## Lab 2: Astronomical Spectroscopy

Jessica Birky, Julian Beas-Gonzalez, Russell Van-Linge

### ABSTRACT

In this lab we determine the wavelength calibration for two different spectrographs: the Ocean Optics USB 2000 fiber optic spectrograph, and the KAST spectrograph mounted at Lick Observatory. Using spectra of several gas lamps (Helium, Neon, and HeHgCd), we compare the pixel values of emission centroids to theoretical emission wavelengths to determine pixel-to-wavelength conversion solutions of $\lambda = .3568(\text{pixel}) + 344.21$ nm for the Ocean Optics instrument and $\lambda = .1549(\text{pixel}) + 329.25$ nm, using linear least-squares regression. We also compute the errors for each emission peak, and find that the error roughly decreases by a $1/\sqrt{n}$ trend, with higher lower intensity peaks having higher errors. Finally, applying our calibration to spectra from several different light sources (incandescent and fluorescent bulbs) and astronomical sources (the Sun, BD+15233, Feige 110, and J0047+0319) we draw conclusions about how these spectra were formed based off of Kirchoff's laws of spectral formation.

## 1. INTRODUCTION

A spectrum measures the flux of photons coming from a source as a function of wavelength. Kirchoff's laws of spectral formation characterizes the three different ways a spectrum can be created: (1) a hot opaque body or dense gas (such as a black body) will produce a continuous spectrum; (2) a hot transparent, low-density gas will produce an emission line spectrum; and (3) a cool, transparent, low-density gas will produce an absorption line spectrum.

In Lab 1 we learned about the properties of charged coupled device (CCD) detectors, and how we can measure the number of photon counts over two spatial dimensions. This lab add several other layers of instrumentation which allow us to disperse light taken over a narrow slit across a region of wavelengths, producing a two-dimensional (1 spatial dimension × 1 wavelength dimension) image of a source. The key component behind spectroscopy is the addition of a diffraction grating (or grating prism called a grism), shown in Figure 1, which causes rays of light at different wavelengths to be split at different angles, as given by the grating equation:

$$\frac{m\lambda}{d} = \sin\alpha + \sin\beta \tag{1}$$

In particular, this lab seeks to classify several different types of spectrum taken from every day sources (such as incandescent/fluorescent lights and the Sun), gas lamps (such as Helium, Neon, and HeHgCd gas), and several astronomical sources. Section 2 covers the particular schematics and
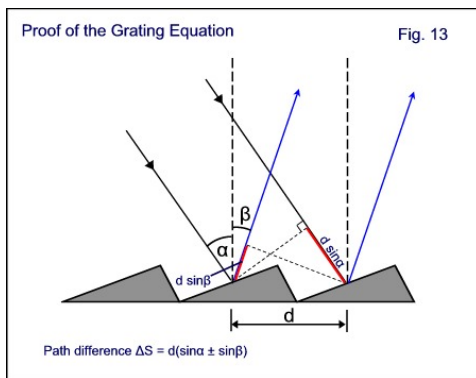
Corresponding author: Jessica Birky (A13002163)
jbirky@ucsd.edu

**Figure 1.** Diagram of a diffraction grating. Light rays hit the grating with angle of incidence $\alpha$ and are reflected with angle $\beta$. The spacing between the grooves of the grating is $d$.

specifications of two different spectrographs: the Ocean Lab spectrometer (which we use to take the spectra of every day sources), and the KAST spectrograph mounted on the Shane 3m Telescope on Lick Observatory (which we use to study astronomical sources).

## 2. OBSERVATIONS

### 2.1. *Spectrographs*

The KAST spectrograph is a double slit spectrograph which has two different observing modes: a red arm with 7 different grisms and a blue arm with 3 different grisms. The image to the right of Figure 2 shows the schematic of the instrument: (1) light from an astronomical source enters the slit at point a, (2) a filter wheel and beamsplitting mirror determine if the light will pass on to the blue or red detector at points b and c, (3) the light goes through a collimator which receives angled light rays and returns parallel rays, (4) a grating or grism is used to disperse light into its separate wavlength components, and (5) the light passes through a camera and finally a CCD detector which counts the number of photons collected in Analog to Digital Units (ADU).

The Ocean Optics spectrograph is a simpler version of the KAST spectrograph: (1) light comes into the entrance aperature at point 1 via a fiber-optics cable, (2) a collimating mirror reflects the light into parallel beams at point 4, (3) the diffration grating at point 5 splits the light into separate wavelenths, (4) the the camera mirror reflects the light onto the CCD detector at point 8.

### 2.2. *Data Collection*

Using the Ocean lab spectrometer we collected spectra of five different sources: the sun, an incandescent light bulb, a fluorescent light bulb, a helium gas lamp, and a neon gas lamp, as well as bias frames taken in a dark room. For each source we colleced 100 frames of data.

