Data Processing Chain for Heterodyne Spectrometer

Matthias Gehnen, Moritz Hefner, Bas Kruisselbrink, Johannes Plett, Felix Rauh

&

Jörn Ungermann, Martin Kaufmann, Benjamin Berkels

Short Description
The Institute of Energy and Climate Research: Stratosphere (IEK-7) launched an initiative to study the Earth atmosphere by light detecting instruments on satellites. The detected images are treated with image processing algorithms presented in this report. The observations of the Spatial Heterodyne Interferometer (SHI) are simulated and moreover calibrated and real orbit measurements are analyzed.

1 Introduction

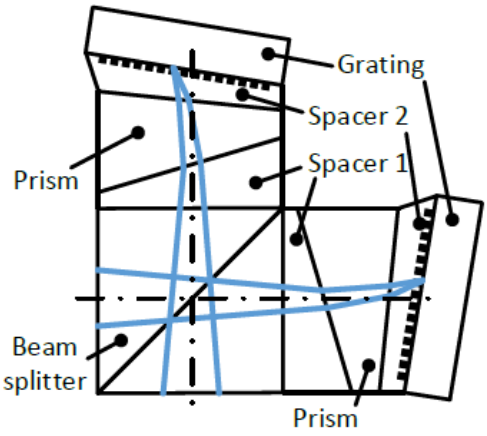
For a better understanding of atmospheric coupling processes it is important to have reliable information about the mesosphere and the lower thermosphere. So far, these parts of the earth's atmosphere located between the stratosphere and the exosphere have been relatively neglected in the research, such that even small research efforts may produce new insights. The Institute of Energy and Climate Research: Stratosphere (IEK-7) is working on remote sensing instrumentation based on Spatial Heterodyne Spectroscopy (SHS) to study the Earth atmosphere by nano-satellites, so called CubeSats. Contributors are Research Center Jülich with University of Wuppertal, the Max-Planck-Institute for the science of light with University of Erlangen, the national German metrology organization and some universities in Canada.
The SHS allows to measure oxygen emissions in the mesosphere and the lower thermosphere, which can be used for deriving temperature profiles. Thus, gravity waves can be detected. Including these in global circulation models is useful for the goal of improving climate predictions and weather forecasts since the extension to higher altitudes makes the respective models more complete.

There are clear advantages of the SHS interferometer configuration which make it particularly favorable for deployment on nano-satellites. First, it can be configurated to have good resolution for the detection of the wavelengths of interest, but at the same time is small in size. Secondly, there are no moving parts built into the entire instrument, in contrast to, e.g., Michelson Interferometers. Therefore, it is more robust during the start of rockets. More information can be found in [1].

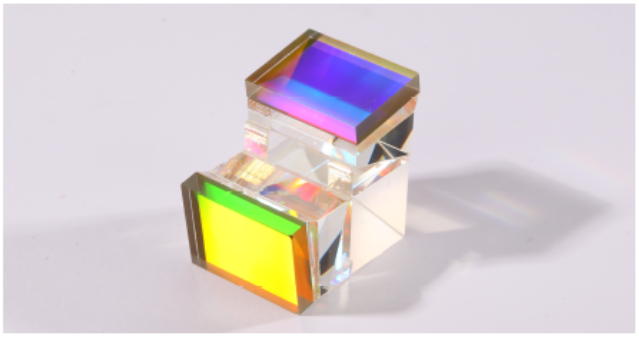
Since the end of 2018, a version of the SHS is deployed on a Chinese communication satellite, which already sent some data which are also considered in Section 5. In the next section, we will introduce the model of the SHS which we implemented and worked with. Then, we describe the different kinds of data we worked with. In Section 4, the developed algorithms are introduced and eventually, in Section 5, the results are summarized.

2 The Spatial Heterodyne Spectrometer

Figure 1 (taken from [1]) shows a schematic of the SHS. A beam of light enters, is splitted, diffracted and eventually creates an interference pattern on the detector. We describe the detector by a matrix where each entry corresponds to a pixel. The value equals the signal intensity of the incoming radiation at the respective pixel. Each row of that matrix is considered separately since each corresponds to a different altitude in the atmospheric layers observed by the SHS. For a fixed row, let x denote the horizontal position and $I_g(x)$ the measured intensity.



(a) Schematic of the SHS.



(b) Assembled SHS for the precursor mission.

Figure 1: The SHS consists of two diffraction gratings, a beamsplitter, and the field-widening prisms. The relative position between these parts is controlled by spacers (Spacer 1 and Spacer 2). The blue lines in show the ray paths for the central beam, focusing on the gratings.

