

Università degli Studi di Padova

Dipartimento di Fisica e Astronomia “G. Galilei”
Master Degree in Astrophysics and Cosmology

FINAL DISSERTATION

The evolution of the night sky spectrum in Asiago

Candidate:
Marco Codato
Badge number 2023377

Supervisor:
Prof. Sergio Ortolani
Co-supervisor:
Prof. Stefano Ciroi

Abstract

Modern sky brightness monitoring techniques aim to precisely measure the total amount of radiation from the observing site but very little can be said about the various sources responsible for such radiation. In this work I use spectra acquired in the last 15 years from the Asiago Observatory to identify the various sources in the sky and study their temporal evolution.

Contents

Abstract	3
1 Introduction	7
1.1 Light pollution	7
1.2 Aim of the work	8
2 The natural sky background	9
2.1 Extraterrestrial sources	9
2.1.1 Zodiacal light	9
2.1.2 Galactic background	11
2.1.3 Diffuse galactic light	11
2.1.4 Extragalactic background	11
2.2 Terrestrial sources	11
2.2.1 Airglow	11
2.2.2 Aurorae	12
2.3 Atmospheric scattering	13
2.3.1 Scattering mechanisms	13
2.3.2 Effects on the sky brightness	14
3 Artificial sky light pollution	17
4 The spectrum of the sky	19
5 Sky spectra reduction	21
5.1 Introduction	21
5.1.1 Software management and reduction steps	21
5.1.2 The dataset	21
5.2 Background extraction	22
5.2.1 Preamble	22
5.2.2 Data extraction	23
5.2.3 Cosmic ray and noise removal	24
5.2.4 Sources identification	25
5.2.5 Sky spectrum export	29
5.3 Background analysis	29

Chapter 1

Introduction

Part I: Sky sources in theory

- Introduction
 - General introduction
 - Aim of the work
 - About the methodologies
- The natural sky
 - Main natural sources
 - And their footprint on spectra
- The light pollution
 - Definitions and aftermaths
 - Mechanism of working
 - Mention to the models in the literature
 - LP footprint in spectra

Part II: Analysis of sky background in spectra

- Software description
 - Bkg extraction
 - Bkg analysis
- Results
- Discussion and interpretation of the results
- Conclusions

1.1 Light pollution

Light pollution (LP) is the alteration of the natural light level due to artificial sources. The resulting increase of the sky brightness has many proven negative effects.

Effects on the human health. Light exposure in night time decrease the natural production melatonin. The effect is proportional to the frequency of light, with bluer radiation producing a stronger decrease of melatonin production.

Melatonin is an important hormone that regulates many biological mechanisms. It is capable of preventing some forms of cancer and is responsible for the sleep regulation. A melatonin deficiency has been proven to be correlated with higher chances of developing breast and prostate cancers and a decrease of sleep time and quality, which typically lead to further health disorders.

Melatonin decrease is proportional both to light intensity and frequency. A greater effect is given by brighter and bluer sources. In this context the spreading of LED lights, with their strong emissions in the blue side of visible spectrum, is considered a concern by many health associations.

Effects on the environment. LP affects other living beings as well as humans. Animals exposed to abnormal level of light at night change their behaviour and habits. Note this form of pollution is probably the most widespread but yet one of the least acknowledged.

Economical effects. When looking at a artificially bright sky one should consider that such photons that brighten the sky are no longer being used for the purpose they were made for, i.e. lighten streets, houses, commercial areas and so on. The energy, and thus the cost, to produce such photons is wasted.

Unluckily in the last years efficient light sources like LEDs allowed to produce powerful lighting systems at low cost making the economical argument less relevant. Since light is cheaper, it is less critical weather part of it is lost toward the sky.

Cultural effects. All the cultures around the world developed myths and legends involving the heavens; night sky inspired artists and philosophers in western cultures for centuries and in general the observation of a starry sky always belonged to the human experiences. Today due to LP FabbriXX estimates that at least the XX% of the world population lives in areas where milky way is not even visible and only a handful of bright stars can stand out of the polluted sky. In terms of traditions and human experience this is a ?great loss, but yet difficult, or impossible, to quantify.

Scientific effects. Of course the increase of sky brightness made astronomical observations more difficult. Observation sites moved from the town centres in the XIX century to the rural areas due to the introduction of the first lighting. With the growing urbanization, many of these sites ended up to be at the limb of the expanding urban areas, heavily limiting the possibility of relevant scientific activities. Nowadays it is likely that in a country no totally dark sites are available, forcing astronomers to build new instruments in very remote areas in poorly populated areas of the world.

A typical example of the effects in the changing of the sky condition is the Asiago observatory. It was built in 1942 in a poorly populated highland, which also offered an adequate shielding from the light of the yet small rural centres in the nearby pianura veneta. When built, the observatory also hosted the largest reflecting telescope in the Europe (Gaileo telescope, 122 m of diameter). With the economic boom in the 50s, industrial and manufacturing activities replaced agriculture in the Veneto flatland. Urban areas significantly expanded making Veneto region one of the most light polluted sites in the whole Europe. At the same time the Asiago highland become one of the most appreciated touristic destination in the surrounding area. The quality of the sky rapidly worsened also with respect to other nearby areas less touched by human activities. In such new condition the Asiago Observatory lost its central role in research activities tough preserving its nature of scientific pole.

1.2 Aim of the work

For all the issues above measuring and monitoring the LP is of crucial important.

Chapter 2

The natural sky background

Even when artificial sources are neglected, there are still several natural background sources. In this chapter each contribution will be described in detail. In the plot on the Figure 2.1 are reported, in logarithmic scale, the main background sources for a wide range of wavelengths, from UV to radio emission. In the next lines I will consider only sources relevant for optical observations.

Photons produced by several background sources interact with dusts and aerosols in the Earth atmosphere leading to extinction and diffusion (scattering) phenomena. The total sky brightness can be quantitatively expressed as

$$I_{\text{sky}} = (I_A + I_{ZL} + I_{ISL} + I_{DGL} + I_{EBL})e^{-\tau} + I_{\text{sca}} \quad (2.1)$$

where A stands for airglow, EZ is zodiacal light, ISL integrated galactic light, DGL diffuse galactic light and EBL extragalactic background light. τ is the extinction coefficients and I_{sca} gathers all the scattering terms, i.e. light scattered from previous sources and from light pollution, from [LBH⁺98].

In this chapter I will analyze in detail all the terms of the equation (2.1), their origin and the relevance and characteristic spectral features. At the end I will show and discuss some examples of sky background spectra taken from the literature.

2.1 Extraterrestrial sources

I will first consider sources of photons outside the earth atmosphere. Using space-based instrument it is possible to study these components without the interference of atmospheric emissions.