We also analyzed several astronomical spectra taken with KAST in 2013. Reduction for the KAST spectrograph requires three types of frames: the scientific frames, as well as bias frames (a zero second exposure capturing instrumental noise), and flat field frames (measuring the pixel-to-pixel response from the detector due to variations in quantum efficiency or optical path differences in the system) which are used for calibration. Table 1 shows the observation log of frames used from the Shane public data repository.
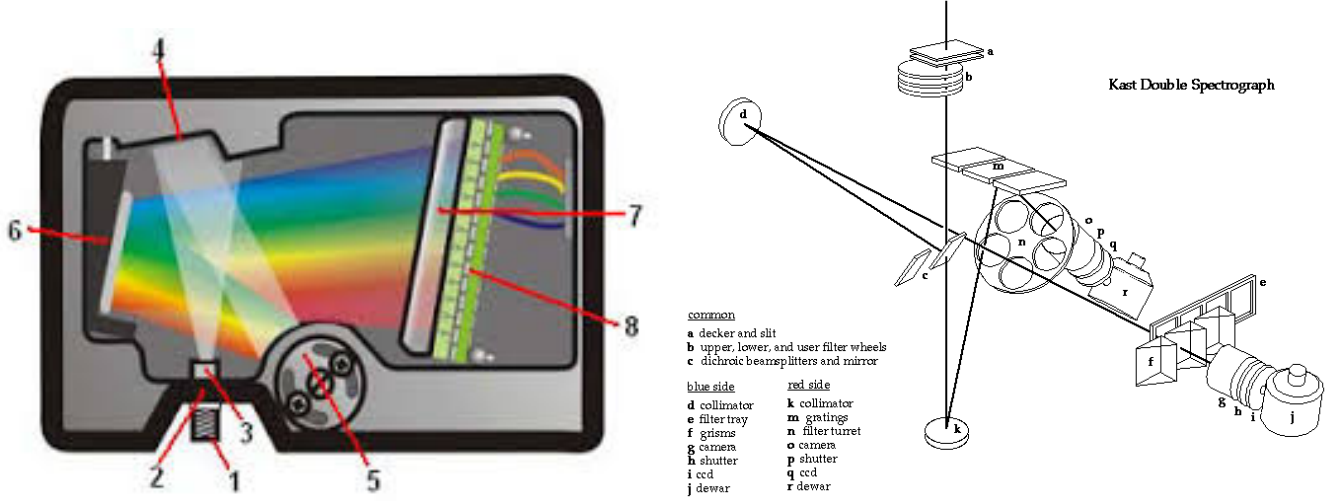
**Figure 2.** Schematic of two spectrographs: Ocean Optics USB 2000 (left) and KAST (right).

| source | RA | DEC | frame # | # frames | exposure time (sec) |
|---|---|---|---|---|---|
| HeHgCd arclamp | – | – | 100 | 1 | 40 |
| bias | – | – | 101-110 | 10 | 0 |
| flat field | – | – | 121-150 | 30 | 90 |
| J0047+0319 | 0:47:34 | 3:25:55 | 155 | 1 | 1200 |
| BD+15233 | 1:34:31 | 15:51:42 | 156 | 1 | 90 |
| Feige 110 | 23:20:35 | -5:02:03 | 160 | 1 | 120 |

**Table 1.** Observation log of bias, flat field and science frames taken on KAST. Frame 100 arclamp is used for wavelength calibration, frame 101-110 bias frames are used for subtracting instrument noise, frame 121-150 dome flat frames are used for calibrating detector sensitivity at each pixel, and frames 155, 156, and 160 are three different astronomical sources.

## 3. DATA REDUCTION & METHODS

### 3.1. *Reduction*

For the Ocean Lab spectra, using the same procedure as Lab 1, we combined all 100 bias frames by taking the median at each pixel. Then for each frame of the solar, lamp spectras we subtracted the bias and combined into one reduced frame (lines $4 - 19$, Section 8.3).

Reducing the KAST data to extract 1D spectrum arrays required several steps: (1) we combined all 10 bias frames (listed in Table 1) by taking the median value at each pixel (lines $15 - 19$ Section 8.3); (2) we computed the normalized flat image (lines $21 - 33$):

$$\mathrm{Flat_{norm}} = (\mathrm{Dome} - \mathrm{Bias})/\mathrm{Median}(\mathrm{Dome} - \mathrm{Bias}) \tag{2}$$

(3) we bias-subtracted the arclamp, dome flat, and science images, and divided each by the normalized flat field (lines $35 - 49$); (4) we sliced the 2D image along the spatial dimension to extract the region that contains the source and took the average across the spatial dimension to produce a 1D spectrum

(photon counts in ADU vs. wavelength pixel index) shown in lines $51-62$. The final bias-stubtracted, normalized science images for BD+15233, Feige 110, and J0047+0319 are shown in 3.

### 3.2. *Wavelength Calibration*

The spectrograph outputs the photon counts as a function of wavelength in arbitrary pixel units, however we want to know the spectrum as a function of its electromagnetic wavelength. In order to do this we must come up with a function that converts pixel to wavelength. This is why we collect emission spectra of several gas lamps: if we know the pixel values of each emission peak, and we can theoretically compute the wavelength produced by an electron transition, then we can match pixel peaks to wavelength peaks and calibrate a solution.

