

AST325 Lab 3: Astronomical Spectroscopy

Shimona Das

Student Number: 1008964943

Group N

December 2024

Abstract

This lab investigated astronomical spectroscopy using a USB650 spectrometer and the LISA telescope spectrograph to analyse emission spectra from the laboratory and celestial bodies. The hydrogen spectrum was calibrated by mapping pixel positions to wavelengths using a linear fit with a slope of 1.020.02 nm/pixel and an intercept of 13.9312.33 nm. This proved to be the best fit when compared to quadratic and cubic fits. The mystery element's spectrum was identified to be neon, based on its characteristic emission lines. Using the LISA spectrograph, the spectra of celestial targets such as Vega, Saturn, Sadalmelik, HD182640, and Almach, were collected and also calibrated. Several challenges were faced such as noise, probe stability, and atmospheric interference, which were mitigated through dark subtraction, optimized integration times, and polynomial calibration techniques. These results aligned with theoretical expectations and confirmed the thoroughness of the calibration process and also the spectrometers' power. This report highlights the importance of accurate calibration and data reduction in spectroscopy and serves as a foundation for future analysis of celestial objects.

1 Introduction

Spectroscopy is an essential tool across all physical sciences and provides insights into the properties and behaviours of substances through their interaction with light. For astrophysicists, spectroscopy allows for the detailed characterization of celestial objects and the universe. Astronomical spectroscopy, in particular, allows for the measurement of chemical compositions and physical conditions, such as temperature, pressure, and magnetic field strength, in celestial bodies including planets, stars, and galaxies. By analysing these spectral lines, astronomers deduce information about distant objects, which makes spectroscopy invaluable for understanding astrophysical phenomena.

This lab focuses on introducing the basics of astronomical spectroscopy through the operation of spectrometers and the interpretation of spectral data. Using a simple visible light spectrometer (350-700 nm), the instrumental calibration was investigated by establishing a wavelength scale and examining the noise characteristics of a Complementary Metal-Oxide-Semiconductor (CMOS) detector. The lab is divided into two parts: calibration and analysis of a classroom-based spectrometer and then the use of the LISA spectrometer on a telescope to collect astronomical spectra from celestial targets. Key techniques included wavelength calibration, noise reduction, and the fundamentals of data reduction, such as dark subtraction and flat fielding which are quite important for accurate spectral analysis.

The first part of the lab used a USB650 spectrometer, the instrument was calibrated using known light sources and wavelength calibration was performed through linear least squares fitting. This process allowed familiarisation with data acquisition, calibration techniques, and the interpretation of spectra. The second part involves observations at night time using a telescope-mounted spectrograph to find spectra from astronomical sources. This calibration was then applied to analyse the spectral features of stars and planets by processing the data to produce clear, interpretable spectra.

This report is structured as follows: an overview of the experimental setup, including spectrometer calibration and data collection methods. It is then followed by data analysis and interpretation of the calibrated spectra from both laboratory and astronomical sources. Finally, a discussion addresses the findings and implications of the spectral features observed, concluding with remarks on the practical applications of astronomical spectroscopy.

2 Data and Observation

Date	Personnel	Notes
07/11/2024	S. Das, M. Pye, J. Hora	Recorded first set of data set on USB650
14/11/2024	S. Das, M. Pye, J. Hora	Retook data with lower integration time
26/11/2024	S. Das, M. Pye, J. Hora	Recorded data for astronomical objects on telescope

Table 1: A summary of observations

Specification	USB650 Red Tide
Wavelength Range	350 - 1000 nm
Detector	Sony ILX511 CCD
Number of Pixels	2048
Sensitivity	75 photons/count at 400 nm
Pixel Size	14 μm x 200 μm
A/D Resolution	12-bit
Optical Resolution	2.0 nm (FWHM)
Dynamic Range	2×10^8 (system)
Integration Time	10 μs to 60 s
Dimensions	89.1 mm x 63.3 mm x 34.4 mm

Table 2: Properties of USB650 Red Tide

Specification	LISA Spectrograph
Wavelength Range	400 - 700 nm (Visible mode) / 650 - 1000 nm (IR mode)
Resolving Power	\sim 600 - 1000 (depends on slit and mode)
Slit Widths	15 μm , 19 μm , 23 μm , 35 μm
Slit Length	4 mm
Collimator Focal Length	130 mm
Camera Lens Focal Length	88 mm
Slit Image Reduction Factor	0.68
Standard Backfocus	41 mm (telescope) / 54.85 mm (CCD) / 17.5 mm (guiding camera)
Calibration Options	Neon lamp, Halogen lamp, White screen with electro-magnet

Table 3: Specifications of the LISA High Luminosity Spectrograph.



Figure 1: Setup of the Red Tide USB650 spectrometer measuring the emission spectrum of a gas discharge tube. The fiber optic probe is carefully positioned near the light source to capture the emitted light without direct contact. The Red Tide USB650 spectrometer is connected via USB to the computer for real-time data acquisition and spectral analysis.



Figure 2: The figure shows the telescope setup used. The telescope is housed within a dome, designed to shield from ambient light and weather conditions. The primary optical tube assembly (OTA) is visible with its supporting equatorial mount, which allows the tracking of celestial objects. Various cables and instruments are attached to the telescope, including a spectrograph or CCD camera for data acquisition. The movable open dome reveals the night sky, enabling unobstructed observation of celestial bodies. This setup was used to capture raw spectral and imaging data for calibration and spectral analysis.

Target Name	Date (YYYY-MM-DD)	Time (HH:MM)	Observers
Vega	2024-11-26	18:30	S. Das, M. Pye, J. Hora
Scheat	2024-11-26	18:40	S. Das, M. Pye, J. Hora
Jupiter	2024-11-26	18:50	S. Das, M. Pye, J. Hora
Saturn	2024-11-26	19:00	S. Das, M. Pye, J. Hora
Algenib	2024-11-26	19:05	S. Das, M. Pye, J. Hora
Almach	2024-11-26	19:10	S. Das, M. Pye, J. Hora
Ring Nebula	2024-11-26	19:15	S. Das, M. Pye, J. Hora

Table 4: Observation log documenting targets, dates, times, and observers present during the session.

During the USB650 experiment, initial challenges included stabilising the probe, which had to be resolved by using a cardboard box as support. We experimented with integration times ranging from 3 ms to 200 ms and found that 3 ms gave a better peak for detection of the mystery element, as higher values caused peaks to flatten to plateaus. Dark samples were collected and subtracted to minimize noise. Additionally, data originally saved in incorrect format (.ocv) was retaken in .txt format.

For the telescope experiment, most of the data were collected with a 10-second integration time along with dark frames of the same duration. For fainter celestial objects like the Ring Nebula, a 600-second integration time was used because of cloud cover and low signal detection.

3 Data Reduction

The data reduction process was important for the preparation of the raw data for analysis, which ensured reliability and accuracy of our results. In the following, steps were taken to clean and process the data, along with an explanation of the motivation and methods used at each stage.

The first part was the **hydrogen and mystery element spectra**. It involved plotting the intensity versus wavelength for the hydrogen spectrum to identify key peaks. A prominence threshold of 20 was applied to detect only significant peaks and filter out the low intensity signals. For the "mystery element", a slightly higher prominence value of 25 was used to filter out even more noise. This choice was motivated by the need to isolate meaningful spectral features. Once the peaks were identified, the pixel position vs. wavelength was plotted and polynomial fits (linear, quadratic, and cubic) were applied to determine the best calibration model. The quadratic fit showed the lowest chi-squared value with a random residual pattern, making it the most appropriate choice for wavelength calibration. This quadratic model was similarly applied for the telescope's spectral calibration which ensured consistency across the experiment.

To refine the row number against **centroid analysis**, the range of rows were limited from 250 to 480, isolating the region where the centroids exhibited a linear trend. This range was visually determined by

inspecting the raw data, which revealed significant outlier rows outside this region. By focusing on this interval, the centroids were accurately determined which minimised noise and reduced uncertainty.

