**PHYS40001: Computing Project**

# Measuring Hubble's Constant

## Introduction

In the field of cosmology, which studies the origin and evolution of the universe, one of the most fundamental observations is that of Hubble's Law. The universe is expanding, which means the distance between objects is increasing with time. Hubble's Law states that the velocity at which objects, such as galaxies, are moving away from the Earth is proportional to their distance from Earth

$$v = H_0 D$$

where $v$ is the velocity, $D$ is the distance and $H_0$ is the value of Hubble's constant.

The distance to a galaxy is inferred from the brightness of the galaxy. The velocity is inferred from the Doppler-shift of light emitted by the galaxy. For a galaxy that is moving away from the observer, the light will appear to the observer to be "redshifted" to longer wavelengths. The relationship between the velocity of the galaxy and the shift in wavelengths is given by the redshift formula

$$\frac{\lambda_o}{\lambda_e} = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}$$

Where $\lambda_o$ and $\lambda_e$ are the observed and emitted wavelengths of light, $c = 2.9979 \times 10^8\ m\ s^{-1}$ is the speed of light and $v$ is the velocity of the emitter. The Hydrogen-alpha spectral line (Hα) is a deep-red visible line with wavelength $656.28\ nm$. This line is commonly observed in the emission spectra from galaxies.

Figure 1 illustrates how $H_0$ can be inferred from measurements of distance and the "redshifted" velocity.

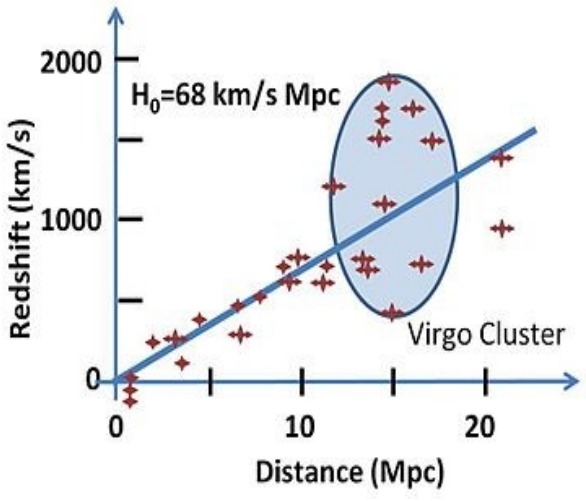



Fig. 1: A plot of velocity inferred from redshift versus distance for a range of galaxies. The slope of the line of best fit gives an estimate for $H_0$.

## Project Objective

Given 30 observations of the redshifted Hα line from different galaxies and the distance to each galaxy, calculate a value for Hubble's constant. The output from your solution code should be a plot showing velocity inferred from redshift vs. distance for each galaxy, the value of Hubble's constant inferred from fitting this plot and its uncertainty.

## Data

Here is a description of the data that you are given.

The file called "Halpha_spectral_data" contains data for the observed shift of the (Hα) spectral shift. It contains 30 files in .csv format. Each file consists of 2 header rows, followed by 2 columns of data. The first column is wavelength in $nm$ and the second column is spectral intensity in arbitrary units (see example in fig. 2). The first row of the header contains the following information: Date at which the observation occurred, the name of the person who made the observation, the observation number and the instrument response for that observation.

The file called "distances_Mpc" is a .csv file in which the first row is a header. It has two columns of data. The first column is the observation number and the second column is the measured distance to that galaxy.

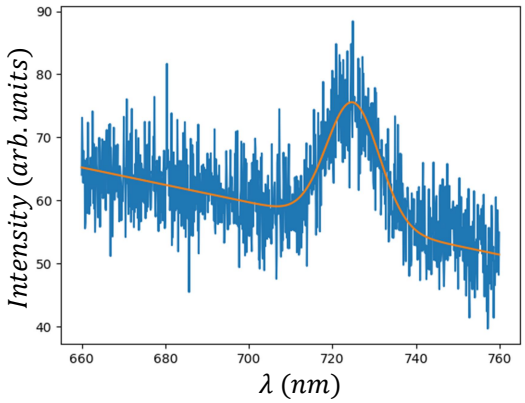



Fig. 2: A plot of spectral intensity vs. wavelength. The blue line shows the observed data. The orange line is obtained by fitting a straight line + Gaussian function to the data.