In this lab, we test a first-order polynomial fit ($y = mx + c$), which as we will see in the next section looks to describe the correlation well. Using least-squares regression we can fit the slope and intercept between our pixel array ($x$) and our wavelength array ($y$), each which contain $N$ elements. The best fit slope, intercept and fit errors are complete linear algebra operations to compute. Lines $1 - 20$ of Section 8.2 show the implementation of the following parameter and error equations:

$$\begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & N \end{bmatrix}^{-1} \begin{bmatrix} \sum x_i y_i \\ \sum y_i \end{bmatrix} \tag{3}$$

$$\sigma^2 = \frac{1}{N-2} \sum_i [y_i - (mx_i + c)]^2 \tag{4}$$

$$\sigma_m^2 = \frac{N\sigma^2}{N \sum_i x_i^2 - \left(\sum_i x_i\right)^2} \tag{5}$$

$$\sigma_c^2 = \frac{\sigma^2 \sum_i x_i^2}{N \sum_i x_i^2 - \left(\sum_i x_i\right)^2} \tag{6}$$

Two separate calibrations are performed for the Ocean Optics and KAST spectrographs. For the Ocean Optics we calibrate using Helium and Neon emission spectra, and for KAST we use the HeHgCd arclamp spectrum.

### 3.3. *Centroid Indentification Routine*

Lines $1 - 33$ of 8.1 show our routine for indetifying emission features. The input to the function is a 2D array where the first element is the pixel array, and the second element is the photon count array. Given a certain 'threshold' value the code finds the indices of all points that lie above a median + threshold*(standard deviation) cut (lines $17 - 20$). From there it groups all indices into separate arrays if they are sequential (lines $22 - 31$), so each array represents a set of indices for one emission feature. To remove small noisy features and only keep those that are signficant, we require that the length of the array my be at least 5 pixels long to be included. We also attempted other methods to indentify centroids, such as taking the slope between adjacent points and finding the points where the slope significantly changes from positive to negative, however we ran into a lot of difficulty deciphering between noise and actual emissions, separating double peaked features, and automating how to find the base of the feature, and so we abandoned this method.

Lines $39 - 50$ of 8.1 shows the implementation of how we find the centroids and compute the root mean squared (rms) width and error of each emission:

$$x_{\text{centroid}} = \langle x \rangle = \frac{\sum_i x_i I_i}{\sum_i I_i} \tag{7}$$

$$s^2 = \frac{\sum_i (x_i - \langle x \rangle)^2 I_i}{\sum_i I_i} = \langle x^2 \rangle - \langle x \rangle^2 \tag{8}$$

Propagating the error for $\langle x \rangle$ as a function of each pixel and assuming Poisson statistics for the intensity distribution we can determine $\sigma^2_{\langle x \rangle}$:

$$\sigma^2_{\langle x \rangle} = \sigma^2_{I_1} \left( \frac{\partial \langle x \rangle}{\partial I_1} \right)^2 + ... + \sigma^2_{I_N} \left( \frac{\partial \langle x \rangle}{\partial I_N} \right)^2 = \sum_i \sigma^2_{I_i} \left( \frac{\partial \langle x \rangle}{\partial I_i} \right)^2 \tag{9}$$

$$= \frac{\sum_j I_j (x_j - \langle x \rangle)^2}{\left( \sum_i I_i \right)^2} \tag{10}$$

## 4. DATA ANALYSIS & MODELING

### 4.1. *Centroid Routine Application*

Figure 4 illustrates how we apply our centroid identification routine described in Section 3.3. On the top plots of each of the three panels, the blue horizontal line shows the median photon count (ADU) and the red horizontal line shows the threshold cuts chosen for identifying emissions. Each threshold was chosen by eye such that the red line cut the spectrum close to the bas of most emission features. The bottom plot of the three panels shows the centroid pixel plotted which the code identifies with green vertical lines.

Tables 3 and 2 give the centroid pixel values we computed and their corresponding intensity levels. For the Ocean Optics calibration we compared matched pixels to wavelengths using lines from the NIST atomic line database[1]. As a visual guide we also compared our helium emissions to a graph of a spectrum with identified lines[2], since it was difficult to tell from NIST alone if the intensities matched up with our measurements. For the KAST calibration we visually compared our emissions the arclamp templates found on the website[3], particularly the template for the blue 600/4310 grism HeHgCd arclamp.

Since we noticed our centroid errors were of small magnitude $(10^{-3} - 10^{-5})$ range, we checked to see if the size of the error depended on the magnitude of the intensity, to see if sharper, higher intensity peaks had smaller error. Figure 7 shows that most of the points with higher errors have lower intensities, with a roughly $1\sqrt{n}$ relationship, which matches the pattern we found in the previous lab.