During the **spectrometer data handling**, stabilising the probe was quite challenging initially. This was resolved by using a foam box to hold the probe steady, which ensured consistent measurements. Integration times between 3 ms and 200 ms were tested and it was observed that a 3 ms integration time was best for detecting sharp peaks, because higher values flattened the peaks into plateaus. To account for noise, dark samples were collected and subtracted from the light samples. Additionally, initial data saved in the incorrect `.ocv` format had to be retaken in the correct `.txt` format.

For the **telescope spectral calibration**, integration times were chosen to match the brightness of the targets and the atmospheric interference. Most celestial objects were captured with an integration time of 10 seconds then followed by dark frames of the same duration for noise subtraction. For fainter targets, such as the Ring Nebula, a longer integration time of 600 seconds was necessary due to cloud cover and low signal levels. These longer exposures allowed us to detect features that would otherwise be lost in noise.

Combining the spectral and spatial calibrations in the **Final Data Preparation**, a 2D wavelength calibration array was created for each astronomical target. The wavelength calibration used a quadratic polynomial derived from earlier steps. The final 1D spectra for each target were extracted by summing across rows in the calibrated images and mapping pixel positions to calibrated wavelengths. This process ensured that the data was properly prepared for analysis, with key spectral features resolved and systematic uncertainties minimised.

To improve the accuracy of the data reduction, **uncertainties** in the hydrogen arc lamp spectrum and linear fit were incorporated into the calibration process. Using the variance of the slope (Var_m), intercept (Var_b), and their covariance (Cov_{mb}), wavelength uncertainties (σ_λ) were propagated and visualised as error bars and confidence bands. This accounted for noise in the spectrometer and the precision of the fit. In the residual analysis, centroid measurement uncertainties (σ_x) were included to evaluate the fit. By incorporating uncertainties throughout the calibration process minimised biases and accounted for instrument and measurement limitations, resulting in a thoroughly evaluated wavelength mapping.

4 Data Analysis

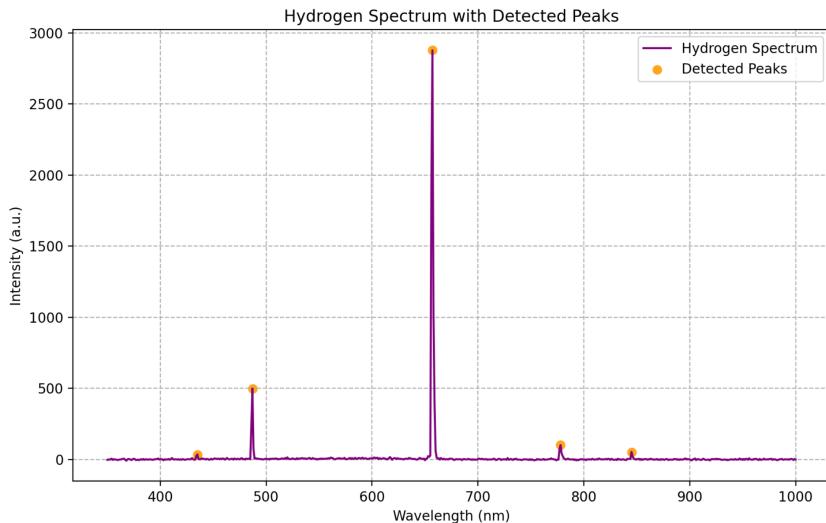


Figure 3: This plot shows the hydrogen spectrum with detected peaks, where intensity (arbitrary units) is plotted against wavelength (in nanometres). The purple curve depicts the raw hydrogen spectrum, displaying several prominent emission lines which are characteristic of hydrogen's atomic transitions. Significant peaks in the spectrum have been identified using a prominence threshold of 20 to isolate important signals while excluding background noise. These detected peaks are marked with orange points, highlighting key wavelengths where the hydrogen transitions occur. The tallest peak at approximately 656 nm corresponds to the H-alpha emission line, one of the most prominent features in hydrogen spectra. Smaller peaks near 486 nm (H-beta) and other wavelengths are also evident, showcasing additional transitions.

The hydrogen spectrum plot in Figure 3 captures the key features of hydrogen's atomic transitions and visualised the significant emission lines. The most prominent peak, corresponding to the H-alpha transition at approximately 656 nm, is detected and emphasised, which highlights its importance in hydrogen spectral analysis. Smaller peaks, such as H-beta near 486 nm, are also visible and demonstrates how the spectrum identifies secondary transitions. However, the noise level in the spectrum, particularly at wavelengths away from the emission lines, could potentially mask weaker transitions or introduce inaccuracies in peak identification. A prominence threshold of 20 was used to appropriately isolate significant peaks, as it may have excluded weaker features. Future analyses could benefit from refining the threshold or applying more advanced noise reduction strategies to improve the clarity of the spectrum. Overall, the spectrum successfully shows the primary hydrogen transitions but improvements could be made in noise reduction.

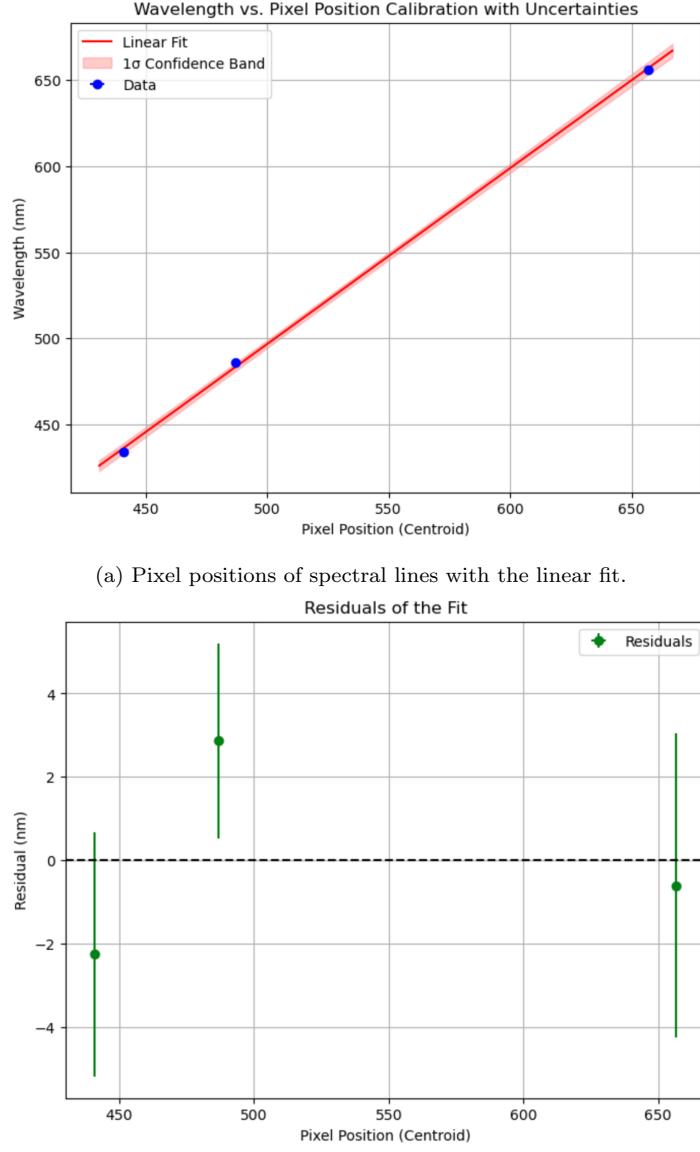


Figure 4: (a) The left plot shows the pixel positions of detected spectral lines plotted against the corresponding wavelengths for the hydrogen spectrum. Blue points represent detected spectral lines and the red line indicates the applied linear fit. The shaded red region indicates the ± 1 confidence band, reflecting the uncertainty in both the slope of 1.02 ± 0.02 and intercept of the fit of -13.93 ± 12.33 . The linear fit is used to establish a preliminary wavelength calibration. (b) The right plot shows the residuals of the wavelength calibration fit as a function of pixel position (centroid). Each green data point represents the difference between the measured wavelength and the fitted wavelength model. The error bars indicate the residual's uncertainty.