We want to deduce the spectrum of the incoming radiation by analyzing this data. Each contained frequency in the spectrum is described by a wavenumber σ and the intensity $L(\sigma)$. As described in [1], the relation to the measured intensity in the point x is

$$I_g(x) = \int_0^\infty 0.5 \cdot L(\sigma) \cdot (1 + \cos(8\pi \tan(\theta_L)\ x\ (\sigma - \sigma_L)))d\sigma.$$

Here, so-called the Littrow angle θ_L describes how the gratings are bent. The Littrow wavenumber σ_L is also a configuration-dependent constant and determines which wave number can be seen in the center of the detector. To get the wavenumber and the intensity from the data we use a fourier transformation which inverts the discretization of the above formula, consisting of the addends

$$I_g(x, \sigma) = 0.5 L(\sigma)(1 + \cos(8\pi \tan\theta_L x(\sigma - \sigma_L))). \tag{1}$$

Each detected point by the interferogram of the SHI is generated by in inverse Fourier transformation from the spectral to the spatial space. The detected image can be transformed back to a discretization of spectral wavelengths by a discrete Fourier transformation (DFT). For an image consisting of rows of N pixel with measured intensities $I_n, n = 0, \ldots, N-1$, the DFT calculates for $l = 0, \ldots, N-1$ the values

$$I'_l = \sum_{n=0}^{N-1} I_n \cdot \exp(-i2\pi ln/N),$$

where I'_l denotes the intensity of light of some specific wavenumber σ_l. By a Discrete Fourier Transformation frequency analysis the spatial frequencies ν_l of the detected interferogram can be derived. Together with the equation

$$\nu_l = 4 \tan \theta_L(\sigma_l - \sigma_L)$$

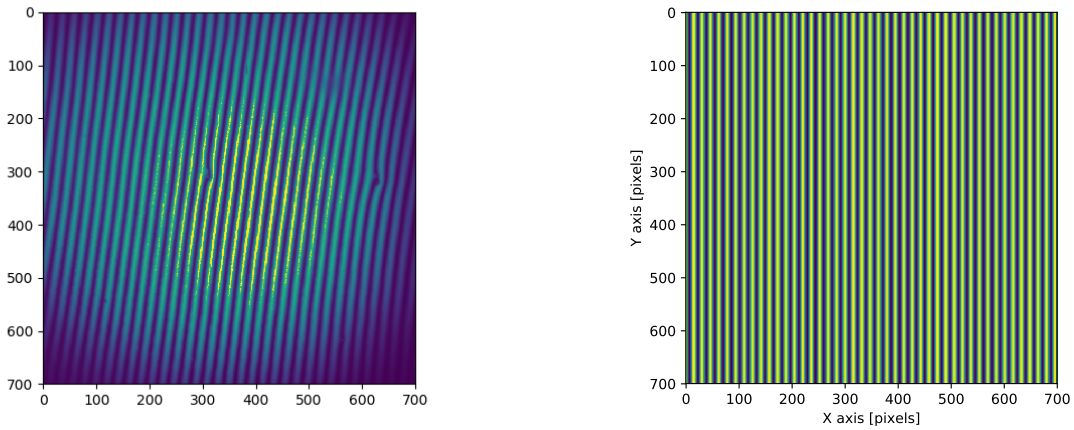
this provides the corresponding wavenumber σ_l to the intensity I'_l [1](4.4.3).
This transformation of the results of the DFT to the sequence of wavenumbers $\sigma_1, \ldots, \sigma_N$ with our given parameters of the SHI does not deliver the input wavenumbers, which we observed after applying our analysis. For an image size of 700 pixel a correction term of 2cm^{-1} may be added to the Littrow wavenumber σ_L to obtain fitting results, but with other imagesizes, this correction term does not work properly.

3 Used Data Sets

The data used can be split into three different types. The first one is the real orbit data, which is the relevant data that we are eventually interested in. The second one is laser data which was generated in a test stand on earth. This means that the incoming radiation (ideally) consists of a single wave length. Considering that the interference pattern of this radiation is a single trigonometric function, these images are very useful to detect sources of errors in the data. In the following, we will briefly introduce the errors which appear in the laser images. Afterwards, in Section 3.2, we will introduce methods for creating synthetic detector data which contain similar errors.

3.1 Errors in the Sample Data

Consider Figure 2 which contains a few properties which we will consider in further detail in Section 4 where we describe how to reduce the described effects.