[1] https://www.nist.gov/pml/atomic-spectra-database
[2] https://www.vernier.com/innovate/a-quantitative-investigation-of-the-helium-spectrum/
[3] https://mthamilton.ucolick.org/techdocs/instruments/kast/images/Kastblue600HeHgCd.jpg

| centroid pixel | centroid intensity | centroid error | nist wave | nist intensity | gas element |
|---|---|---|---|---|---|
| 132.3649283 | 236.48 | 0.001898 | 388 | 60-300 | He |
| 288.9587617 | 249.46 | 0.002074 | 447 | 25-200 | He |
| 412.1908462 | 98.46 | 0.005039 | 492 | 20 | He |
| 437.7643945 | 271.16 | 0.001998 | 501 | 100 | He |
| 677.5521828 | 3106.25 | 0.000248 | 588 | 120-500 | He |
| 908.0143522 | 564.74 | 0.001505 | 668 | 200 | He |
| 1021.697778 | 867.6 | 0.001039 | 706 | 20-100 | He |
| 673.712 | 1696.45 | 0.001759 | 585.25 | 200 | Ne |
| 700.126 | 617.38 | 0.0045 | 594.48 | 50 | Ne |
| 722.492 | 181.31 | 0.003963 | 602.99 | 100 | Ne |
| 747.974 | 1339.68 | 0.005689 | 614.3 | 100 | Ne |
| 775.796 | 290.88 | 0.003017 | 621.73 | 100 | Ne |
| 820.02 | 3428.72 | 0.005194 | 640.22 | 200 | Ne |
| 860.822 | 1265.04 | 0.001739 | 650.65 | 150 | Ne |
| 886.09 | 419.28 | 0.002222 | 659.9 | 100 | Ne |
| 913.44 | 767.24 | 0.005151 | 667.82 | 50 | Ne |
| 982.83 | 635.18 | 0.00158 | 692.94 | 1000 | Ne |
| 1013.238 | 834.96 | 0.00133 | 705.91 | 100 | Ne |
| 1134.814 | 80.86 | 0.00923 | 748.87 | 300 | Ne |

**Table 2.** Centroid pixels for the Ocean Optics Helium and Neon calibration spectra. The corresponding wavelengths for each emission are matched by comparison to lines from the NIST atomic line database. Units for centroid intensity are ADU, and wavelengths are in nm.

| centroid pixel | centroid intensity | centroid error | ref wave | gas molecule |
|---|---|---|---|---|
| 23.50092286 | 30772.15763 | 0.000285 | 326.105 | HeHgCd |
| 184.1577234 | 37829.79081 | 3.10E-05 | 361.051 | HeHgCd |
| 229.5492333 | 40120.66215 | 0.000152 | 365.015 | HeHgCd |
| 477.9985327 | 11444.21891 | 4.00E-05 | 404.656 | HeHgCd |
| 639.6619601 | 30172.47648 | 1.60E-05 | 435.833 | HeHgCd |
| 949.4331504 | 35609.38425 | 1.60E-05 | 479.992 | HeHgCd |
| 1257.08592 | 5622.686548 | 1.00E-04 | 505.882 | HeHgCd |
| 1372.254406 | 10136.48802 | 6.60E-05 | 546.074 | HeHgCd |
| 1638.903321 | 10132.41605 | 7.80E-05 | 587.562 | HeHgCd |

**Table 3.** Centroid pixels for the KAST HeHgCd arclamp calibration spectrum. The corresponding wavelengths for each emission are matched by comparison to template spectrum found on the Lick Observatory website. Units for centroid intensity are ADU, and wavelengths are in nm.

4.2. *Wavelength Solution*

## 5. DISCUSSION

As seen in Figure , the solar spectrum and lamp spectrum show the approximate profile of a blackbody. Blackbody radiation

$$B_\lambda(T) = \frac{hc^3}{\lambda^5} \frac{1}{e^{hc/\lambda k_b T} - 1} \tag{11}$$

## 6. CONCLUSION

In conclusion we find the wavelength solution for the Ocean Optics spectrometer to be

However further analysis into the physics of the solar or lamp spectra taken on the Ocean Optics spectrograph would require correcting for detector biases.

