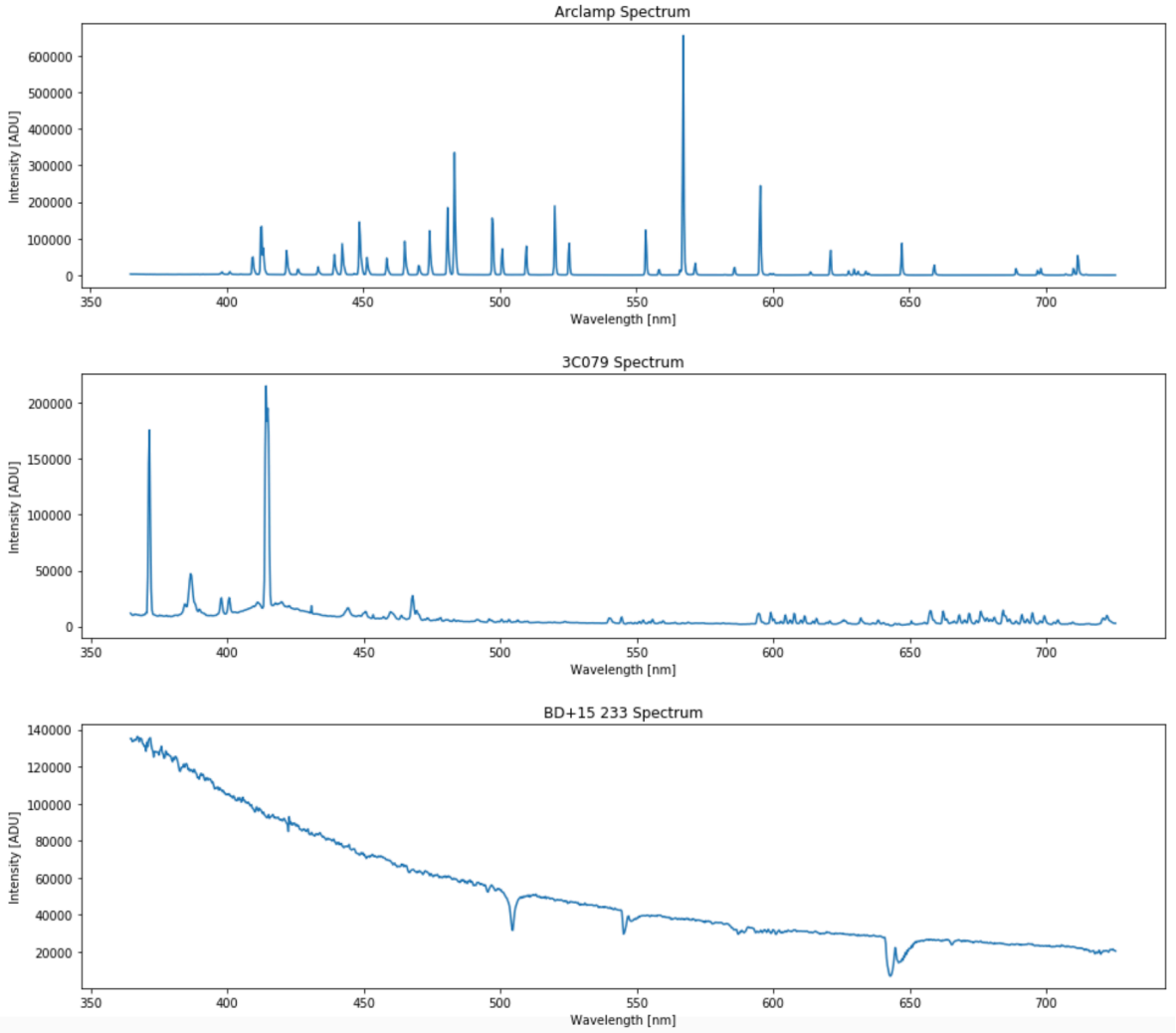



**Figure 7**: Final spectra with wavelength in nanometers on the x-axis and intensity in ADU on the y-axis for reduced frames of **(A)** the arc lamp in Kast, **(B)** the galaxy 3C079, and **(C)** the star BC+15 233.

## Discussion

The final plots of 3C079 and BD+15 233 indicate an emission line and absorption line spectrum, respectively. 3C079 is a Seyfert galaxy. Seyfert galaxies are characterized by having

an active galactic nucleus (AGN) at their core, making the galaxies hot and bright sources of electromagnetic radiation. In particular, 3C079 is a type II Seyfert galaxy, meaning it has strong emission lines as is apparent in our final plot of its spectrum in **Figure 7B** at the more blue end between about 370 nm and 450 nm. The final resultant emission line spectrum aligns with the theoretical expectation Kirchhoff's second law, which is what would be expected of this type of galaxy since the dispersion of gas and dust creates a transparent source, and the AGN provides a high temperature. There are three major spectral features in the spectrum we chose to assess for the chemical composition at the wavelengths 371.4 ± 0.4 nm, 386.6 ± 0.3 nm, and 414.6 ± 0.6 nm. These wavelengths are likely associated with the elements vanadium, iron, and another line of vanadium, respectively. The elements in the final spectra are not present in the arc lamp used, but the wavelength solution should theoretically be applicable to all elements. The application allows for a reliable analysis of these features due to low standard deviation..