When examining the two presented plots in Figure 4a and Figure 4b, it becomes clear that the linear model provides a reasonable representation of the wavelength and pixel relationship. The first plot shows a linear fit that visually tracking the data points, and the relatively narrow confidence band suggests that the fit parameters (slope, 1.02 ± 0.02 , and intercept, -13.93 ± 12.33) are well constrained. However, the residuals plot shows that while some data points lie near zero residual, others deviate by a few nanometers, indicating that the model does not perfectly capture all variations in the observed data. The size of these deviations with their associated error bars, hint at the presence of systematic effects. In particular, the uneven spread and magnitude of residuals across different pixel positions raise queries about the completeness of the calibration model. Ultimately, while the linear fit provides a good first-order approximation, the residuals highlight that further refinement may be needed for more precision.

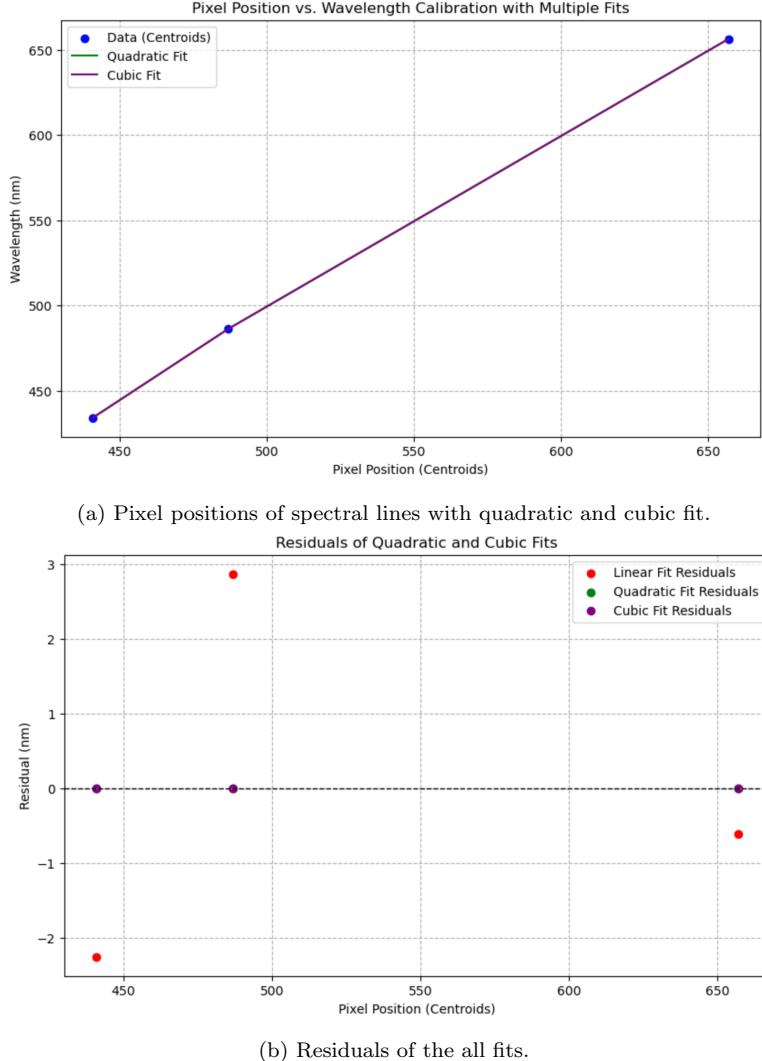


Figure 5: The left plot shows the pixel position against wavelength calibration for a quadratic and cubic fit. The blue markers represent the measured data points. Overlaid on these data are a quadratic fit (green line) and a cubic fit (purple line). Both polynomial models are fitted to the same set of calibration points, demonstrating that increasing the polynomial degree can produce a curve that still closely matches a nearly linear trend. Although higher-order fits add flexibility, the limited calibration points shown here lead to minimal differences between quadratic and cubic solutions, suggesting that the linear model may be sufficient. The right plot shows a comparison of residuals from multiple polynomial fits to the wavelength calibration data. The dashed horizontal line at zero residual marks the ideal fitted model scenario.

The residual comparison in Figure 5b highlights the limitations and advantages of different polynomial calibration models for the given data. While the linear fit's residuals are visibly larger and more

scattered, indicating that a simple one-parameter slope and intercept do not fully capture the wavelength-pixel relationship, the higher-order show a straight line pattern around $y=0$. Both the quadratic and cubic residuals remain close to zero, implying that these models could theoretically accommodate for non-linearity in the calibration data. However, the indistinguishable performance of the quadratic and cubic fits suggests that increasing the model complexity beyond a quadratic does not improve much. This outcome brings practical considerations: a more complex model cannot be justified if it does not significantly reduce improve accuracy, especially given that additional parameters can introduce unnecessary complexity and also potential overfitting. In essence, the figure demonstrates that while linearity is sufficient for precise calibration, a quadratic model introduces fitting. However, this may not be appropriate due to the straight line pattern formed.

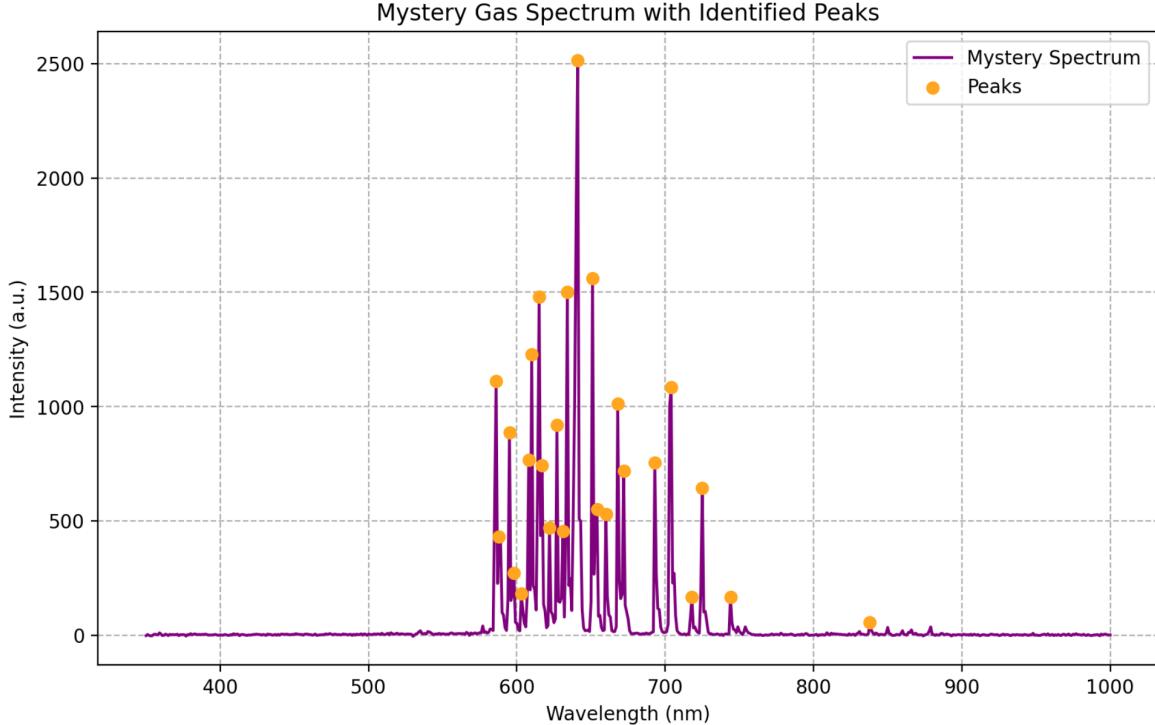


Figure 6: This plot shows the spectrum of a "mystery gas" with intensity (arbitrary units) plotted against wavelength (in nanometers). The purple curve represents the raw spectrum of the gas with several prominent emission peaks characteristic of the gas's atomic or molecular composition. Significant peaks are identified and marked with orange points. These peaks correspond to specific wavelengths where the gas exhibits strong emission, which are used for identification and analysis. The tallest peak is at around 656 nm. Other notable peaks are distributed primarily between 600 and 700 nm. The identification and analysis of these peaks are crucial for determining the chemical composition of the gas.