## 7. AUTHOR CONTRIBUTIONS

This project was done in collaboration with Julian Beas-Gonzalez and Russell Van-Linge (Group E), and we divided the tasks for this lab. I worked on writing a routine to isolate arrays that contain emission features in the gas lamp spectra, match lines to the Helium and HeHgCd spectra, and applied the wavelength calibration using linear regression. Julian worked on computing the centroids and errors, and matching lines to the Neon spectrum for calibration. Russell worked on reducing the 2D KAST images (bias subtraction, normalization, and slicing) to produce 1D spectra, and tested alternative centroid identification methods. Besides manually matching arclamp centroids to NIST/template spectrum, our code is mostly automated.
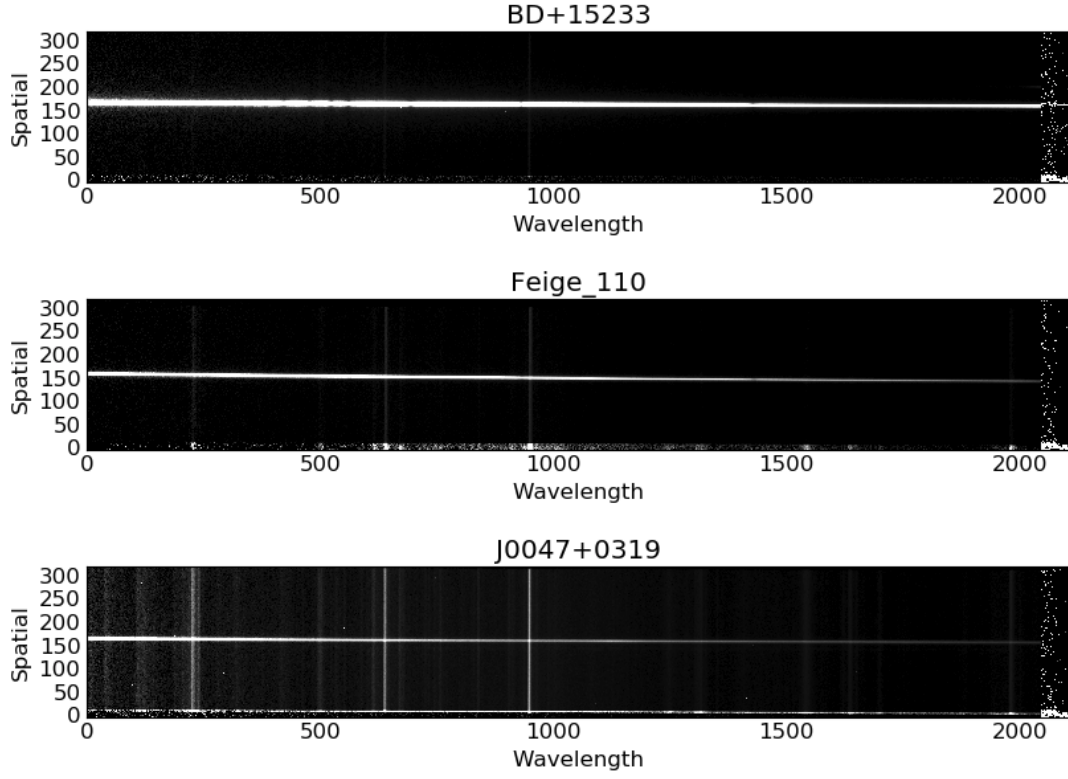
**Figure 3.** Bias-subtracted, normalized, 2D images of three different spectra taken on the KAST spectrograph: BD+15233, Feige 110, and J0047+0319. X-axis shows the wavelength dimension, y-axis shows the spatial dimension. Horizontal white line marks the source, and vertical white lines mark emission features in Earth's atmosphere. Dark bands across the source are absorption features, due to electron transitions from atoms/molecules in both the star's atmosphere and Earth's atmosphere.
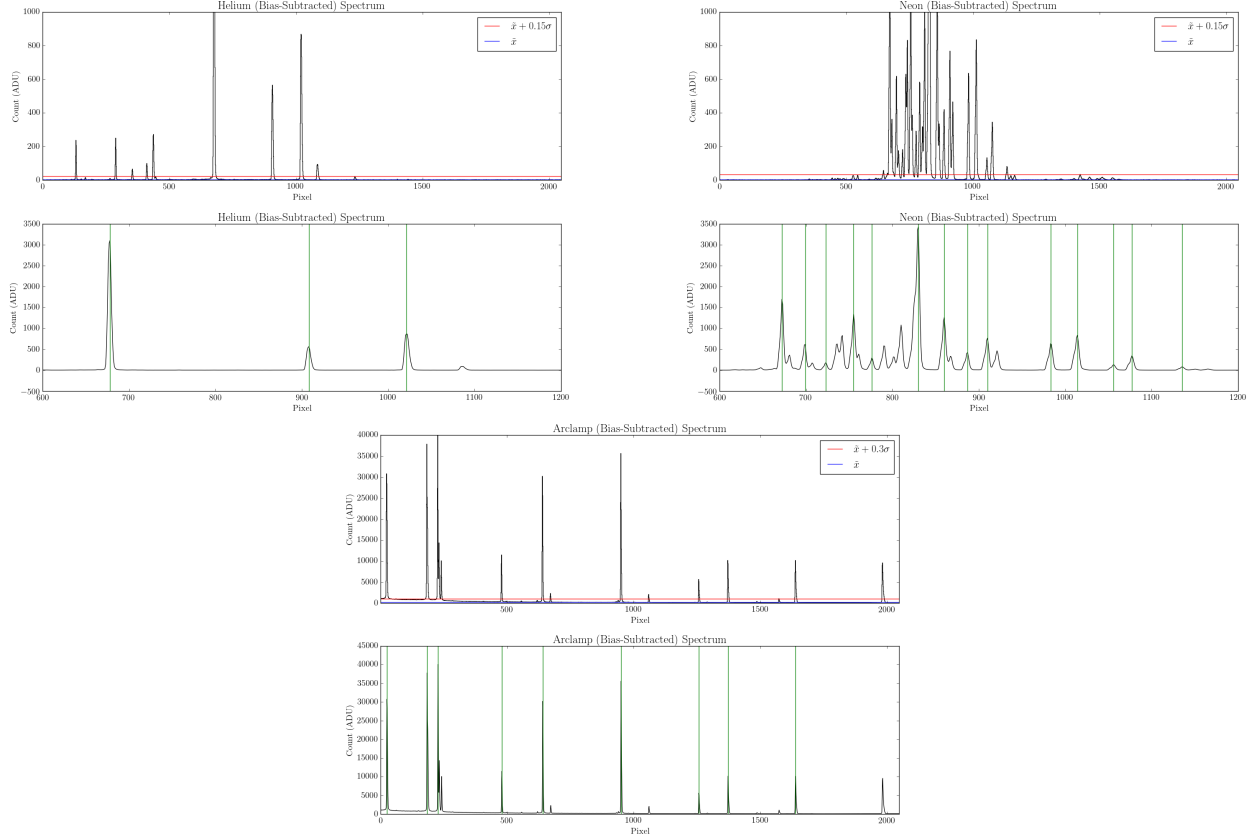
**Figure 4.** Centroid finding routine applied to the Helium, Neon, and HeHgCd arclamp emission spectra.
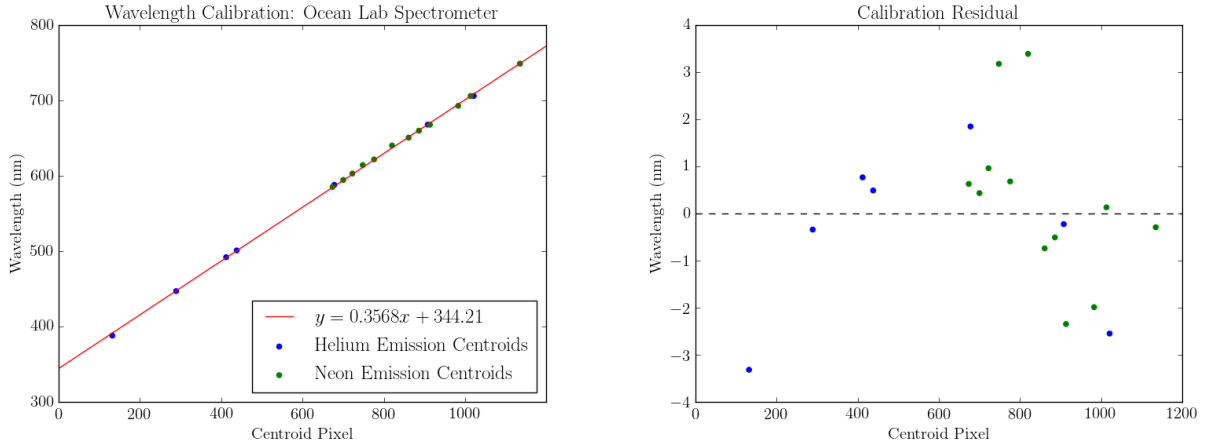