2.1.1 Zodiacal light

Zodiacal light consists on sunlight scattered by interplanetary dust particles [Lei75]. From the Earth it looks like a white glow visible during the twilight and extending from the Sun in the zodiacal region.

Angular distribution. The figure 2.3, adapted from [FHLT74], describes the angular distribution of the zodiacal light in ecliptic coordinates. Such light is maximum along the ecliptic and close to the Sun. A fainter local maximum is present in direction opposite to the Sun. It is known as *gegenschein* and is produced by back-scattered solar light. Zodiacal light brightness varies from about 10^{-6} erg/s/cm²/sr/Å on the ecliptic at 30° from the sun to about 10^{-9} erg/s/cm²/sr/Å. Gegenschein maximum brightness reaches about 10^{-7} erg/s/cm²/sr/Å. After the airglow (see §XX), this is the second brightest background source in optical bands.

The contribution of zodiacal light to the optical background is maximum during the twilight, after sunset in spring or before sunrise in autumn, from the northern hemisphere.

Spectral energy distribution. Being essentially reflected sunlight, the optical zodiacal light energy distribution has the same shape of the solar one.

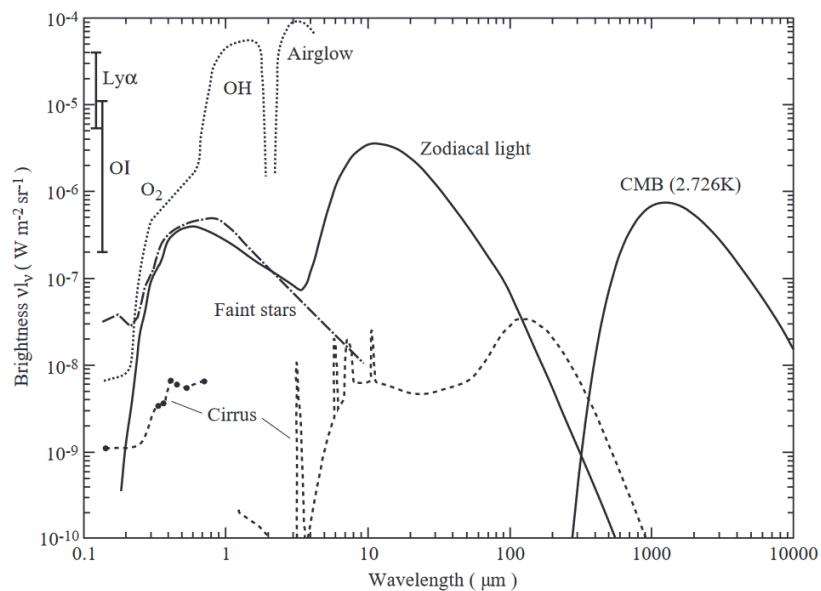


Figure 2.1: Different sky brightness contributions in different electromagnetic domains. In the optical band most relevant contribution are Airglow, zodiacal light and faint stars. From [LBH⁺98].



Figure 2.2: Zodiacal light after sunset at La Silla, Chile. Source: eso.org.

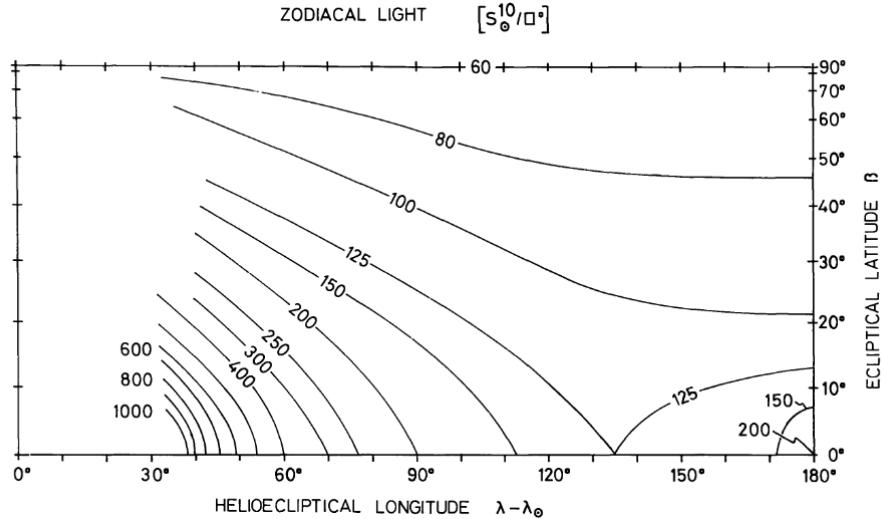


Figure 2.3: Isophotal map of the zodiacal light at 7100 Å. As a reference, according to [LBH⁺⁹⁸], $1 S_{10}^{10}/\Delta^2 = 9.21 \times 10^{-10} \text{ erg/s/cm}^2/\text{sr}/\text{\AA}$ at those wavelengths. From [FHLT74].

2.1.2 Galactic background

In optical bands a significant contribution to the background level is provided by unresolved stars in the Galaxy. The contribution of such sources depends on the ability of resolve the brightest stars [LBH⁺⁹⁸], i.e. on the limiting magnitude of the instrument.

Angular distribution. Unresolved stars background follows the morphological structure of the Milky Way. The signal is higher toward the galactic plane and the galactic center. Its spectrum follows typical optical stellar spectra with the characteristic black-body emission. It is the third most relevant contribution to the optic continuum with an emission that spans from peak values of $10^{-6} \text{ erg/s/cm}^2/\text{sr}/\text{\AA}$ in the most crowded areas to $10^{-8} \text{ erg/s/cm}^2/\text{sr}/\text{\AA}$ toward the galactic poles.

2.1.3 Diffuse galactic light

Similar to the zodiacal light, it is the result of the scattering of stellar emission with interstellar dust particles. [LBH⁺⁹⁸] estimate its contribution as between 20% and 30% of the total integrated light from the galaxy. This estimation is rather uncertain because of the faintness of the radiation and the contamination of direct stellar light. There are no comprehensive maps for the diffuse galactic light but it is very likely this emission to be concentrated along the galactic disk, analogously to the direct stellar component. Its spectral energy distribution is comparable with stellar spectra, since its nature of stellar reflected light.

2.1.4 Extragalactic background

A much smaller contribution is led by the extragalactic background, i.e. emission of faint and or unresolved galaxies. It is very difficult to quantify the resulting brightness and in many cases are available only the upper limits for extragalactic background. The main estimation difficulties are due to the faintness of the signal and with respect to the other sources. Typical values of intensity are of the order $10^{-9} \text{ erg/s/cm}^2/\text{sr}/\text{\AA}$. No reliable information about spatial distribution is available.