(a) Interference pattern generated by 763.109nm laser

(b) Ideal interference pattern

Figure 2: Mitigating the saturation error

There are two obvious discrepancies between the expectation and the reality. The minima and maxima of the interference pattern are shifted to the right. This can be explained by a rotation of the gratings or the detector. Furthermore the intensity decreases towards the edge, which is a known lense effect (vignetting).
Moreover, at some central pixels, there is the problem, that the detector's sensitivity is exceeded. Our goal is to eliminate or reduce these effects in order to get better data. As a first step, we reproduce these errors in synthetic data.

3.2 Generating Synthetic Data

The given measurements of the SHS of monochromatic light still include some errors, due to optical reasons like the behaviour of a lense or because of technical failures. In order to test our later developed algorithms and to get an insight on how they behave, we generated synthetic images of what the SHS should measure. The aim was also to simulate some of the observed effects on the detected images. More specifically, we considered the effect of a tilting by a certain angle φ and a flattening of the detected light intensities towards the edges of the detector.
Without the errors, the incoming intensity of light with a specific wavenumber σ and an amplitude of $L(\sigma)$ is described by (1). In particular, the light intensity on the detector is independent from the vertical distance of the center. A tilting of the image can be achieved by adding a phase shift term in the cosine function. To generate a tilting by an angle φ we adjust the formula (1) to

$$I'_g(x, y, \sigma) = 0.5L(\sigma)(1 + \cos(8\pi \tan\theta_L x(\sigma - \sigma_L) + 2\pi\varphi y)), \tag{2}$$

where y is the vertical distance from the detectorcenter. To also include a flattening of the measured light intensities towards the edges of the detector, we factor the function I_g by a parabolic function $f(x, y)$. For an image with a maximal distance of Δ between the pixel and the center c_0 of the image, f is described by

$$f(x, y) = 1 - \frac{\|(x, y) - c_0\|}{\Delta},$$

so the value of f is 1 in the center of the image and 0 on the four corners. Then we modify (2) by adding f as a factor, so our final function

$$I_g(x, y, \sigma) = f(x, y)0.5L(\sigma)(1 + \cos(8\pi \tan\theta_L x(\sigma - \sigma_L) + 2\pi\varphi y)) \tag{3}$$

describes a synthetic detector image of a specific wavenumber σ, if the function values are generated for some grid values for x and y.

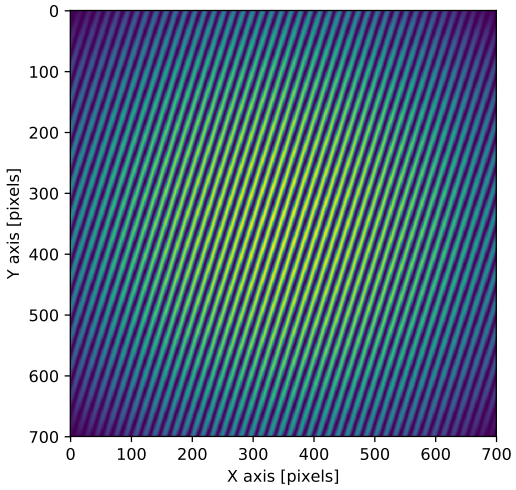


Figure 3: Synthetic tilted and flattened detector image with wavelength 763.109nm

Moreover this method can be used to generate an interference pattern for a set of wavenumbers $(\sigma_1, \ldots, \sigma_n)$ with corresponding intensities $(L(\sigma_1), \ldots, L(\sigma_n))$ by calculating the values $I_g(x, y) = \sum_{i=1}^{n} I_g(x, y, \sigma_i)$ for grid values of x and y.

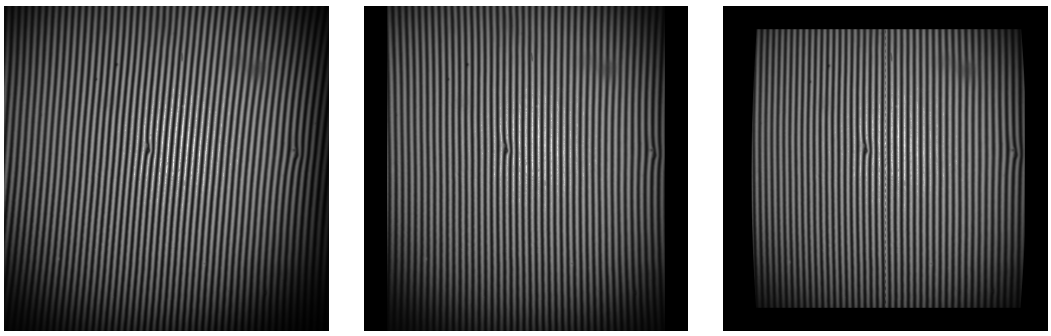
4 Image corrections

In this section, we explain how we try to remove the errors described in Section 3.1 from the data. We will show our algorithms starting with an grayscale sample image from the test with lasers (Figure 4a).