**Figure 5.** Wavelength calibration for the Ocean Optics spectrometer. Here we determine the pixel-to-wavelength conversion to be $\lambda = .3568(\text{pixel}) + 344.21$ nm (left). Lack of obvious correlations in the residual plot (right) indicates that a first order polynomial fit is sufficient for describing the wavelength-pixel relationship.
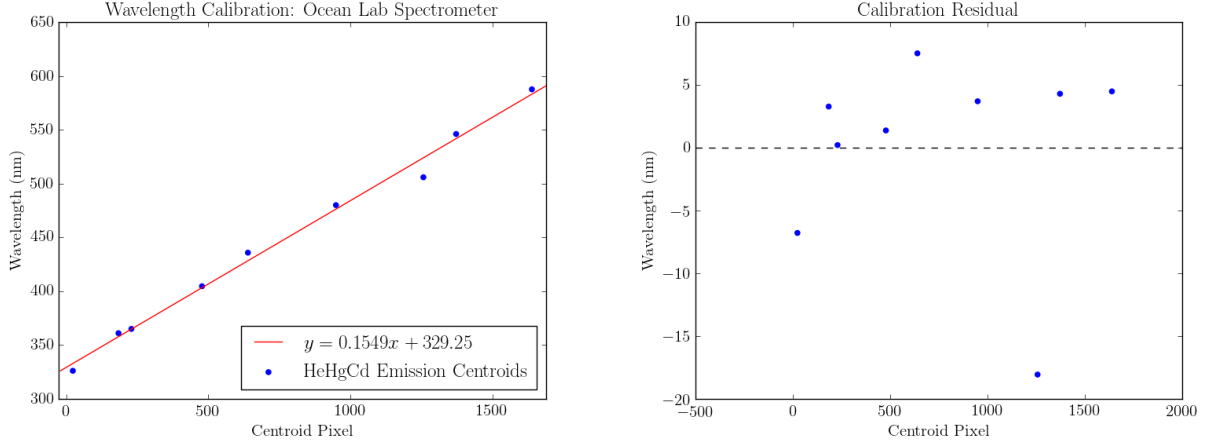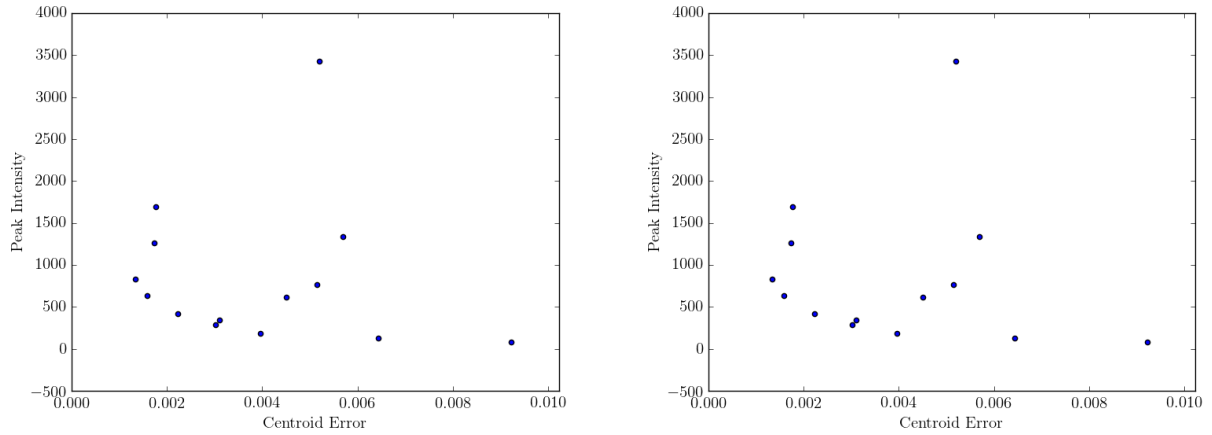
**Figure 6.** Wavelength calibration for the KAST spectrometer. Here we determine the pixel-to-wavelength conversion to be $\lambda = .1549(\text{pixel}) + 329.25$ nm (left). Lack of obvious correlations in the residual plot (right) indicates that a first order polynomial fit is sufficient for describing the wavelength-pixel relationship.