2.2 Terrestrial sources

Terrestrial sources are those capable of producing visible photons in the atmosphere.

2.2.1 Airglow

The airglow is the faint emission on the higher layers of the atmosphere, produced by the interaction between atoms and the particles from the solar wind or by the chemical interaction between atoms. High



Figure 2.4: Oxygen (green) and sodium (orange) airglow emission, photographed from the ISS. Source: eol.jcs.nasa.gov.

energy solar particles collide with the atmospheric atoms exciting their electrons to higher energy levels; when the electrons jump back to the initial states they release energy in form of photons, leading to the characteristic emission spectrum. Another possible emission channel is by chemical recombination: when atomic oxygen collide with nitrogen or hydrogen atoms a single molecule (NO or OH) is created and a photon is released. Atomic oxygen or nitrogen are produced by photodissociation of the respective molecules during the day by solar radiation.

Main components. We can subdivide the airglow sources as a function of the height of the emitting layer. A first layer between 85 and 100 km is provided by molecular oxygen, sodium (respectively Herzberg bands and Fraunhofer D line) and OH transitions. Going higher, up to 300 km, forbidden atomic oxygen lines are produced. The outermost layers of the atmosphere, above 1000 km, are usually referred as geocorona; in this region faint but detectable hydrogen lines are produced.

Being produced by thin and homogeneous layers, the airglow emission is relatively uniformly distributed in the sky sphere, with an increase of brightness at high zenithal distances due to the increase of geometric depth along the line of sight. Maximum brightness is achieved at about 10° above the horizon after that the overall brightness is dimmed by atmospheric extinction. Brightest lines can produce a brightness up to 10^{-5} erg/s/cm²/sr/Å

Variations in airglow emission. Airglow emission varies in time, both on short and long timescales, following the behavior of the atmosphere and the solar activity [LBH⁺98]. Emission is also related to the geomagnetic latitude: is maximal in the sub-polar region, at a latitude of about $60^\circ - 80^\circ$ after which it significantly drop. In the polar region airglow emissions are substituted with auroral emission. In the low latitude regions emissions are generally low with a slight increase toward the equator [Eat69].

2.2.2 Aurorae

Aurorae are bright light bands observable at polar latitudes. They are produced by the excitation of atoms in the high layers of the atmosphere by the solar wind. At high latitudes interplanetary high energy charged particles can penetrate the magnetosphere ad reach the atmosphere where they collide and excite atmospheric elements. Excitation energy is then released in form of a photon, responsible for the observed radiation. Auroral spectrum is constituted by emission lines. Colors ranges from green and orange (typical of oxygen transitions) to blue or purple (trace of nitrogen emission), see figure 2.5.

The occurrence and intensity of this phenomenon is strictly regulated by the solar activity, and an increase of auroral emission can be observed during solar storms or period of high activity. The



Figure 2.5: Aurora borealis in the northern Finland. Credits: Martincco.

phenomenon is observable only in polar regions, namely above 80° of latitude (see [Eat69]) and for this reason it has a limited impact on the total optical sky background only in that geographic area. Nevertheless in case of intense solar activity like solar storms, aurorae can be observed at lower latitudes. Historical sources even report the sporadic observation of aurorae up to temperate latitudes.

2.3 Atmospheric scattering

Atmospheric scattering refers to the interaction of light with aerosol particles in the atmosphere. Unlike reflection or refraction, when a light beam get scattered its photons are deflected in random directions. Note scattering is somehow the complementary of extinction: when a light beam is dimmed due to extinction it means that part of its photons has been deflected (thus scattered) away from the beam direction. Extinction occurs when looking directly at a light source, while scattering is the diffuse radiation around the source.

2.3.1 Scattering mechanisms

There are two main scattering mechanism, depending on the size of the particle with respect to the wavelength of the incident radiation: Mie and Rayleigh scattering.

Rayleigh scattering. This is due to the interaction of light with particles smaller than the wavelength: if the particle is small enough the phase difference of the beam along the particle is negligible and the it “sees” in each moment an homogeneous electromagnetic field. The particle will be forced by the beam to emit as an oscillating dipole, in phase with the beam and with the same wavelength, but in a random direction. From the dipole theory, when a light beam of intensity I_0 travels through a medium, meeting N particles of polarizability α , the scattered intensity of the beam is

$$I = I_0 \frac{8\pi^4 N \alpha}{R^2} \frac{1 + \cos^2 \varphi}{\lambda^4} \quad (2.2)$$

where λ is the wavelength of the radiation and φ the angle between the beam and the observing direction. Note that it steadily decreases at increasing wavelengths ($\propto \lambda^{-4}$), i.e. Rayleigh scattering is stronger in the blue than in the red bands. Concerning the angular distribution of scattered light, the intensity is maximum both along the original beam direction (*forward scattering*) as well as in the opposite direction, back to the source (*back scattering*). Minima are located at 90° from the maxima.

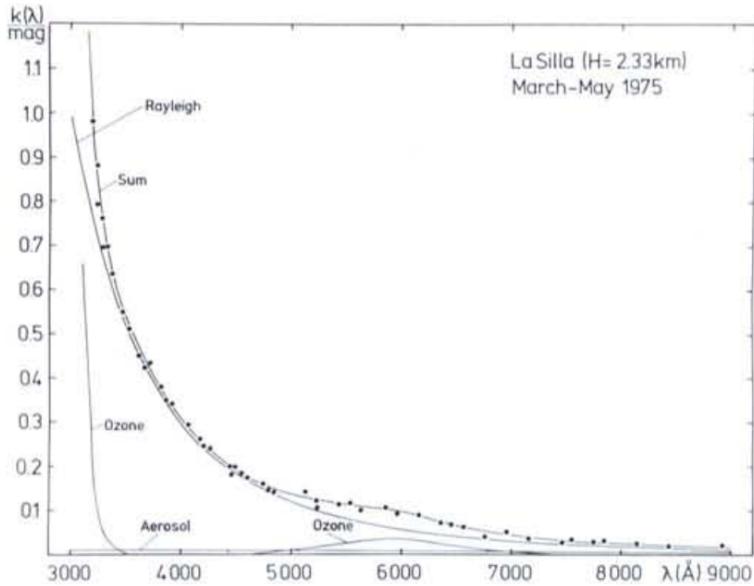


Figure 2.6: Extinction coefficients for Rayleigh and aerosol (Mie) scattering. The former component dominates over the latter at all the wavelengths, in particular for shorter ones. From [Tüg77].