The plot of BC+15 233 in **Figure 7C**, on the other hand, is an absorption line spectrum. BC+15 233 is an F0-type star, which is only slightly more luminous and massive than the sun, and lies in the next spectral classification above GV of the sun. Compared to all types of stars in the universe, F0 stars are not very bright so they are difficult to detect at far distances. Because of the absorption features, our original centroid routine could not be applied, and the three main features had to be estimated manually using local minima. Using a local minima loop shown in the Appendix resulted in many different centroids, but the main three occur at 504.9 ± 131.9 nm, 545.5 ± 169.8 nm, and 643.1 ± 259.6 nm. These wavelength values are much less reliable than the values found for 3C079 due to the less precise routine and the wide absorption features in the spectrum. The large standard deviation makes it unclear which definitive lines are present in the spectrum. There is likely interference from other objects in space. There are various possible causes of broadening of lines in the spectrum as light approaches the Shane 3-m telescope from far distances in space, causing BC+15 233 to remain unclassified until these anomalies can be understood and addressed in future research. However, we still find our methodology to be sound for emission line spectra.

The implementation of spectroscopy in this lab has allowed for analysis of distant astronomical sources with unproven chemical compositions. Spectra in general are a powerful tool in astronomy through comparison to experimentally proven features of elements to better understand what exists in space. The results for well defined spectra from this method are trustworthy because of the mean standard deviation from the fit of about 0.6 centroid pixel values and of about 12 nm for the arc lamp spectrum. These values are less than the values from the USB 2000, which are 3.9 pixel values and 19.7 nm respectively, most likely due to the device error corrections using bias and dome flat frames for Kast data. It is also possible that the more precise results come from the use of an arc lamp with four elements rather than one, meaning the results would be more accurate if a more chemically diverse lamp was connected to the Kast blue arm. Results could also be improved in the future by using more precise wavelength inputs for the solution, particularly from one unified scientific source rather than separate experiments. Calibration for absorption line spectra could also be greatly improved by accounting for sources of line broadening. A final improvement in future experimentation could be the use of a more complicated best-fit line than linear least squares, such as a non-linear fit, in order to account for the varying responses of the CCD at different wavelengths, rather than assume a 1-to-1 relationship to pixel number.

# Appendix

Our lab group again operated through group messaging as questions arose in our individual analyses. Joman contributed researched helium wavelengths to produce a best-fit, but I elected to use a different primary source. I contributed my centroid finding routing using a mask and values above double the mean of the spectra and Megan contributed an alternative using local extrema. I used a version of Megan's alternative for the absorption line spectrum of BD+15 233 since my routine of finding peaks was not applicable to this shape. The other group members were unsure how to properly code the linear-least squares fit, so I also contributed my code for accomplishing this goal. Megan contributed images of the bias frames from Kast, but I elected not to include these in my report. Joman discussed various versions of how final plots should look with colors and amounts of information per plot. Angelo shared his arc lamp plot and discussed having unsupported data types.

The source used for helium emission lines in the visible spectrum used in order to create the linear least squares wavelength solution for USB 2000 were found in https://edisciplinas.usp.br/pluginfile.php/4822251/mod_resource/content/1/Espectro%20do%20helio%20-%20presto.pdf.

The elements present in the arc lamp spectrum of Kast were fit using helium lines from the same source as USB 2000, from https://en.wikipedia.org/wiki/Ion_laser for argon, and from http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/atspect2.html for neon and mercury.

Information about 3C079 was found at http://simbad.u-strasbg.fr/simbad/sim-basic?Ident=3C079&submit=SIMBAD+search and https://www.britannica.com/science/Seyfert-galaxy. Element comparisons were found at https://physics.nist.gov/PhysRefData/ASD/lines_form.html.

Information about BD+15 233 was found at http://simbad.u-strasbg.fr/simbad/sim-basic?Ident=BD%2B15233&submit=SIMBAD+search and https://en.wikipedia.org/wiki/F-type_main-sequence_star.