**Figure 7.** Error vs. Intensity at centroid of each identified emission peak. We see that the majority of point generally appear to lie along $1/\sqrt{n}$ pattern, with some outliers.
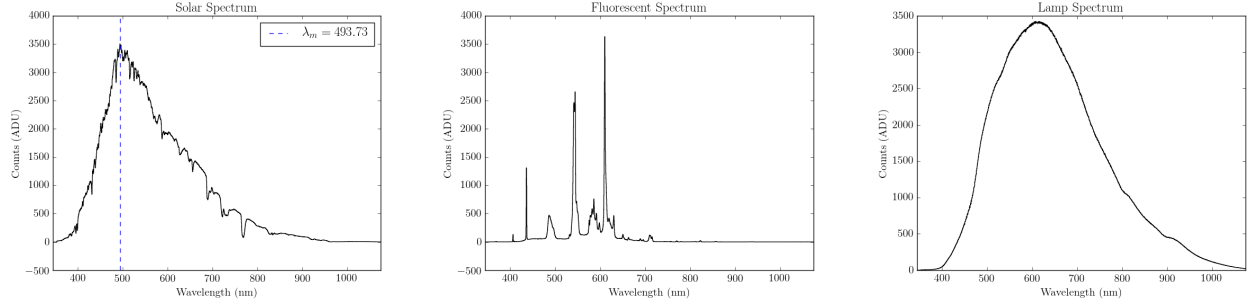
**Figure 8.** Three spectra taken on the Ocean Lab Spectrometer: the Sun, a fluorescent light bulb, and an incandescent lamp. Wavelength scale determined by calibration with emission centroids in Helium and Neon lamps to known emission peaks from the NIST atomic line database.
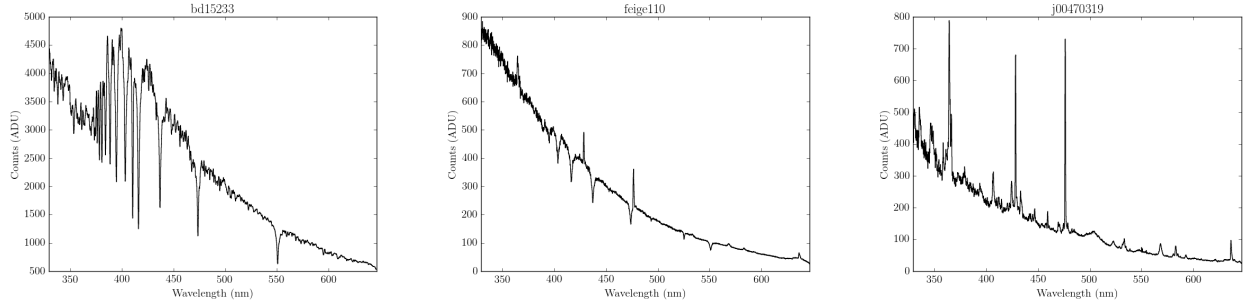


**Figure 9.** Three astronomical spectra taken on the KAST spectrograph: BD+15233, Feige 110, and J0047+0319. Wavelength scale determined by calibration with emission centroids in HeHgCd arclamp spectra.