In the optical regime, Rayleigh scattering is produced by the air molecules themselves. This form of scattering is responsible for the blue sky in daytime and for twilight. It is also the main responsible for atmospheric extinction, in particular at shorter wavelengths.

Mie scattering. Mie scattering occurs when light meets particles of size comparable or larger than the characteristic wavelength. In this case different parts of the particle experience different intensities of the electromagnetic field; each portion of the particle will emit as a single dipole, locally in phase with the beam, and the total emission will be the result of the interference pattern between all the portion of the particle. In this case only the wavelength of the original beam is preserved, while phase is distorted and direction is random. The numerical expression of Mie scattering is rather complex and involves the solution of Maxwell equation; several numerical or empirical relation are available in the literature. For our purposes it is relevant to say that the intensity of Mie scattering is proportional to $\propto \lambda^{-0.8}$. Moreover, most of the light is scattered along or close to the beam direction, i.e. the forward component is dominant (the larger the particle, the more relevant the forward component).

In the atmosphere Mie scattering is produced by solid particles suspended in the air (*aerosols*), such as dust grains, smoke or water drops. In the figure 2.6, from [Tüg77], are reported the extinction coefficients, a proxy the total scattered light, for Rayleigh and Mie scattering (the latter labelled as “aerosols”). It is clear that Mie scattering is negligible with respect to Rayleigh one, in particular at shorter wavelengths. This is true only in clear sky conditions, with an atmosphere locally without fog, smoke or other pollutants, that may significantly rise the impact of Mie scattering.

Note that to the Ozone extinction component in the figure 2.6 does not correspond any scattered component since it is produced by molecular absorption. In this case the radiation absorbed, instead of being re-emitted, is used to increase the kinetic energy of each ozone molecule, heating the ozone layer. This process is responsible for the temperature gradient inversion in the stratosphere.

2.3.2 Effects on the sky brightness

Every light source is responsible for a certain amount of light scattering. Depending on the site and the observation time, the effects of scattering on the total sky brightness can vary significantly.

Dark clear sites. In a dark site most of the scattered light is produced by airglow, zodiacal light and the galactic background, i.e. the brightest “direct” light sources. According to [LBH⁺98] scattered light from the natural sources accounts for a brightness of the order of 10^{-8} to 10^{-7} erg/s/cm²/sr/Å.

Light polluted sites. Atmospheric light scattering is the main mechanism responsible for light pollution when far from the light source. Due to Earth curvature, geographic features and atmospheric extinction the direct light contribute to pollution only when very close to the light source. As reported in the chapter XX, brightness of scattered artificial light from a city varies with the population and the distance.

Scattered sunlight and moonlight. As said before Rayleigh scattering is responsible for the twilight: even if the Sun is below the horizon some residual scattered light still brightens the sky. Conventionally astronomical twilight ends when the sun is 18° below the horizon and other light sources, such as zodiacal light, becomes more relevant.

A similar effect is provided by scattered moonlight. Such light brightness depends on the lunar phase and on the distance between the Moon and the observing position in the sky. According to [KS91] and similar sources, in optical bands scattered moonlight can lead to an increase of the sky brightness of the order of 5 mag/arcsec^2 .

Chapter 3

Artificial sky light pollution

Chapter 4

The spectrum of the sky

Chapter 5

Sky spectra reduction

In this chapter are reported the steps in the data reduction that I designed to extract spectra of the sky from frames originally taken for scientific purposes. After a brief overview I describe accurately the data reduction process that I built. To improve the readability of this writing, I will report only the most significant pieces of code.

5.1 Introduction

5.1.1 Software management and reduction steps

In this work I developed some pieces of Python (v. 3.10) code to manage all the steps of the data reduction. Note that most of the data pre-reduction was already done and was not necessary to use old software such as IRAF or its python version PyRAF. In this work I tried to heavily automatize the script in order to be able to analyze all the frames with a single run. Many efforts were spent to build a robust code, capable of working correctly for spectra with very different features, without the necessity to fine-tune the software settings every single time a new frame is processed. All of this effort was made in order to be eventually able, in the future, to rapidly analyze new frames.

I decided to divide the source code into different independent scripts, each one devoted to a specific task. It follows a brief description of each step of the data reduction process.

Background extraction Starting from the original spectra I have to separate the scientific targets from the background regions. Once identified the spectra of the targets and the cosmic rays, the relative regions are masked. The remaining area contains the spectrum from the sky background only and is extracted to a new file.

Background analysis The sky spectrum is averaged along the slit direction. From the regions that do not present lines is estimated the shape of the sky continuum emission. Prominent lines or lines of interest are identified and the relative equivalent width is computed. The output of background analysis is the estimation of the continuum emission and the list of the widths of the emission lines.

Line analysis The width of the same line is compared in the different frames. Particular attention is devoted to the line intensity with the epoch of observation and the direction in the sky.

Continuum analysis Continuum intensity in different bands of the spectra is computed and correlated again with the epoch and direction of observation.

5.1.2 The dataset

This work is based on 35 spectra taken between 2006 and 2020 in the Osservatorio Astrofisico di Asiago, Asiago, northern Italy. Spectra were collected by professor Stefano Ciroi and collaborators to collect data on studied astronomical objects and were taken with the 1.22 m reflective telescope “Galileo Galilei” equipped with the grating spectrograph “Boller&Chivens”.

Each frame has been pre-reduced by Ciroi and its work group: data has been corrected for bias and flat field and calibration on both flux and wavelength was performed. Cosmic rays were not removed as well as the sky background. Before November 2011 frames had a spatial scale on the CCD of 0.63 arcsec/px while on later data the scale was 1.0 arcsec/px. For all the object, grating with a line density of 300 tr/mm

was used while the grating angle varied between 0° and 5.25° according to the type of target. Similarly slit aperture size varied from a minimum of 200 to a maximum of 400 μm while the exposure times ranged between 300 and 3600 s.

5.2 Background extraction

The data available was not taken with the aim of monitoring the sky condition and thus contain astronomical objects. The first step in the data reduction is to identify the regions of the spectra where there are the scientific targets. Most of the effort described below is to automate the process.

On the figure 5.1 is shown as an example the spectrum obtained in the 6 November 2020 pointing at the active galaxy labeled 4FGL J0512. Two main astronomical sources are visible as long strips in the top and bottom side of the frame. The other bright vertical lines and bands represent the emission of the sky diffuse background. Looking carefully some cosmic rays can be also spotted toward bluer wavelengths. This frame will be used later to show the different steps in the data process.