The graph in Figure 8 represents the spectrum of a mystery gas, and its prominent emission lines provide information about the gas' chemical identity. The detected peak wavelengths are distributed between values of 586 nm and 744 nm, with a notable cluster of emission features in the 600-700 nm range. To identify the gas, the detected peak wavelengths were compared to known emission wavelengths of common elements in the NIST atomic spectra database. Upon comparison, the array of wavelengths matches quite closely with the known spectral lines of neon, which is a noble gas commonly used in calibration lamps. For example, the detected wavelengths at approximately 586 nm, 588 nm, 595 nm, and 598 nm correspond to strong neon emission lines, and the broader distribution of peaks in the red and orange regions (such as from 640-660 nm) further aligns with the characteristic neon spectrum.

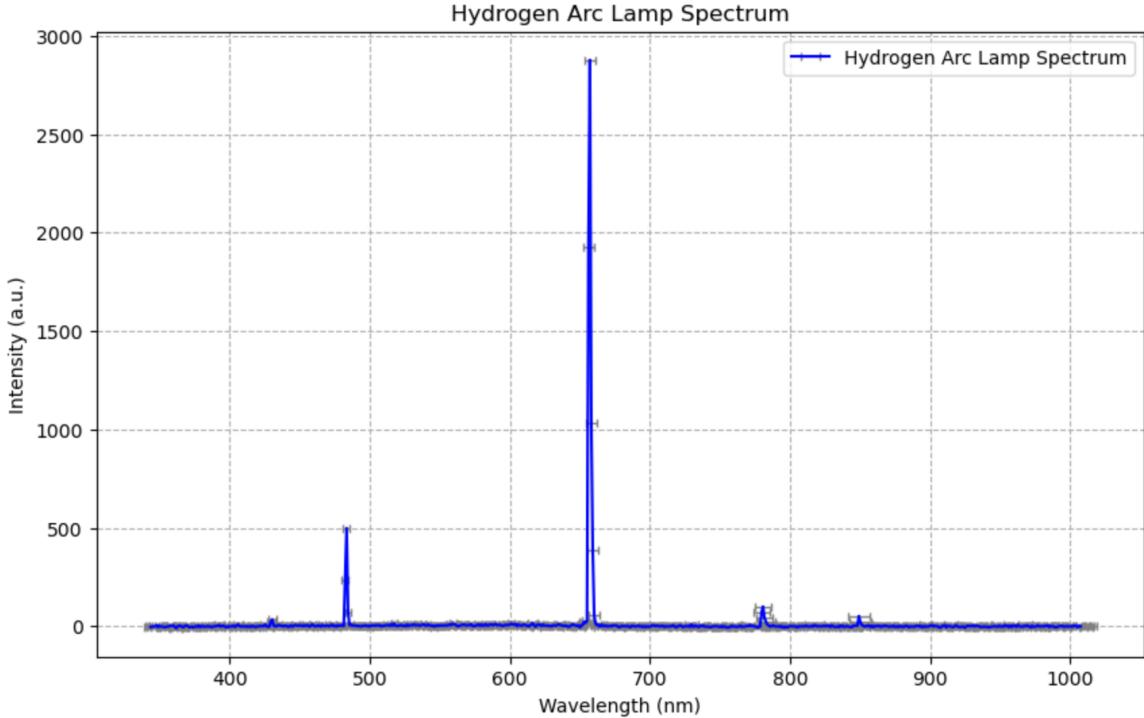


Figure 7: The plot displays the emission spectrum of a hydrogen arc lamp, with wavelength (in nanometers) along the x-axis and intensity (in arbitrary units) for the y-axis, and wavelength calibration uncertainty. The peaks represent prominent spectral lines characteristic of hydrogen's atomic transitions. The tallest peak, at 784 nm, corresponds to the most intense emission line, while additional smaller peaks are at other wavelengths. The spectrum has been dark-subtracted to remove background noise. The data provides a basis for wavelength calibration, with the known emission lines mapped to pixel positions in the spectrometer.

This spectrum in Figure 7 provides a clear visualization of the emission lines from a hydrogen arc lamp, extending over a wide wavelength range. The primary features, such as the prominent peak near 656 nm (H-alpha), are consistent with expectations for hydrogen's Balmer series. However, the inclusion of data beyond 800 nm introduces regions with minimal intensity and noise, which may not potentially contribute significantly to the analysis and could be cropped for better clarity. The alignment and width of the peaks suggest a well-calibrated spectrometer. Future iterations could focus on refining the x-axis range and annotating key spectral features. The uncertainty in the hydrogen spectrum's wavelength calibration was computed using the propagated uncertainties of the linear regression parameters by mapping pixel positions to wavelengths. The wavelength variance (σ_λ^2) was derived from the covariance matrix of the fit, which accounts for the slope variance (Var_m), intercept variance (Var_b), and their covariance (Cov_{mb}). This method incorporates the linear fit's uncertainties and the pixel positions (x) input on the mapped wavelengths. The standard deviation (σ_λ) was displayed as error bars, effectively demonstrating the uncertainty at each point. However, it should be noted that this approach assumes a linear model and negligible pixel noise, which may omit potential systematic errors.

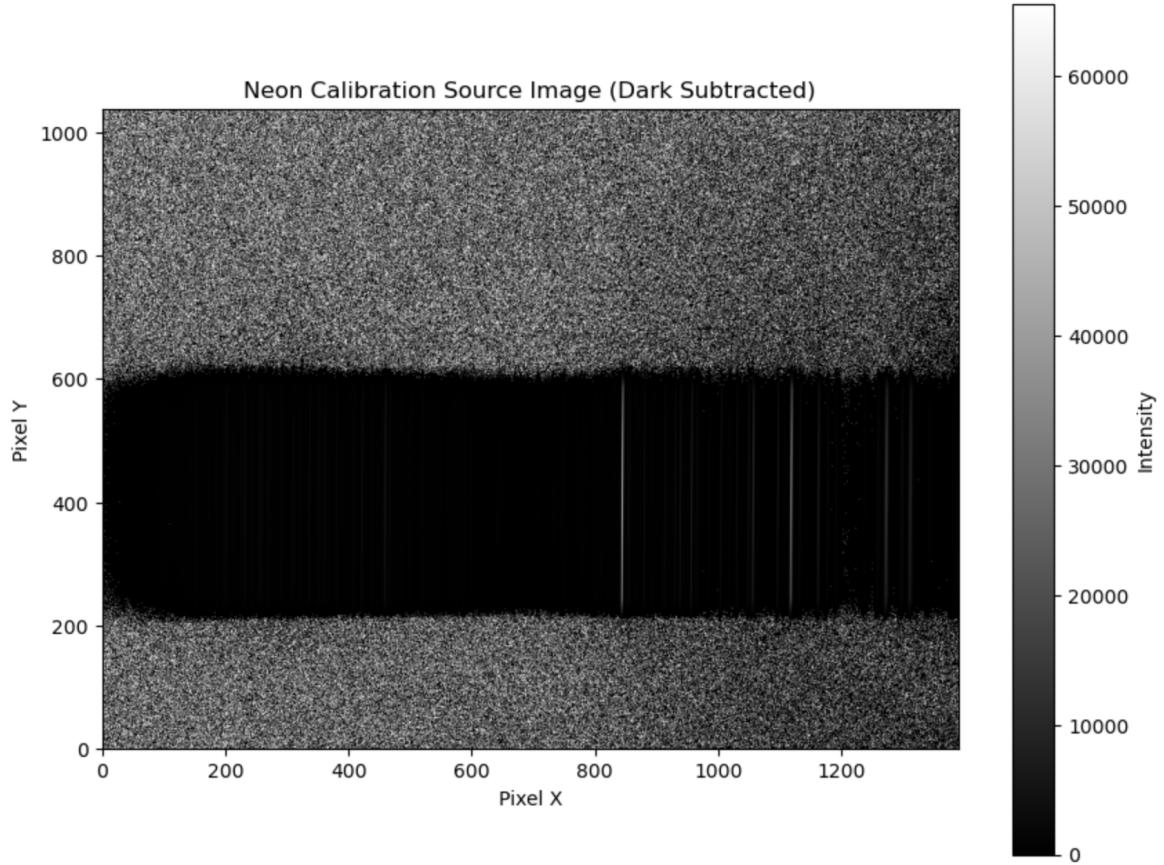
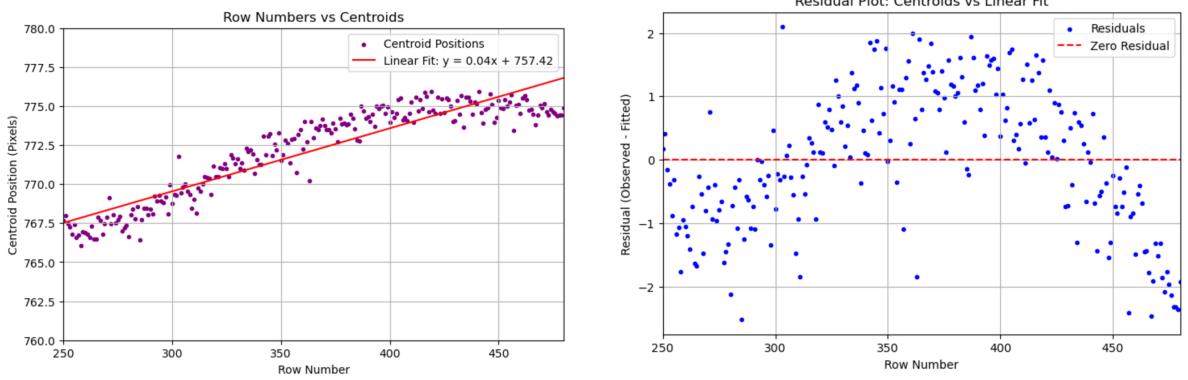