## 8. APPENDIX

### 8.1. *Centroid Identification Routine*

```python
def emission(data, **kwargs):
    """
    Input:  'data': 2D array [x,y]
            'thres': standard deviation threshold
    Output: 'emission': an array of arrays which contain the index of
                each emission which lies above some threshold cut
    """
    thres = kwargs.get('thres', 1)

    x, y = np.array(data[0]), np.array(data[1])
    Npix = len(x)

    med = np.median(y)
    std = np.std(y)
    cut = med + thres*std

    count_cut = []
    for i in range(Npix):
        if y[i] >= cut:
            count_cut.append(i)

    emission = []
    arr = []
    for i in range(len(count_cut)-1):
        if (count_cut[i+1] - count_cut[i]) == 1:
            arr.append(count_cut[i])
        else:
            arr.append(count_cut[i])
            if len(arr) > 5:    # only keep arrays that have >5 pixels
                emission.append(np.array(arr))
            arr = []

return np.array(emission)

"""
Produces float values for the centroids, calculating their error and width
Credit: Julian
"""
centroids, errors, widths, intensities = [], [], [], []
for i in np.arange(len(cent)):
    inte = []
    for a in emiss[i]:
        inte.append(data[1][a])
    centr_f = sum(emiss[i]*inte)/sum(inte)
    err_f = sum(inte*((emiss[i]-centr_f)**2))/(sum(inte))**2
    width_f = sum(inte*((emiss[i]-centr_f)**2))/sum(inte)

    centroids.append(centr_f)
    errors.append(err_f)
    widths.append(width_f)
```

```
51      intensities.append(max(data[1][emiss[i]]))
```

## 8.2.  *Wavelength Calibration*

```
1  def linear_regression(x, y):
2      """
3      Input:  x, y: 1D arrays
4      Output: [m, c], [m_err, c_err]: slope and intercept best fit and error
5      """
6      N = len(x)
7      x, y = np.array(x), np.array(y)
8
9      A = np.array([[np.sum(x**2), np.sum(x)], \
10                    [np.sum(x), N]])
11     a = np.array([np.sum(x*y), np.sum(y)])
12
13     fit = np.dot(np.linalg.inv(A), a)
14
15     sig_sq = np.sum(y - (fit[0]*x + fit[1]))**2/(N + 2)
16     m_err = np.sqrt(N*sig_sq/(N*np.sum(x**2) - (np.sum(x))**2))
17     c_err = np.sqrt(sig_sq*np.sum(x**2)/(N*np.sum(x**2) - (np.sum(x))**2))
18     err = np.array([m_err, c_err])
19
20     return fit, err
```

## 8.3.  *KAST Data Reduction*

```
1  """
2  Credit: Russell
3  """
4  def readData(folder):
5      '''
6      Reads all the fits files in a folder and creates a 3D array
7      '''
8      files = os.listdir(folder)
9      array3D = []
10     for ff in files:
11         arr = fits.getdata(folder+ff,ignore_missing_end=True)
12         array3D.append(arr)
13     return np.array(array3D)
14
15 def combineFrame(data_array):
16     '''
17     Avergaes the 3D into 2D array
18     '''
19     return np.median(data_array,axis=0)
20
21 def norm_flat(bias_folder,flat_folder):
22     '''
23     Input the bias and flat folders and creates
24     bias,flats, and normalized flats
25     '''
26     flat3d = readData(flat_folder)
27     bias3d = readData(bias_folder)
28
```

```
29      flat2d = combineFrame(flat3d)
30      bias2d = combineFrame(bias3d)
31
32      norm_flat = (flat2d-bias2d)/np.median(flat2d-bias2d)
33      return norm_flat,flat2d,bias2d
34
35  def science_image(science_folder,flat_folder,bias_folder):
36      '''
37      Creates a science image and outputs the 2D array to be used to make spectra
38      '''
39      # create flats and bias
40      normflat, flat, bias = norm_flat(bias_folder,flat_folder)
41
42      science_3d = readData(science_folder) # reducing down to 2d
43      science_2d = combineFrame(science_3d)
44      science_final = (science_2d - bias)/normflat # subtracting bias and normalizing
45      # Plots the science image
46      plt.imshow(science_final,origin='lower',interpolation='nearest',\
47          cmap='gray',vmin=10,vmax=1000)
48      plt.show()
49      return science_final
50
51  def science_spectra(science_2d,start_slice,end_slice,file_name='file_name'):
52      '''
53      Creates a spectra by slicing up the science image at the input spatial
54      pixels, saves array, and outputs plot
55      '''
56      spectra = np.mean(science_2d[start_slice:end_slice,:],axis=0)
57      spectra = spectra[spectra>0]
58      x = np.arange(0,len(spectra))
59      plt.plot(x,spectra,color='black')
60      plt.show()
61      np.save('%s_spectra'%file_name,science_2d)
62      return spectra
```