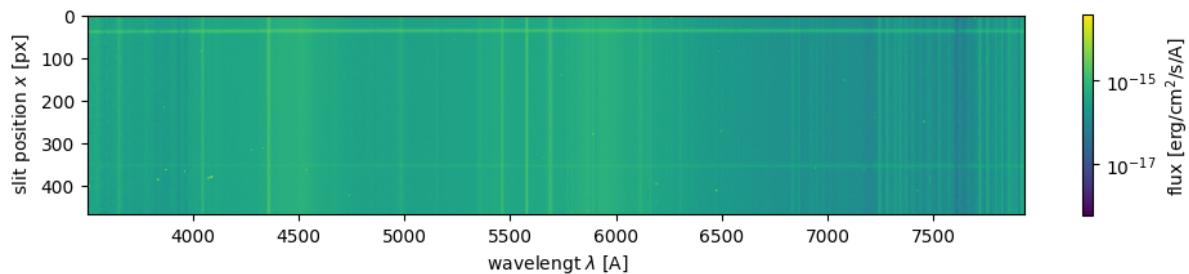


Figure 5.1: An example spectrum from the dataset. The frame contains the active galaxy “4FGL J0152”. Main components are the galaxy and another source (horizontal stripes), the diffuse background (vertical lines and bands) and some cosmic rays (especially below 4000 \AA).

5.2.1 Preamble

Before the actual data analysis I need to load all the necessary.

Modules, options and parameters. First I need to load packages and set parameters to allow the rest of the script to work correctly.

```
import numpy as np
import matplotlib.pyplot as plt
from astropy.io import fits
from scipy.signal import find_peaks, peak_widths
from datetime import datetime
import glob
import os
from wotan import flatten
from scipy.optimize import curve_fit

#####
#OPTIONS
save_plots = True
save_FITS = False
plot_profile = True
plot_spec = True
show_ima = False

#
#PARAMS
peak_height = 0.05 #height above the bkg level
data_col_frac = .75 #minimum fraction of valid pixels in a column
width_mult = 2 # interval to exclude around a source, wrt center in FWHM units
cr_width = 2.5 # cr trace spatial width
cr_prominence = 5 #threshold height wrt average column level
cr_pad = 1 # number of px to exclude around cr, in a fixed column
LAMBDA_lim = 3500 #Å, limit blue wavelength
#####
```

Options are some special parameters that allow to enable or disable some debugging diagnostics useful in the development phase. Parameters are some special values that can be changed to control the final output result. Each of these lines will be explained later when appearing in the code.

Data import and handling. Each spectrum is stored in a `.fits` file (acronym of Flexible Image Transport System, see [biblio FITS format](#)). For this work I need both the actual spectrum data and the auxiliary information about the acquisition and reduction processes, contained in the header section.

To automate the process of data import I used the `glob` function from the homonyms module ([citare glob package](#)) to identify all the files with the FITS extension in a given path directory. I also needed the function `basename` from the `os.path` module ([citare os.path](#)) to retrieve the actual name of each file, since `glob.glob` provides only the full path of the files. Then the file research is implemented with the following code.

```
#browse all the *.fc.fits files in a directory and its subdirectories
main_path = './Asiago_nightsky/'
file_ls = glob.glob(main_path+'/**/*.*.fc.fits', recursive= True)
names = [os.path.basename(x) for x in file_ls]
```

`main_path` can be eventually adapted to scan new files in different directories. The choice to search for files that ends with `.fc.fits` is due to the fact that Ciroi and collaborators use the convention to append the suffixes `.f` and `.c` to indicate frames corrected for flat field and calibrated respectively.

5.2.2 Data extraction

For each file found in the `main_path` directory I need to extract the information contained both in the header and the data unit(s). This task is performed by the `astropy.fits.io` module and is implemented in the following way:

```
#process all the files found
for name,file in zip(names,file_ls):

    #open a FITS file
    hdul = fits.open(file)
    hdr = hdul[0].header
```

The header is a Python dictionary from which I can extract the quantities that I will need as new variables.

Wavelength data. The first thing to extract is the wavelength information in the following way.

```
#extract wavelenght information from the header
NAXIS1, NAXIS2 = hdr['NAXIS1'], hdr['NAXIS2']
LAMBDAO, DELTA = hdr['CRVAL1'], hdr['CDELT1']

#generate the lambdas array
if hdr['CTYPE1'] != 'LINEAR':
    print('WARNING: no linear wavelength calibration')
    LAMBDA = np.arange(LAMBDAO, LAMBDAO+NAXIS1*DELTA, DELTA)
    if len(LAMBDA) == NAXIS1+1:
        LAMBDA = LAMBDA[:-1]

#remove extreme blue wavelengths
LAMBDA_start_id = len(LAMBDA)-len(LAMBDA[LAMBDA>LAMBDA_lim])
LAMBDA = LAMBDA[LAMBDA_start_id:]
```

In this way the array `LAMBDA` contains the wavelength associated to each column of the spectrum, assuming a linear calibration. In the case the original calibration is not linear, a warning message is printed.

I decided to limit the wavelengths in the blue hand of the spectrum to the limit value `LAMBDA_lim` which was set to 3500 Å. Below this threshold the signal tends to be too noisy and low quality, due to the strong correction in the flux calibration phase and the intrinsic low sensitivity of CCDs in the blue domain. Also `LAMBDA_lim` is meant to be a parameter that can eventually be changed to fit new data.

In the figure 5.2 is show the bluest hand of the spectrum of the planetary nebula NGC2392 (also known as “The Eskimo nebula”), taken on the 1 February 2006: the signal is extremely noisy due to the extreme calibration necessary at those wavelengths. Note also the presence of several cosmic rays (red circles) that provide further spurious signals to the spectrum.

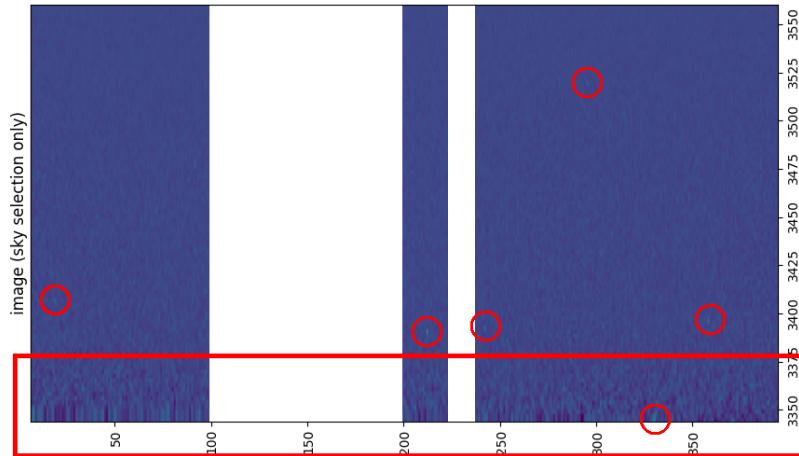


Figure 5.2: Extreme blue/UV domain in the spectrum of the nebula NGC2392. The area in the rectangle is totally useless due to the noise. Cosmic rays have been highlighted too.