Figure 8: The image displays the 2D intensity map of the neon calibration source after the dark frame have been subtracted. The pixel intensity values, represented on the grayscale color bar to the right, correspond to the light captured by the detector during exposure. The dark subtraction removes unwanted thermal noise and electronic readout patterns, leaving only the relevant signals from the neon source. The horizontal band in the middle of the frame corresponds to the illuminated region which contains the emission lines of the neon gas. These lines are spread across the pixels due to the dispersion introduced by the spectrometer, which separates the different wavelengths. This image serves as a starting point for wavelength calibration, where the pixel positions of bright emission features are used to associate specific pixel locations with the corresponding wavelengths. The dark regions above and below the illuminated band indicate areas of the detector not exposed to significant light.

Figure 8 represents the dark-subtracted neon calibration source which is a critical step in preparing data for accurate wavelength calibration. While the horizontal band highlights the emission region effectively, the surrounding noise indicates the detector's sensitivity to background signals even after dark subtraction. The sharp emission lines visible within the band suggest that the neon source provided distinct and resolvable features which are ideal for calibration. However, any inconsistencies in the intensity distribution in the band could potentially affect the precision of the calibration. Ensuring consistent illumination for this and reducing residual noise further could improve the quality of the calibration process.

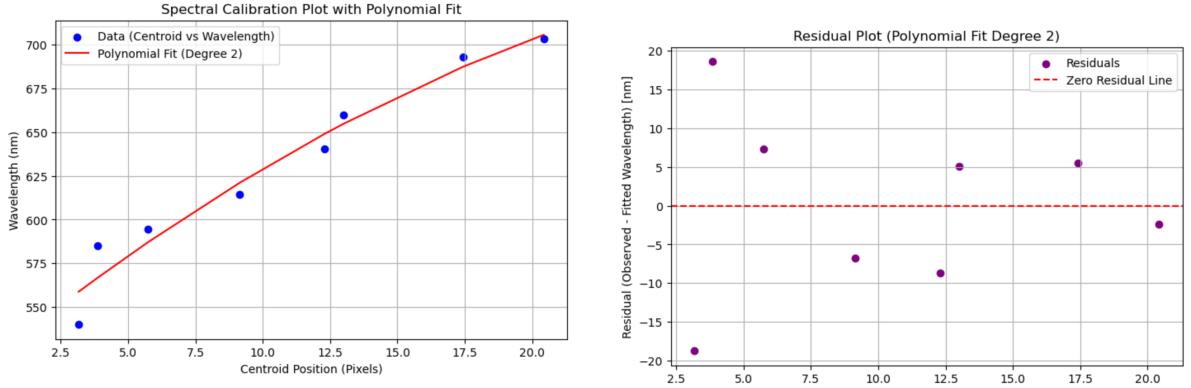


(a) Row numbers versus centroid positions which shows the linear relationship. Purple points represent the observed centroid positions while the red line indicates the linear fit with the equation $y = 0.04x + 757.42$. The linear fit achieves an R^2 value of 0.8560, suggesting a strong correlation.

(b) Residual plot showcasing the linear fit. The blue points represent the residuals for each data point, while the red dashed line indicates zero residual. The mean residual is approximately zero, and the standard deviation is 1.1046, indicating moderate scatter around the fit.

Figure 9: Combined plots show the linear fit of row numbers versus centroid positions and the corresponding residuals. The residuals highlight deviations from the linear model, while the centroid plot emphasizes the strong correlation between rows and centroid positions..

The combined plots in Figure 10 demonstrate the relationship between row numbers and centroid positions, along with the residuals from the linear fit. The centroids plot demonstrates a linear trend, supported by the fit's high R^2 value of 0.8560, indicating that the model captures most of the data's variability. However, the residual plot reveals scattered deviations around zero, with a standard deviation of 1.1046. While the residuals are generally small, their distribution suggests minor inconsistencies in the linear model's ability to model the data. Overall, the analysis confirms the linear trend is a reasonable approximation.



(a) Spectral calibration plot showing centroid positions (in pixels) against corresponding known wavelengths (in nanometers) for Neon emission lines. The blue points represent the observed data, while the red curve shows the quadratic polynomial fit applied to the data. The polynomial coefficients for the fit are $-1.7e-01$ $1.3e+01$ $5.2e+02$, indicating the calibrated relationship between pixel positions and wavelengths.

(b) Residual plot for the spectral calibration, which shows the differences between the observed wavelengths and predicted polynomial fit. The purple points represent the residuals at each centroid position, while the red dashed line indicates the zero-residual level. The distribution of residuals suggests the polynomial fit effectively captures the relationship, with minimal deviations.

Figure 10: Combined plots for the spectral calibration analysis. (a) Shows the relationship between centroid positions and known wavelengths, with a quadratic polynomial fit applied. (b) Displays the residuals of the polynomial fit with the accuracy and consistency of calibration.

The spectral calibration plot and its residual plot in Figure 10 provide insights into the accuracy of the polynomial fit applied to the centroid positions and associated wavelengths. In the spectral calibration plot in Figure 10a, the quadratic polynomial fit captures the general trend of the data and aligns well with the observed centroids. However, the residual plot in Figure 10b shows areas where the fit deviates

from the data, particularly at the edges of the centroid range, with residuals peaking near ± 20 nm. While residuals are relatively small and randomly distributed which suggests the fit is robust for most of the range, the higher deviations at the extremes can indicate non-linearities and systematic errors in the calibration data.

