

Summary of the Galaxy comparison code

Francisco Machado

March 25, 2014

0.1 Introduction

The goal of this document is to provide an understanding of the code written to do the comparisons between galaxies. I will cover each file and explain how the input is taken and then the workings of the program.

0.2 Format_galaxies.py

The purpose of this file is to preprocess the simulation galaxy's SED, transforming the data from Luminosity into AB magnitude at a distance of 10 pc.

0.2.1 Input

The input for this program is:

```
python3 Format_galaxies.py Input_Folder Output_Folder
```

The Input_Folder is the origin place of the files and the Output_Folder is the directory where the modified files will be stored.

The program runs through every .txt file and runs the appropriate transformations. Right now it assumes that the file consists **only** on two columns, the first one containing the wavelength (λ) and the other the Luminosity per wavelength (L_λ) at a distance of 50 Mpc in SI units (Wm^{-1}).

Moreover the code also assumes that when a row is split by the spaces (' ') using the string.split(' ') method from python, the second column corresponds to the FILE.SECOND_COLUMN position in the resulting array. Changes in the input file format must be taken into consideration by changing this value.

0.2.2 Data processing

To turn the data into AB magnitudes the program follows the following steps.

Firstly the value of Luminosity per wavelength is transformed to luminosity per frequency (L_ν), using the equation:

$$L_\nu = L_\lambda \frac{\lambda^2}{c} \quad (1)$$

This value will have units of WHz^{-1}

Then the luminosity is transformed into flux (F_ν) at a radius of $R = 10 \text{ pc}$, by dividing it by the area of a sphere of such radius.

$$F_\nu = L_\nu \frac{1}{4\pi R^2} \quad (2)$$

F_ν will now have units of $\text{WHz}^{-1}\text{m}^{-2}$

Then we make the conversion into flux in microJanskys (S_ν):

$$S_\nu = F_\nu \times 10^{32} \quad (3)$$

The units of S_ν are $\mu \text{ Jy}$.

Then we use the formula to calculate the AB magnitude from the flux in $\mu \text{ Jy}$:

$$m_{AB} = 23.9 - \log_{10}(S_\nu) \quad (4)$$

m_{AB} will not have units but determines the apparent AB magnitude of the galaxy.

Since we calculate the flux at a distance of 10 pc this apparent magnitude matches the absolute magnitude (M_{AB}).

$$M_{AB} = m_{AB} \quad (5)$$

For the remainder of the article let's define the absolute magnitude of a simulation galaxy as M_S .

0.2.3 Saving the Data

The code then saves each file into a new file of the same name of the original on the Output_Folder. Each file consists on two columns separated by a space (' ') where the first column is the wavelength (λ) and the second will be the absolute AB magnitude for that wavelength (M_{AB}).

0.3 FastComp.py

This program has compares an input galaxy with a set of galaxies, by their absolute AB magnitude, using the χ^2 minimization method. It returns a file with the value of χ^2 for every galaxy in the comparison set, as well as drawing a graph and histogram with the same information.

0.3.1 Input

The input for this program is:

```
python3 FastComp.py Input_File Data_Folder
```

The Input_File is a file containing the information about the galaxy we wish to compare with our set. The file must be formatted in the following way. The first line must contain a single number which will correspond to the redshift of the galaxy. The remainder of the file must contain 3 columns, which will be the wavelength (λ), the flux in μ Jy (S_ν), and the error in the flux (δS_ν) also in units of μ Jy. Each of the columns is separated by a space (' ').

0.3.2 Data Pre-Processing

The program starts by reading the information of the file and storing the redshift on the variable called (appropriately) **redshift**. The remaining values are stored on an dictionary of arrays where the 'wl' corresponds to the wavelengths, 'fl', the fluxes and 'err' the error in the flux.

Then the distance in Mpc (D_{Mpc}) is calculated using the cosmocalc package, using the following command (included to specify the constants):

```
(cosmocalc.cosmocalc(redshift, H0=70.4, WM=0.2726, WV=0.7274))['DL_Mpc']
```

Then the distance is transformed into pc, (now D_{pc}) and used to calculate the distance modulus (μ) which will then be used for calculating the variation in the magnitude of the Input_File.

$$\mu = 5 \cdot (\log_{10}(D_{pc}) - 1) \quad (6)$$

Now for each data point in the Input_File we tranform the flux (S_ν) into an absolute AB magnitude (M_{AB}) (AB magnitude measured at a distance of 10 pc). To do this, the flux is firstly used to calculate the apparent AB magnitude (m_{AB}) and then we make use of the modulus distance (μ) to transform from apparent to absolute.

$$m_{AB} = 23.9 - 2.5 \cdot \log_{10}(S_\nu) \quad (7)$$

$$M_{AB} = m_{AB} - \mu \quad (8)$$

The error will also be propagated.

$$\begin{aligned} \delta m_{AB} &= 2.5 \frac{\delta S_\nu}{S_\nu} \\ \delta M_{AB} &= \delta m_{AB} \end{aligned}$$

Here we are assuming that there is no error in the distance since we have no error in the redshift which we can take into account.

For the remainder of the articles let's define the absolute magnitude of an input galaxy as M_I , and its error as δM_I .

0.3.3 Comparison Method

To compare the Input_File with the set of galaxies we use the χ^2 minimization method by comparing to every galaxy in our set and select the galaxies which present the least value of χ^2 .

For this we open and read each file and one at a time calculate the χ^2 term.

If the values of the wavelengths match with the ones we have for the set of galaxies, then we simply calculate the χ^2 term directly:

$$\chi^2 = \frac{1}{N} \sum_{k=1}^N \frac{(M_S - M_I)^2}{\delta M_I} \quad (9)$$

If the wavelengths do not match, then we must do an interpolation between the two neighboring points to gather the correct value we should use for M_S .

Let λ_S^+ and λ_S^- be the closest values of the wavelengths to λ_I such that $\lambda_S^+ > \lambda_I > \lambda_S^-$. Also let M_S^+ and M_S^- be the values of the absolute magnitude of the simulation galaxy for wavelengths λ_S^+ and λ_S^- respectively.

Then we can define an interpolated value for the absolute magnitude $M_S^{interpol}$ and then do the comparison with that term.

$$M_S^{interpol} = \frac{M_S^+ - M_S^-}{\lambda_S^+ - \lambda_S^-} (M_I - M_S^-) + M_S^- \quad (10)$$

$$\chi^2 = \frac{1}{N} \sum_{k=1}^N \frac{(M_S^{interpol} - M_I)^2}{\delta M_I} \quad (11)$$

Note, if there are values of wavelengths in the input file greater than the largest wavelength in our simulated galaxies' files or smaller than the smallest one in our simulated galaxies' files, then we have no way of doing a precise comparison with those and they are disregarded.

0.3.4 Saving the Data

The program then saves the resulting χ^2 of all the file comparisons to a file in the Results folder. The program draws a bar graph of this information, with the χ^2 for each galaxy, saving it into the Graphs folder. The program also draws a histogram of the distribution of galaxies with a certain χ^2 value, saving it to the Hist folder.

Finally the program also displays the top 4 best and worst fitting galaxies in the terminal. This is meant to be a quick and easy way to display the most relevant information so that it can be quickly plotted using the `Plot_mag.py` program.