Slit and scale data. For my script I will also need information about the angular scale of the spectra. In particular I want to retrieve the angular scale on the final spectrum (in arcsec/px) and the size of the slit on the detector (in px). Implementation is the following:

```
year = hdr['DATE-OBS'][:4]

#aperture information from the hdr
SLIT = hdr['SLIT'] #microns
try:
    BINX, BINY = hdr['BINX'], hdr['BINY'] #binning factors
    TELSCALE = hdr['TELSCALE'] #arcsec/mm
    CCDSCALE = hdr['CCDSCALE'] #arcsec/px
except KeyError:
    BINX, BINY = hdr['HBIN'], hdr['VBIN']

print(' WARNING: no scale info in the hdr (using defaults)')

TELSCALE = 10.70 #arcsec/mm #TO BE CHECKED!!!
CCDSCALE = 0.63 #arcsec/px #TO BE CHECKED!!!

SLIT_angular = SLIT/1000 * TELSCALE #slit size in arcsec
SLIT_px = SLIT_angular / CCDSCALE / BINX #slit size in px
```

Note the TELSCALE and CCDSCALE entries in the file header are available only from 2009. For earlier files I had to set by hand these two variables; in this case a warning message is also displayed. As reported in the code, as stated in [citare proprietà CCD prima del 2009](#) before 2009 the telescope scale was of 10.70 arcsec/mm while the scale on the detector was 0.63 arcsec/px.

5.2.3 Cosmic ray and noise removal

The next step in the sky extraction process is to remove the cosmic rays and the residual noise in the blue domain. These two sources of photons provide spurious signals that may interfere with the detection and removal of the astronomical signal from the background. See for example the figure 5.2 at page 24.

Implementing an automatic process of recognition and masking of the spurious signals from cosmic rays and photon noise is rather complex and requires quite advanced analysis tools. The steps I made are the following:

1. I scan each column of the spectrum, i.e. for each fixed value of the wavelength. The average flux level of the column is computed as a reference of the brightness of the spectrum at the analyzed wavelength.
2. In each column I look for sharp peaks in the flux due to the presence of cosmic rays or strong noise. The peaks are identified with the function `find_peaks` from the module `scipy.signal`. In particular I restricted only to peaks with a width below the value of `cr_width= 2.5 px` and a

prominence¹ higher than `cr_prominence= 5` times the average value of the column. These are good settings to filter only peaks of noise and cosmic rays that are typically very sharp and much brighter than the astronomical signal.

3. I mask the pixels around the peaks. The number of pixels to mask is computed from the width of the peaks, which is provided by the function `peak_widths` from the module `scipy.signal`. I used this function to compute the FWHM of each peak, since this was a more solid width estimator with respect to the full width, which may be biased due to the noise fluctuations around the peak. I assumed that the total width of the peak was $2 \times \text{FWHM}$. To be sure all the cosmic rays traces or the noise fluctuations were contained in the detected width, I conservatively decided to increase the masked interval of a further `cr_pad= 1 px` in both directions along the columns.
4. I remove all the columns where too many pixels have been masked, i.e. too noisy columns. The minimum fraction of saved pixels in a column must be higher than the value `data_col_frac= 0.75`.

After the removal of sharp peaks and noisy columns I expect to have a clean frame which contains only the spectra of the astronomical targets and the background sky. The procedure described above is implemented by the following lines.

```
#####
#bkg level estiamtion
raw_data = hdul[0].data[:,LAMBDA_start_id:]
raw_integr = np.sum(raw_data, axis = 1)
bkg_est = np.nanmedian(raw_integr)

#####
#remove cosmic rays and UV noise
x = np.arange(len(raw_integr))
data = np.copy(raw_data)

cr_col_frac = np.zeros(len(LAMBDA)) #fraction of remaining px
for cr_col,col in enumerate(data.T):
    col_avg = np.nanmean(data[:,cr_col])
    cr_line,_ = find_peaks(col,
                           prominence = cr_prominence*col_avg,
                           width = (0,cr_width))

    cr_widths = peak_widths(col, cr_line, rel_height=0.5)[0]

    #set left and right boundaries of the source region along the slit
    left_width = cr_line-cr_widths - cr_pad
    right_width = cr_line+cr_widths + cr_pad

    #scan each column and remove peaks
    cr_sel = np.zeros(np.shape(col), dtype=bool)
    for i in range(np.shape(col)[0]):
        for peak,width in zip(cr_line,cr_widths):
            if abs(i-peak) < width+cr_pad:
                cr_sel[i] = True

    #counts how many pixels are left in a column
    saved_px = (NAXIS2 - np.sum(cr_sel))/NAXIS2
    cr_col_frac[cr_col] = saved_px
    if saved_px >= data_col_frac: #if enough, take the masked column
        data[cr_sel, cr_col] = np.nan
    else: #else discard the entire column
        data[:, cr_col] = 0.
```

5.2.4 Sources identification

The main idea to find scientific target is to integrate the (cleaned) spectrum in the wavelenghts and obtain the integrated luminosity profile of telescope field through the slit. Astromical sources are spatially limited and thus appear as bright peaks over the flat homogeneus background. Unluckyly in some real spectra the continuum is not homogeneus but present some gradients, probably due to some calibration biases; consequently the background integrated luminosity profile may not be flat at all, complicating the

¹info prominence significato

detection of astronomical signals. If $F(x)$ is the total luminosity profile, we can imagine to split the two components

$$F(x) = A(x) + B(x) \quad (5.1)$$

where A is the signal due to the astronomical sources and B the one from the background.

The full implementation of the identification process follows the steps below:

1. Integrate the spectrum along the wavelengths to retrieve the luminosity profile along the slit. Estimate the typical luminosity value of each profile as the median one; median value is much less sensitive to peaks in the luminosity profile and for this reason is preferable with respect to other estimations such as the average.