## Project Notes

The data for spectral intensity versus wavelength is noisy, as can be seen in fig. 2. In order to use it to calculate the velocity, the data should first be fitted with a combination of a straight line and a Gaussian function. A similar fitting procedure was covered in the final exercise of Core Worksheet 2. Once the data is fitted, the mean value of the fitted Gaussian is a good estimate for the observed wavelength, $\lambda_o$, of the shifted Hα line.

The instrument response for each observation is either "Good" or "Bad". A bad instrument response can occur if, for example, we have low signal intensity, there is atmospheric interference of the signal or there is a drift in wavelength calibration due to thermal expansion of the instrument. These observations should not be used in the calculation of $H_0$.

Finding the distance for each observation involves matching the observation number in the header of the spectral data file with those listed in the first column of the distance data. An elegant solution to both this problem and the instrument response mentioned above is to write code which can read the header in the spectral data files. There are some hints for doing this in the Supplementary Python Knowledge section at the end of this document.

## Project Submission

You are required to submit your source code (.py file), a plot of redshift velocity vs. distance and the fitted value of Hubble's constant and its uncertainty. The deadline for submission is 5pm on Thursday 14th November. Full details on how to submit will be given out shortly.

# Supplementary Python Knowledge

This section describes some Python commands that were not covered in the Computing Lab sessions but which could be useful for the project. You are not required to use these commands but they could make your code neater and more elegant.

## *Context Manager and Opening Data:*

When we loaded in data during the lab session we used the sp.loadtxt command. This is good when we want to load in all the data at once. However, there are instances where we want to open a file and then only read a few lines, and maybe not the whole file (e.g. to read the header but not the data in the file). We can do this with the "readline()" function, which is a file variable associated function.

The following code block is known as a context manager, and it allows us to have more control over how we manipulate our files. We open our data called filename (and use the 'r' keyword to tell python we are going to read this data) and then assign it to the file variable. We use a colon and indent our code to show it is part of the context manager (just like for, if and while loops). This can be followed by the readline() function, as shown below. This will read a single row from a file. We can chain these together to read multiple lines from a file. Note that once we leave the context manager the file is automatically closed.

```
with open('filename','r') as file:
    line1=file.readline()
    line2=file.readline()
```

## *String Manipulation:*

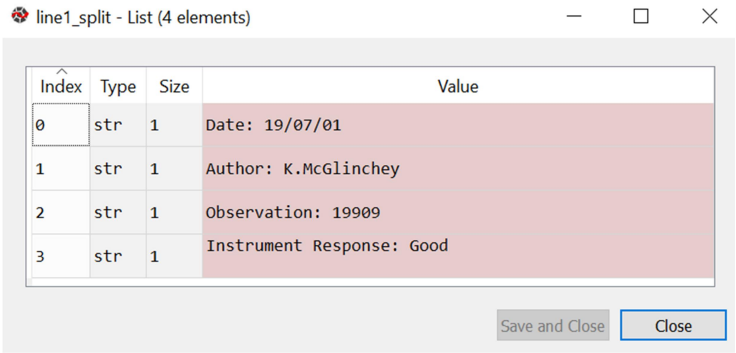
When we read the first line of the file we opened with the content manager, we end up with a variable that looks like:

| | | | |
|---|---|---|---|
| line1 | str | 1 | Date: 19/07/01 ,Author: K.McGlinchey ,Observation: 19909 ,Instrument R ... |
| line2 | str | 1 | Wavelength (nm) ,Intensity(A.U.) |

The whole first line of the file has been read in to a single string variable, with no separation for the different cells in the csv file. However, there is important information in this line that you may want to extract as its own separate variable, such as the observation number and if the instrument response is good. To accomplish we can use the following functions on strings to manipulate them:

```
line1_split=line1.split(',')
```

Which then leads to:



The split command takes an input argument which is the value to split the string on, and then returns a list of the split string.

Another useful command is the .strip() command, which removes whitespace and any special characters from your string. For example, in the following code the str2 variable will be a string with the whitespace removed from the beginning and end of str1.

```
str1 = ' Instrument Response: Bad '
str2 = str1.strip()
```

## *Automating File Readin:*

In this project you will have to query lots of different files, and opening up each one individually is not an efficient way to code. There exists the functionality in python to search for the patterns in a folder that contains files and return the names of the files that match the pattern. This is accomplished using the glob module.

```
import glob as glob
file_list=glob.glob('*.csv')
```

The glob.glob command will search for the patterns outlined in the parantheses. For this case, the * is a special character and essentially means 'match number of characters. The command will search the current folder for all files containing the .csv file extension. This list can then be iterated over with a for loop and each file loaded in in this manner.