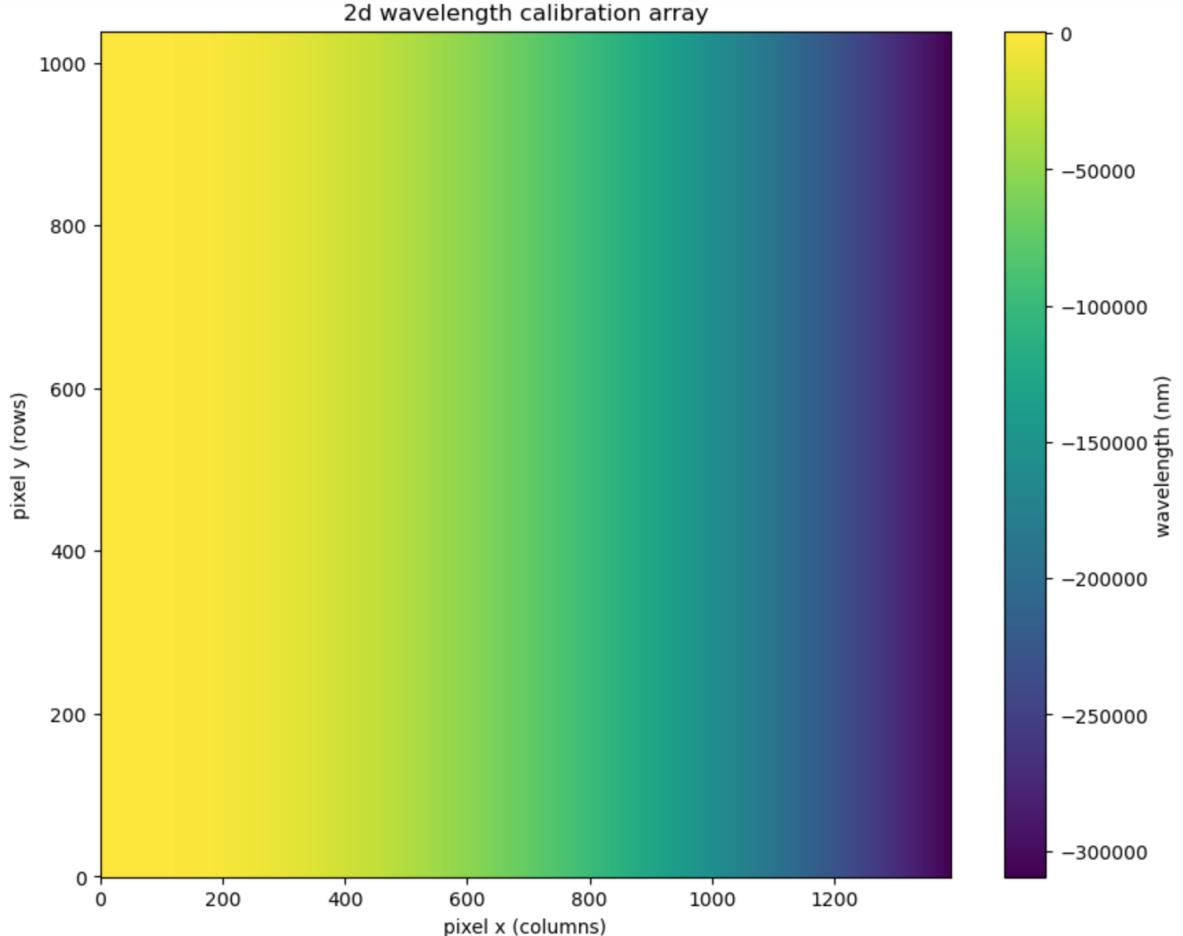


Figure 11: This figure displays a two-dimensional wavelength calibration array generated from a polynomial fit, applied to the centroid positions of spectral emission lines. X-axis represents the pixel columns, corresponding to the dispersion axis of the spectrograph, Y-axis represents the pixel rows, corresponding to the spatial axis. The colour scale indicates the wavelength values (in nanometers), which transitions from yellow for shorter wavelengths to purple for longer. This calibration array is a mapping tool, enabling each pixel's intensity value in the spectrograph image to be assigned an accurate wavelength. The gradient reflects the continuous nature of the polynomial calibration across the array. This map is essential for extracting spectra from astronomical targets.

The 2D wavelength calibration array in Figure 11 provides a mapping between the pixel positions of the detector and corresponding wavelengths, which is derived from a polynomial fit to spectral emission line centroids. The gradient across the array seems quite smooth and continuous, indicating a good polynomial model but the accuracy of the calibration heavily depends on the degree of the polynomial fit and the quality of the initial centroid and wavelength data. Potential sources of uncertainty include errors in centroid calculations or inaccuracies in the known reference wavelengths.

This array can be used to plot spectra from real astronomical targets through mapping the pixel intensities in a spectrograph image to their corresponding wavelengths. For a given region of interest such as a bright spectral row or aperture, the calibrated array provides the wavelength values for each pixel along the dispersion axis. By summing or averaging the intensities along the spatial axis, a 1D spectrum can be extracted. This can then be plotted as intensity versus wavelength which enables the identification of spectral features. This process is critical for interpreting the physical properties of astronomical objects based on spectra.

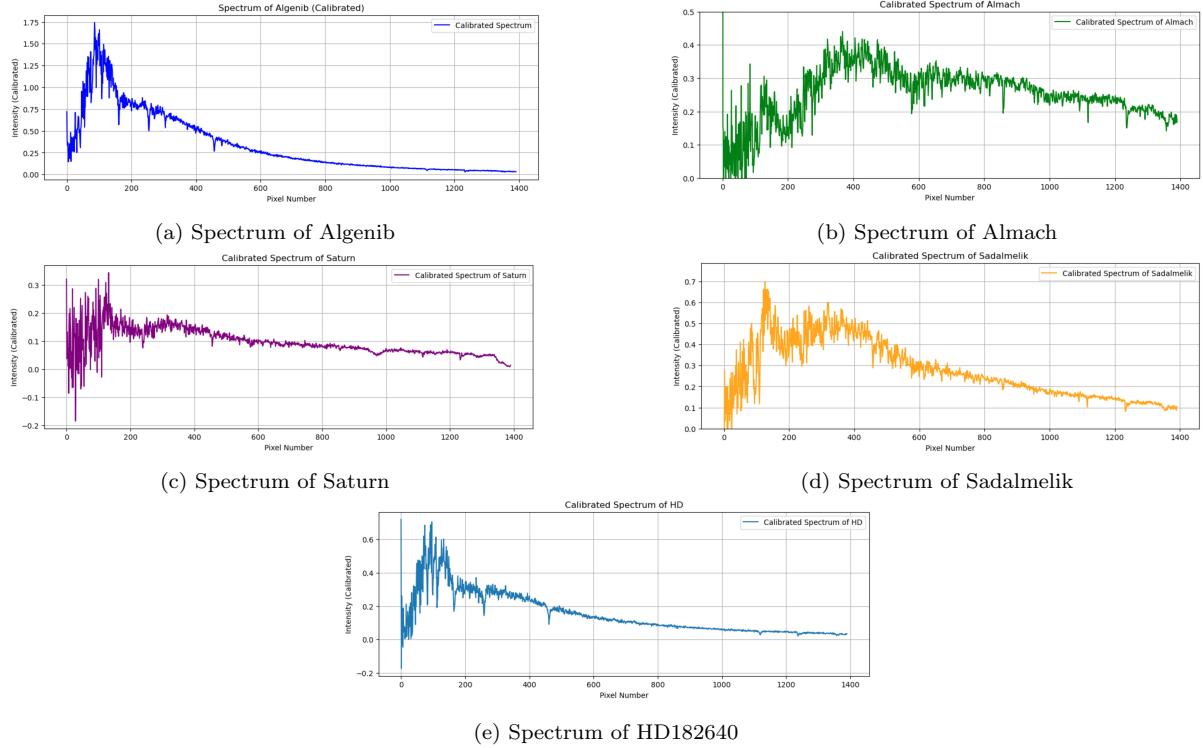


Figure 12: Combined calibrated spectra of astronomical targets: Algenib, Almach, Saturn, Sadalmelik, and HD182640. Each spectrum is individually calibrated, showing pixel number on the x-axis and calibrated intensity on the y-axis.