If $F(x, \lambda)$ is the flux in a pixel of the spectrum in the position x along the slit and wavelength λ , then the luminosity profile $F(x)$ is obtained as

$$F(x) = \int_{\lambda_{min}}^{\lambda_{max}} F(x, \lambda') d\lambda' \quad (5.2)$$

while typical luminosity value is

$$\hat{F} = \underset{x}{\text{median}}\{F(x)\} \quad (5.3)$$

2. Estimate the general trend of the background B , i.e. regardless the presence of peaks of the astronomical targets. I used the function `flatten` from the `wotam` package ([citare wotam](#)) to perform a bi-weighted detrending of the signal. This technique consist on estimating the smoothed value of the signal in the position x as the median value of the data centered in x and a window of total span w . A more robust indicator is obtained when data are filtered such that points with a larger displacement to the median value are less relevant. Given the value of the residual a of the data with respect to the median value, there can be several different types of relative weight functions; in the `wotam` package it is used the Tukey function which is expressed as

$$L(a) = \begin{cases} \left[-\left(\frac{a}{c} \right)^2 \right]^2 & \text{if } |a| < c \\ 0 & \text{else} \end{cases} \quad (5.4)$$

where c is a tuning parameter that specifies the steepness of the weight scale.

I used a windowing of 50% the total profile length and a fine-tuning parameter `cval=c=10`. Such large window can only partially smooth down the peaks signals. Called $\mathcal{S}_{w,c}$ the biweighted filter operator, with window w and tuning parameter c then the smoothed luminosity profile is given by

$$F'(x) = \mathcal{S}_{w,c}\{F(x)\} \quad (5.5)$$

3. A better approximation for the background B can be obtained by masking the data around the astronomical sources. I used as threshold the detrended profile F' , vertically shifted by a factor equal to the 5% of the average bkg level \hat{F} . The new profile obtained is

$$F''(x) = F'(x) \cdot \theta(x) \quad \text{where} \quad \theta(x) = \begin{cases} 1 & \text{if } F(x) \leq F'(x) + 0.05 \cdot \hat{F} \\ \text{nan} & \text{else} \end{cases} \quad (5.6)$$

The new masked profile F'' is much closer to B than the original profile as the contribution of astronomical sources A has been drastically decreased.

4. The final background profile estimation is obtained by a further filtering of the masked trend F'' . I used the `flatten` function with a window of 10% the profile length. This allows to smooth the noise fluctuations in the background. Formally

$$B(x) \approx \mathcal{S}_{W,c}\{F''(x)\} \quad (5.7)$$

where $W \ll w$ is the new windowing factor, while the value c has been kept constant. Once estimated B I can focus on the profile of the astronomical sources A which can be isolated simply as $A = F - B$.

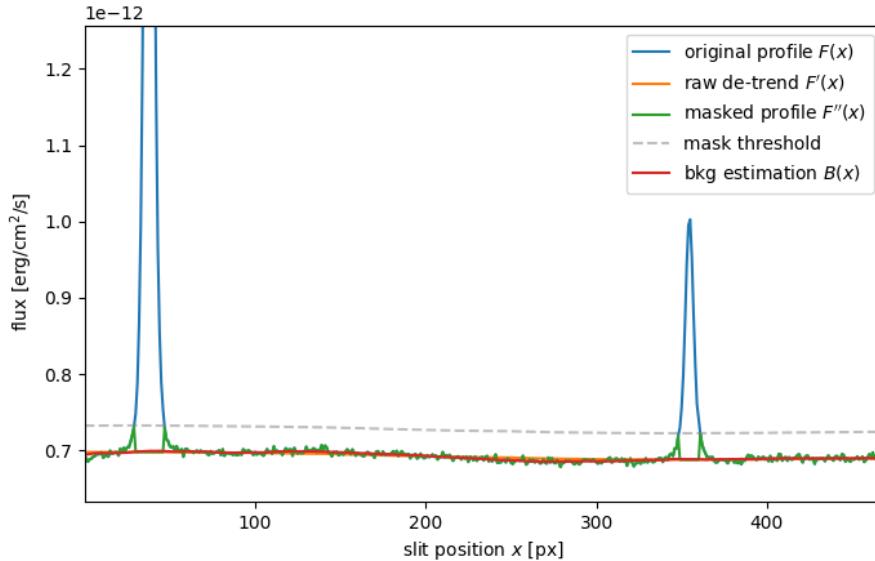


Figure 5.3: The different luminosity profiles to estimate the background B on my test frame. Thanks to the good quality of the data (uniform slit illumination, high signal-to-noise ratio) the identification of the sources is particularly easy.

In the figure 5.3 are plotted the various luminosity profiles used to retrieve the final background B in my example frame (see above). In this case the background is quite constant along the slit and the signal from the sources stands out the background easily; even the preliminary filtered profile F' would have been a good approximation of the background, without the necessity of further computation, which is instead necessary for frames with lower quality data.

5. Identify the peaks in the profile of the astronomical sources A with the function `find_peaks` (see above). I selected peaks higher than 5% of the reference level \hat{F} and broader than `cr_width`; the first condition ensure not to include faint peaks due to background noise fluctuations, the second one, prevent from including sharp peaks produced by cosmic rays.
6. Estimate the width of the peaks with `peak_widths` that provided the FWHM. Many astronomical sources such as galaxies or planetary nebulae present a bright core and faint extended wings. This means that the FWHM can underestimate the total width of a peak if wings are very broad. In a conservative approach I decided to assume the peak width as $2 \times \text{width_mult} \times \text{FWHM}$ where `width_mult = 2` is another tunable parameter of the script. With the resulting widths I masked the regions centered around the peaks.

In the figure 5.4 are shown the areas associated with the astronomical sources that has been removed, in my example frame (red bands) and the final background signal (green dots).

The masked intervals in the luminosity profiles becomes masked rows, when back to the bidimensional spectrum, which should now have only counts from the sky background.

On the figure 5.5 is finally reported the background sky spectrum extracted from the example frame. The white areas have been masked due to the presence of astronomical sources or cosmic rays.

The full code to implement source extraction is the following.

```
#use noise/bkg info to find peaks
integr = np.nansum(data, axis = 1)
bkg_est = np.median(integr)

#detrend: global trend (including peaks)
_,trend_raw = flatten (x,
                      integr ,
                      method ='biweight',
                      window_length =0.5*NAXIS2 ,
                      cval = 10, return_trend = True )

#trim removing peaks, i.e. data fare above the global trend
```