Code for identifying and appending centroids in emission line spectra:

```
mew = np.mean(int_He)
mask = np.where(int_He > mew*2, 1, 0)  # Using a mask to define peaks as having a mininmum height requirement of 2x
                                       # mean intensity in order to specify only the most prominent spectral features
pix_masked = []
tmp_peak = []
for i, mask_val in enumerate(mask):
    if mask_val == 1:
        tmp_peak.append((pix_He[i], int_He[i]))  # List of pixel values and intensities
    elif len(tmp_peak) > 0:
            pix_masked.append(tmp_peak)        # Appending values where peak exists
            tmp_peak = []
```

```
centroids = []
sig_p = []

# Creating an array of centroid values and on of associated erroe for each peak according to equations shown below

for peak in pix_masked:
    numerator = sum([a*b for a, b in peak])
    denom = sum([b for a, b in peak])
    x_c = numerator / denom
    centroids.append(x_c)

for i, peak in zip(centroids, pix_masked):
    errnum = sum([b*(a-i)**2 for a, b in peak])
    errdenom = sum([b for a, b in peak])
    err = errnum / errdenom
    sig_p.append(np.sqrt(err))
```

Code for linear least squares fit:

```
N = len(centroids)     # Number of centroid values for N input below

M1 = np.array([[np.sum((centroids)**2), np.sum(centroids)],[np.sum(centroids), N]])  # First matrix on right hand side
M2 = np.array([[np.sum(centroids*wl_He)],[np.sum(wl_He)]])  # Second matrix on right hand side
```

```
M1_inv = np.linalg.inv(M1)  # Inverse of first matrix
M = np.dot(M1_inv, M2)      # Dot product of two matrices for final result matrix
```

```
mfit = M[0,0]                          # Slope value from final matrix
cfit = M[1,0]                          # Intercept value from final matrix
y = [mfit*a + cfit for a in centroids]  # Linear least squares best-fit line
```

Code for correcting frames and plotting **Figure 4** images from Kast Spectrograph:

```
fnames = ['r107.fits', 'r131.fits', 'r101.fits', 'r134.fits', 'r133.fits']  # File names

bdata = fits.getdata(fnames[0])            # Extracting data from FITS files
fdata = fits.getdata(fnames[1]) - bdata    # Subtracting bias frames from each of the other three types
adata = fits.getdata(fnames[2]) - bdata
s1data = fits.getdata(fnames[3]) - bdata
s2data = fits.getdata(fnames[4]) - bdata
```

```
adata[adata > 65000] = 1    # Correcting for saturated pixels in arclamp and science frames
s1data[s1data > 65000] = 1
s2data[s2data > 65000] = 1
```

```
fnorm = fdata / np.median(fdata)  # Creating normalization factor
fnorm[fnorm == 0] = 1             # Correcting for 0's in normalizations factor to avoide dividing by 0
```

```
# Plotting images of final science and arclamp frames
# vmin and vmax for best resolution found by trial-and-error

aflat = adata / fnorm
plt.figure()
plt.imshow(adata,origin='lower',interpolation='nearest',cmap='viridis', vmin=0, vmax=200)
plt.title('Arc Lamp')
plt.xlabel('pixels(x)')
plt.ylabel('pixels(y)')

s1flat = s1data / fnorm
plt.figure()
plt.imshow(s1data,origin='lower',interpolation='nearest',cmap='viridis', vmin=0, vmax=750)
plt.title('3C079')
plt.xlabel('pixels(x)')
plt.ylabel('pixels(y)')

s2flat = s2data / fnorm
plt.figure()
plt.imshow(s2data,origin='lower',interpolation='nearest',cmap='viridis', vmin=0, vmax=50)
plt.title('BD+15 233')
plt.xlabel('pixels(x)')
plt.ylabel('pixels(y)')
```

Code and results for standard deviations associated with Kast wavelength solution:

```
sig = np.sqrt(1/(N-2) * np.sum((wl_tot - y)**2))
print(sig**2)  # Overall Gaussian standard deviation of fit
```

```
263.7103667315316
```

```
sig_M = np.sqrt((N*sig**2) / ((N*np.sum((centroids)**2))-(np.sum(centroids))**2))
print(sig_M**2)  # Slope standard deviation
```

```
6.703785919071918e-05
```

```
sig_c = np.sqrt((sig**2*np.sum((centroids)**2)) / ((N*np.sum((centroids)**2))-(np.sum(centroids))**2))
print(sig_c**2)  # Intercept standard deviation
```

```
26.95653351788386
```

Code for absorption line spectra dips adapted from Megan's centroid-finding routine:

```
raw_mins = []  # Finding minima in BD+15 233 spectrum
sig_p = []
index = 0

for i in np.diff(np.sign(np.diff(s2spec))):
    index += 1
    if i > 0 and s2spec[index - 10] > s2spec[index + 1] < s2spec[index + 4]:
        raw_mins.append(wl_final[index + 1])

for i, a, b in zip(raw_mins, wl_final, s2spec):
    errnum = sum([b*(a-i)**2])
    errdenom = sum([b])
    err = errnum / errdenom
    sig_p.append(np.sqrt(err))

raw_mins = np.array(raw_mins)  # Converting centroids to an array for easier numpy operations
sig_p = np.array(sig_p)
```