The calibrated spectra of the astronomical targets (Algenib, Almach, Saturn, Sadalmelik, and HD182640) in Figure 12 display intensity distributions as functions of pixel numbers and also reflect their distinct spectral characteristics. These spectra were achieved after calibrating each target. For this, the average of five dark noise files and five flat field files were taken in order to deduct noise from the actual data as shown in the Appendix. Algenib’s spectrum (a) shows a pronounced peak at lower pixel numbers, with a steep decline in intensity across the range to 1400 which is characteristic of a star with a strong emission in the blue or ultraviolet part of the spectrum which is expected due to Algenib’s high surface temperature of 21,000 K. Almach (b) exhibits a broader, more elevated intensity curve, suggesting a more uniform emission distribution across its observable spectrum. This is typical of complex stellar compositions such as Almach. Saturn’s spectrum (c) has lower intensity overall, with a lot more noise especially at the beginning, which could be due to weaker signal from a planetary body compared to stars. Sadalmelik’s spectrum (d) is more similar to Algenib’s but has a higher intensity plateau which indicates different stellar characteristics, possibly related to its spectral type. This is again expected due to Sadalmelik being a cooler G-type supergiant, whereas Algenib is a hotter B-type subgiant. Lastly, HD182640’s spectrum (e) shows a steep initial rise followed by a gradual decline, indicating a strong emission at certain wavelengths which would be influenced by its classification as a bright star.

These results do align with expectations to a certain extent, as each target’s spectral behaviour reflects its astrophysical nature. The variations in intensity and noise levels are also consistent with the diversity in brightness, distance, and type of the observed objects. However, the high noise levels in Saturn’s spectrum could potentially indicate calibration challenges or observational limitations like cloud cover or atmospheric turbulence. This is the most likely scenario as there were cloudy conditions especially during the data collection of Saturn. Further refinement of calibration processes could enhance the accuracy and consistency of these spectra and ensure that all features align better with theoretical expectations.

5 Discussion & Conclusion

This lab explored key techniques in astronomical spectroscopy such as spectral calibration, noise reduction, and data interpretation, using a USB650 spectrometer and the LISA telescope spectrograph. The

results emphasize the importance of calibration and uncertainty analysis in deriving accurate spectra for both laboratory and astronomical targets. The hydrogen spectrum with its prominent Balmer series lines, provided a reference for this calibration. The linear fit offered a reasonable approximation of the wavelength and pixel relationship, as shown by the narrow confidence band and minimal residuals near the centre of the range which were then compared to other higher-order polynomial models, with the quadratic fit proving the most effective due to its lower chi-squared value and random residual pattern.

The incorporation of uncertainties into the calibration process from the propagated variances of the linear regression parameters (slope of 1.02 ± 0.02 and intercept of -13.93 ± 12.33), added more rigour to the analysis. These uncertainties which were shown as confidence bands and error bars, accounted for noise in the spectrometer and the fit precision. This method was particularly important for telescope observations, where atmospheric interference and faint signals brought up additional challenges. The identification of the mystery gas as neon, based on the close match of detected emission lines to known neon wavelengths, further validated the resolving power of the spectrometer.

Despite this, several challenges and limitations were encountered. Stabilising the USB650 probe required improvisation and variations in integration times significantly influenced spectral clarity. Higher integration times flattened spectral peaks whereas shorter times introduced noise. During telescope observations, weather conditions such as cloud cover affected the quality of faint targets like the Ring Nebula which required longer integration times of 600 seconds that amplified noise. Additionally, the limited number of calibration points constrained the accuracy of higher-order polynomial fits, while detector noise, even after dark subtraction, remained a source of uncertainty in faint spectral regions. The residuals from the linear calibration, particularly at the extremes of the pixel range, revealed systematic effects that even the higher order models could not eliminate.

Future work could address these challenges through several improvements. Expanding the number of reference wavelengths used in calibration would significantly improve the reliability of polynomial fits especially for higher-order models. Advanced noise reduction techniques such as using adaptive filtering, could better isolate faint spectral features, while live feedback during data acquisition through a graphical interface, could stabilize the probe and optimise integration times, especially for the longer integration times. For telescope observations, better weather monitoring and scheduling would mitigate atmospheric effects to help improve the signal to noise ratio for faint targets. Other methods such as perhaps machine learning techniques could be employed to refine peak detection and centroid calculations, reducing systematic errors and improving spectral resolution.

In conclusion, this lab explored astronomical spectroscopy, showcasing the effectiveness of spectral calibration and data reduction techniques. The results aligned with theoretical expectations and literature values but highlighted the limitations of linear models and the need for consideration of uncertainties. The methodologies developed in this lab could also form a solid foundation for further studies in spectroscopy and demonstrate the potential for these techniques to improve the understanding of celestial bodies. With refinements in calibration and observation techniques, this approach can be extended to more complex astronomical datasets.

6 Bibliography

Department of Astronomy and Astrophysics, *Thermal Radiation & the Statistics of Noise*: Lab Manual for AST325/326 (Fall 2024), University of Toronto, October 2024.

<https://www.rasc.ca/sites/default/files/publications/JRASC-2009-08-hr.pdf>

Appendix:

Detector Calibration The calibration process involves subtracting dark noise and normalizing the detector's response using flat-field images. The equation used for calibration is:

$$P_i = \frac{R_i - D_i}{F_i - D_i} \quad (1)$$

where:

- P_i : The i^{th} pixel of the calibrated image.
- R_i : The i^{th} pixel of the raw image data for the target.

- F_i : The i^{th} pixel of the flat-field image.
- D_i : The i^{th} pixel of the averaged dark image.

Each step in this calibration is as follows:

- Subtracting the dark frame (D_i) from the raw data (R_i) removes fixed pattern noise and other artifacts introduced by the detector itself, such as thermal noise.
- Normalizing by the flat-field image ($F_i - D_i$) compensates for uneven illumination and pixel-to-pixel sensitivity variations, ensuring a consistent response across the detector.

A Derivation of Centroid and Wavelength Uncertainties

A.1 Centroid Calculation

Consider a spectral line recorded on a CCD or a similar device, where the pixel positions along the dispersion direction are given by x_i , and the corresponding recorded intensities by I_i . The centroid C of the line is defined as the intensity-weighted mean position:

$$C = \frac{\sum_i x_i I_i}{\sum_i I_i}. \quad (2)$$

A.2 Uncertainty in the Centroid

If we assume that each intensity measurement I_i follows Poisson statistics, then the variance in I_i can be approximated as $\sigma_{I_i}^2 = I_i$. Under this assumption, the variance in C (denoted as σ_C^2) can be derived by standard error propagation. Let $I_{\text{tot}} = \sum_i I_i$. Small variations in I_i will induce a variation in C . This leads to:

$$\sigma_C^2 = \frac{\sum_i (x_i - C)^2 I_i}{(I_{\text{tot}})^2}. \quad (3)$$

Taking the square root, the centroid uncertainty is:

$$\sigma_C = \sqrt{\frac{\sum_i (x_i - C)^2 I_i}{(I_{\text{tot}})^2}}. \quad (4)$$

A.3 Wavelength Calibration and Uncertainty

Assume a linear relationship between pixel position x and wavelength λ :

$$\lambda = mx + b, \quad (5)$$

where m (slope) and b (intercept) are determined by fitting known spectral lines at known wavelengths to their measured centroid positions. If the uncertainties in m and b are represented by σ_m and σ_b respectively, and there is a covariance term $\text{Cov}(m, b)$ between them, then the uncertainty in the wavelength λ , at a given pixel position x , can be derived from error propagation. For the linear model:

$$\lambda(x) = mx + b,$$

the variance in λ is given by:

$$\sigma_\lambda^2 = (x^2)\sigma_m^2 + \sigma_b^2 + 2x \text{Cov}(m, b). \quad (6)$$

Taking the square root gives the wavelength uncertainty:

$$\sigma_\lambda = \sqrt{(x^2)\sigma_m^2 + \sigma_b^2 + 2x \text{Cov}(m, b)}. \quad (7)$$