4.1 Tilting

As we see in the interference pattern of the heterodyne spectrometer, the lines indicating the maxima and minima are not vertical. To correct this error, we first have to identify how big the shift is. Second, we may modify the interferogram by shifting pixel such that the interference pattern becomes vertical.

For finding out how far the interferogram is shifted, we first choose a set of pixel at the lower and at the upper side of the detector. Then we consider the lines between the pixel from the lower to the upper border: The line with the biggest sum of measured intensities will probably be a maximum



(a) Testimage by a laser (b) Tilted image (c) Tilt and pincushion corrected

of the wave. By taking this line, we see how much the interference is shifted away from a vertical line.

Second, we can modify our interferogram by moving every pixel on the x-axis, so the line becomes vertical. Then our image looks like Figure 4b, after tilting it by up to 60 pixel.
This is the first step we take, to get some structure in our data. More important and a bit more difficult is the next step:

4.2 Pincushion

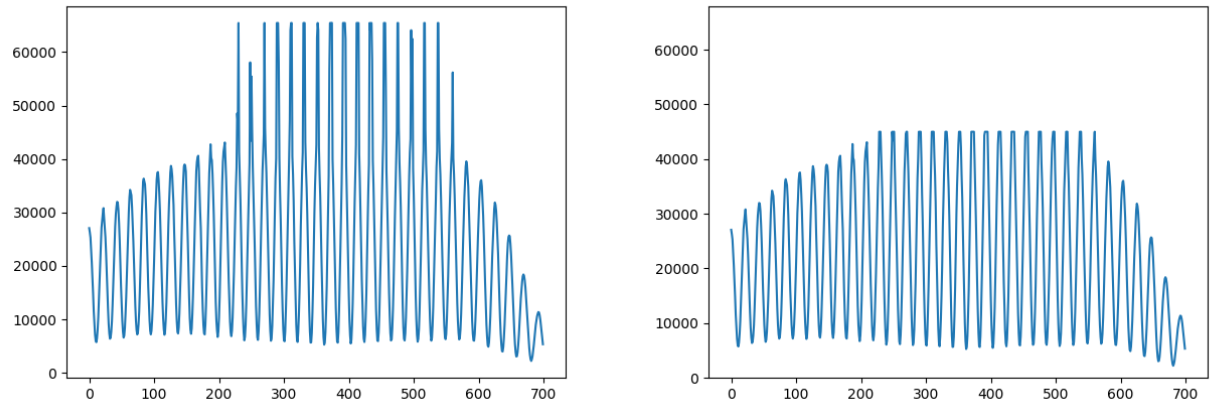
The more complex error is called pincushion, and it describes a kind of distortion. Here, lines that do not go through the middle of the image (here: through the point $350, 350$), because all our images have the size $700 * 700$) are bent towards the middle of the image. The more outside a line is, the more it is bent.
In Figure 4b the effect is easily visible. The error exists not only in the horizontal axis, but it also exists at the bottom and at the top. In our pictures, this is not easy to see. But it should be mentioned, that of course we must correct the mistake here as well as the horizontal mistake.
If we want to correct this error, we have to move the bent lines back to straight lines. For doing this, first we have to detect in which way the lines are bent. Because pincushion is a typical optical effect which is caused by the same reasons as in other situations, it is well known that the lines are bent quadratically.
So, we just have to find out the coefficient of the parabola. This can easily be done in the same way in which we have identified the coefficients of the tilting: We search for the parabola which goes through the pixel with the highest sum of intensities. With this parabola, we simply can read out the coefficients for the pincushion. Now, for every pixel we can calculate where the correct data of this pixel is saved in the picture. We move the pixel to the correct position and get as a result in our example the image 4c .
The calculation will be presented in detail in an extended report.