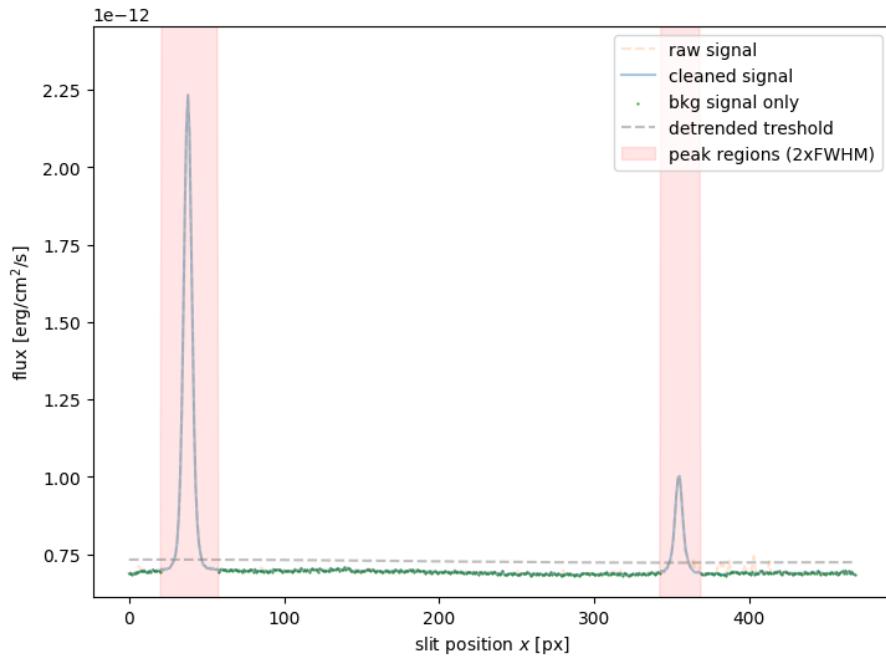


Figure 5.4: The astronomical sources identified by the script in the test frame. The original profile (broken orange line) has been filtered to remove cosmic rays and noise, contributing to smooth out the profile into the cleaned one (blue solid line). Red areas on the cleaned profile are the regions occupied by the sources and are going to be masked. The remaining usable signal from the background is represented by the green dots, each one representing a row in the bidimensional spectrum.

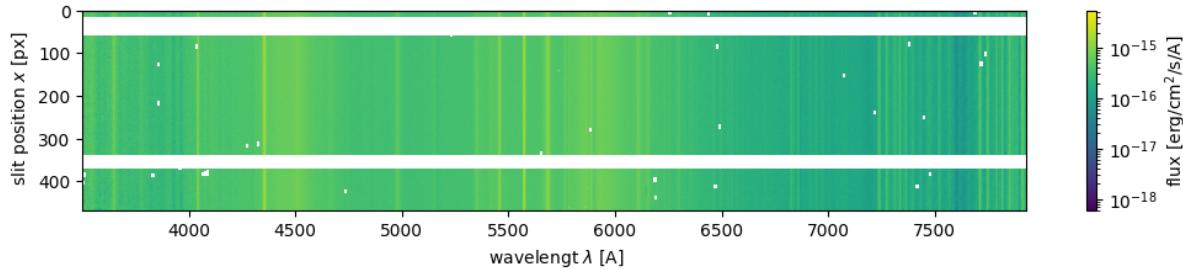


Figure 5.5: Final background spectrum extracted from the example frame.

```

integr_trim = np.where(integr <= trend_raw+0.05*bkg_est,
                      integr, trend_raw)

#detrend the trimmed data, much less sensitive to the peaks
_,trend = flatten(x,
                   integr_trim ,
                   method ='biweight',
                   window_length =NAXIS/10. ,
                   cval = 10, return_trend = True )

#detrend residuals: original peaks are highlighted wrt the bkg profile
diff = integr-trend

#find peaks
peaks,properties = find_peaks(diff, height=0.05*bkg_est, width = cr_width)
peak_FWHM = peak_widths(integr, peaks, rel_height=.5)[0]/2.

if len(peak_FWHM)== 0:
    print( " WARNING: no sources were detected")

```

```

#generate a boolean mask True outside the peaks
bkg_sel = np.full(np.shape(x), True)
for i,peak in enumerate(peaks):
    width = (int(peak_FWHM[i])+1)*width_mult
    for w in range(-width,width):
        bkg_sel[peak+w]=False
    w = width -1
    while integr[peak+w] >= trend[peak+w]:
        bkg_sel[peak+w] = False
        w += 1
    w = width
    while integr[peak-w] >=trend[peak-w]:
        bkg_sel[peak-w]=False
        w += 1

#####
#integrated spectrum (along the slit)
total = np.nanmean(data, axis = 0) #integration along the slit
sky = np.nanmean(data[bkg_sel,:], axis = 0) #integration of bkg rows only

total[total == 0] = np.nan

#####
#extract only the bkg rows

ma_data = data #set masked data
for i,row in enumerate(bkg_sel):
    #cancel data from the source rows
    if row == 0:
        ma_data[i,:] = np.nan

```

5.2.5 Sky spectrum export

Once masked all the contaminants, I can extract the sky spectrum and save it for further analysis. For each frame I create a new .fits file that contain the sky spectrum and the masked data from other sources. The header of the new file contain the same information of the original one, plus further additional information about the time of creation and the value of the limited wavelength LAMBDA_lim (respectively “BKGEXTR” and “UVLIM”). Each file is saved in the firectory of the original one and to its name is appended the .bkg. suffix.

```

#####
#save masked data in a new FITS file
if 1 == save_FITS:
    now = datetime.now()
    now_str = now.strftime("%Y-%m-%d %H:%M:%S")

    hdr.set('BKGEXTR', now_str, 'Time of bkg extraction')
    hdr.set('UVLIM', LAMBDA_lim, 'A')
    hdr['NAXIS1']=len(data[0])
    new_hdu = fits.PrimaryHDU(ma_data)
    new_hdul = fits.HDUList([new_hdu])
    new_hdul[0].header = hdr

    file_new = file[:-5]+'.bkg.fits'
    new_hdul.writeto(file_new, overwrite=True)

```

5.3 Background analysis

Bibliography

- [Eat69] RH Eather. Latitudinal distribution of auroral and airglow emissions: The ‘soft’auroral zone. *Journal of Geophysical Research*, 74(1):153–158, 1969.
- [FHLT74] A Frey, W Hofmann, D Lemke, and C Thum. Photometry of the zodiacal light with the balloon-borne telescope thisbe. *Astronomy and Astrophysics*, 36:447–454, 1974.
- [KS91] Kevin Krisciunas and Bradley E Schaefer. A model of the brightness of moonlight. *Publications of the Astronomical Society of the Pacific*, 103(667):1033, 1991.
- [LBH⁺98] Ch Leinert, S Bowyer, LK Haikala, MS Hanner, MG Hauser, A-Ch Levasseur-Regourd, I Mann, K Mattila, WT Reach, W Schlosser, et al. The 1997 reference of diffuse night sky brightness. *Astronomy and Astrophysics Supplement Series*, 127(1):1–99, 1998.
- [Lei75] Christoph Leinert. Zodiacal light—a measure of the interplanetary environment. *Space Science Reviews*, 18(3):281–339, 1975.
- [Tüg77] H Tüg. Vertical extinction on la silla. *ESO Messenger*, 11:7–8, 1977.