B Code

B.0.1 Code to Plot Wavelength against Pixel Position

```
1 # import the required libraries
2 import numpy as np
3 import matplotlib.pyplot as plt
4
5 # create the variables for the wavelenghts and the centroids
6 centroid_pos = np.array([440.85061644, 486.82908805, 656.8346545])
7 wavelen = np.array([434.05, 486.13, 656.28])
8 cent_errors = np.array([0.10, 0.10, 0.10])
9
10 # performing a linear fit with covariance and using wavelength = slope * centroid + intercept
11 p, cov = np.polyfit(centroid_pos, wavelen, 1, cov=True)
12 slope, intercept = p
13 var_slope = cov[0, 0]
14 var_intercept = cov[1, 1]
15 sigma_slope = np.sqrt(var_slope)
16 cov_slope_intercept = cov[0, 1]
17 sigma_intercept = np.sqrt(var_intercept)
18
19 # finding the fitted line and uncertainties
20 x_fit = np.linspace(min(centroid_pos)-10, max(centroid_pos)+10, 200)
21 y_fit = slope * x_fit + intercept
22 # calculating the error on the fit
23 y_fit_err = np.sqrt((x_fit**2)*var_slope + var_intercept + 2*x_fit*cov_slope_intercept)
24
25 # plot data, fit, and uncertainty
26 plt.figure()
27 plt.errorbar(centroid_pos, wavelen, xerr=cent_errors, fmt='o', color='blue', label='Data')
28 plt.plot(x_fit, y_fit, 'r-', label='Linear Fit')
29 plt.fill_between(x_fit, y_fit - y_fit_err, y_fit + y_fit_err, color='red', alpha=0.2, label='1 Confidence Band')
30 plt.title('Wavelength vs. Pixel Position Calibration with Uncertainties')
31 plt.xlabel('Pixel Position (Centroid)')
32 plt.ylabel('Wavelength (nm)')
33 plt.grid()
34 plt.legend()
35 plt.show()
36
37 print(f"Slope (m): {slope:.4f}    {sigma_slope:.4f}")
38 print(f"Intercept (b): {intercept:.4f}    {sigma_intercept:.4f}")
39 print("Covariance Matrix:")
40 print(cov)
41
42 # for the residual
43 fitted_vals_at_data = slope * centroid_pos + intercept
44 residuals = wavelen - fitted_vals_at_data
45 # propagating uncertainty to residuals:
46 residual_errors = np.sqrt((centroid_pos**2)*var_slope + var_intercept + 2*centroid_pos*cov_slope_intercept + (slope**2)*(centroid_errors**2))
47
48 plt.figure()
49 plt.errorbar(centroid_pos, residuals, xerr=cent_errors, yerr=residual_errors, fmt='o', color='green', label='Residuals')
50 plt.axhline(0, color='black', ls='--')
51 plt.title('Residuals of the Fit')
52 plt.xlabel('Pixel Position (Centroid)')
53 plt.ylabel('Residual (nm)')
54 plt.grid()
55 plt.legend()
56 plt.show()
```

Listing 1: Python code for plotting data and residuals and calculating uncertainties.

B.0.2 Code to Plot Hydrogen Arc Lamp Spectrum

```

2 # import the libraries
3 import numpy as np
4 import matplotlib.pyplot as plt
5
6 # input slope, intercept and covariance from previous calcs
7 slope = 1.0213
8 intercept = -13.9339
9 Var_m = 5.28317840e-04
10 Var_b = 1.51941321e+02
11 Cov_mb = -2.79042401e-01
12
13 # get pixel positions and intensities
14 x = hydrogen_corrected["Wavelength"].values
15 I = hydrogen_corrected["Intensity"].values
16
17 # pixels to wavelengths using the linear fit
18 map_wave = slope * x + intercept
19
20 # find the wavelength uncertainty for each point
21 var = (x**2)*Var_m + Var_b + 2*x*Cov_mb
22 unc = np.sqrt(var) # standard deviation (1 ) in wavelength
23
24 # plotting the spectrum
25 plt.figure(figsize=(10, 6))
26 plt.title('Hydrogen Arc Lamp Spectrum')
27 plt.errorbar(map_wave, I, xerr=unc, fmt='-', color='blue', ecolor='gray', capsize=2,
28               label='Hydrogen Arc Lamp Spectrum')
29 plt.xlabel('Wavelength (nm)')
30 plt.ylabel('Intensity (a.u.)')
31 plt.grid(ls='--')
32 plt.legend()
33 plt.show()

```

Listing 2: Python code for plotting hydrogen arc lamp spectrum with the wavelength uncertainties.

B.0.3 Code to Plot Neon Calibration Source Image

```

1 # import all required libraries
2 from astropy.io import fits
3 import numpy as np
4 import matplotlib.pyplot as plt
5
6 # define paths and load data
7 neon_file_path = '2024-11-07_Neon_20s.fit'
8 dark_file_path = 'Darks_Night-0002_20s_.fit'
9 neon_data = fits.getdata(neon_file_path)
10 dark_data = fits.getdata(dark_file_path)
11
12 # do the dark subtraction
13 calibrated_neon_data = neon_data - dark_data
14
15 # show the Neon Calibration Source Image
16 plt.figure()
17 plt.title('Neon Calibration Source Image (Dark Subtracted)')
18 plt.imshow(calibrated_neon_data, cmap='gray', origin='lower')
19 plt.colorbar(label='Intensity')
20 plt.xlabel('Pixel X')
21 plt.ylabel('Pixel Y')
22 plt.show()

```

Listing 3: Python code for plotting Neon Calibration Source Image.

B.0.4 Code to Plot Neon Calibration Source Image

```

1 # import all required libraries
2 from astropy.io import fits
3 import numpy as np
4 import matplotlib.pyplot as plt

```

```

6 # define paths and load data
7 neon_file_path = '2024-11-07_Neon_20s.fit'
8 dark_file_path = 'Darks_Night-0002_20s_.fit'
9 neon_data = fits.getdata(neon_file_path)
10 dark_data = fits.getdata(dark_file_path)
11
12 # do the dark subtraction
13 calibrated_neon_data = neon_data - dark_data
14
15 # show the Neon Calibration Source Image
16 plt.figure()
17 plt.title('Neon Calibration Source Image (Dark Subtracted)')
18 plt.imshow(calibrated_neon_data, cmap='gray', origin='lower')
19 plt.colorbar(label='Intensity')
20 plt.xlabel('Pixel X')
21 plt.ylabel('Pixel Y')
22 plt.show()
23

```

Listing 4: Python code for plotting Neon Calibration Source Image.

B.0.5 Code to calculate centeroids

```

1 # first have to initialise array
2 rows = np.arange(calibrated_neon_data.shape[0])
3 centroids = []
4
5 # then calculating centroid for each row
6 for row in rows:
7     row_data = neon_data[row, :]
8     total_intensity = np.sum(row_data)
9     # avoiding division by zero
10    if total_intensity > 0:
11        centroid = np.sum(np.arange(len(row_data)) * row_data) / total_intensity
12    else:
13        # if no other data
14        centroid = np.nan
15    centroids.append(centroid)
16 centroids = np.array(centroids)
17

```

Listing 5: Python code to calculate centeroids

B.0.6 Code to Plot Spectra of Algenib

```

1
2 # load dark frame average
3 dark_frame = fits.getdata('master_dark_40s.fits')
4
5 # loading in raw target data
6 raw_target = pd.read_csv('Algenib_40s.csv', header=None).values
7 # first column is pixel numbers
8 raw_target_pixels = raw_target[:, 0]
9 # second column is intensity
10 intensity = raw_target[:, 1]
11
12 # load in the flat frame data (CSV)
13 # the master flat has average of the other flat
14 flat_frame = pd.read_csv('master_flat.csv', header=None).values
15 flat_frame_pixels = flat_frame[:, 0]
16 flat_frame_intensity = flat_frame[:, 1]
17
18 # do the calibration using the equation
19 calibrated_intensity = (intensity - dark_frame[0, :len(intensity)]) / (
20     flat_frame_intensity - dark_frame[0, :len(flat_frame_intensity)])
21
22 plt.figure(figsize=(12, 4))
23 plt.title("Spectrum of Algenib (Calibrated)")
24 plt.plot(intensity, calibrated_intensity, label="Calibrated Spectrum", color="blue")
25

```

```
26 plt.xlabel("Pixel Number")
27 plt.ylabel("Intensity (Calibrated)")
28 plt.legend()
29 plt.grid()
30 plt.show()
```

Listing 6: Python code to Plot Spectra of Algenib