4.3 Vignetting and Sensitivity Function

In optics and photography the vignetting of an image is the darkening or shadowing on the edges and corners as already seen in Figure 2a. This effect is caused by the lenses of the optics in the heterodyne spectrometer. Taking the measured monochromatic interference pattern into account, the loss of intensity at the edges and corners of the data causes other wavelengths to appear on the spectrogram. If we fourier transform the monochromatic laser data there should only be one wavelength. To ensure that the data collected in orbit is not measuring any wavelengths that got created by the vignetting, we have to counter the effect. We do this by modeling the vignetting as a curve which minimizes the data in least-square sense. We do this for every row each in 1D, because the orbit interference pattern has the most important information on small fraction of the height in the image. But before we can start fitting our curve, we have to resolve another problem.

4.3.1 Saturation of the pixel

In the monochromatic laser data collected on earth a saturation of the pixel in the middle of the image can be found. This is caused by exceeding the limits of the capacitors in the detectors. This caused the detector pixel to misinterpret the intensity of the incoming photons. Looking at the collected data the saturation manifests as a plateau with a constant intensity as seen in figure 5. There are two reasons to cut these peaks off. First of all this is corrupted data that gives only false information and therefore cannot be used. Secondly, if we fit a curve to counter the vignetting with these peaks caused by the saturation, the curve deforms and increases the error after the correction of the real data. The peaks get cut off to the same level of the intensities next to the corrupted data. This way, the error caused by the saturation is only a fraction compared to before.



(a) Saturation corrupting data

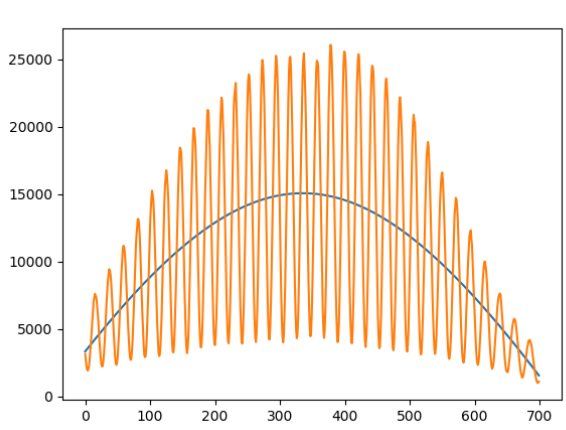
(b) Peaks cut off at height of neighbors

Figure 5: Mitigating the saturation error

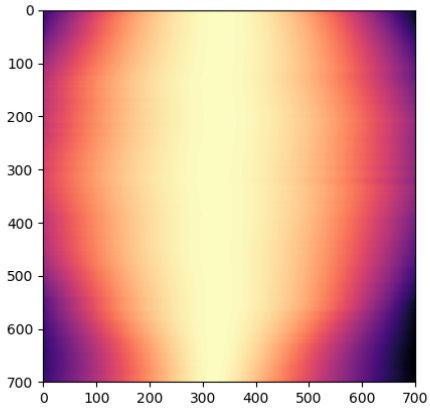
4.3.2 Curve fitting in least square sense

As mentioned before, to counter the vignetting, we fit a curve through our data on each row of the interference pattern. We do this by changing the parameters of the curve $f(x)$ to minimize $\sum_x (f(x) - p(x))^2$ where $p(x)$ is the intensity at the pixel x of the chosen row. If we have the parameters, the next step is to calculate the sensitivity function. We scale the function so that the maximum of the function equals 1. Then we calculate the sensitivity function by inverting the downscaled function: $s(x) = \frac{1}{f(x)}$. If we multiply every intensity $p(x)$ with the corresponding value $s(x)$, the fitting curve is constant one. That way the monochromatic interference pattern of a laser should be converted from a bell curve to constant height.

First, we started using a parabolic function $f(x) = ax^2 + bx + c$ as a fitting curve. Although it serves as a good approximation, the downside to parabolic functions or even any polynomial function is that the dropoff at the edges is too steep. Using an polynomial function to create the sensitivity function makes the corrected data points oscillate strongly at the edges as in Figure 7. This means that the parabolic function overachieves the correction of the data points at the edges and makes them useless. Because of this effect, the next idea was to use an hyperbolic function $g(x) = a\sqrt{1 + \frac{(x+b)^2}{c^2}} + d$. These have almost the same properties as parabolas near the maximum, but they fall off asyomptotically linear on each side. This linear decrease stops the appearing of oscillating data points. Fitting this curve on every row of the laser image, calculating the sensitivityfunction and correcting the vignetting gives great results with the hyperbolic function. If we use this method of cutting off the saturation peaks and counter the vignetting with a hyperbolic sensitivity function on Figure 2a we obtain Figure 8. Because we only did the curvefitting for every row

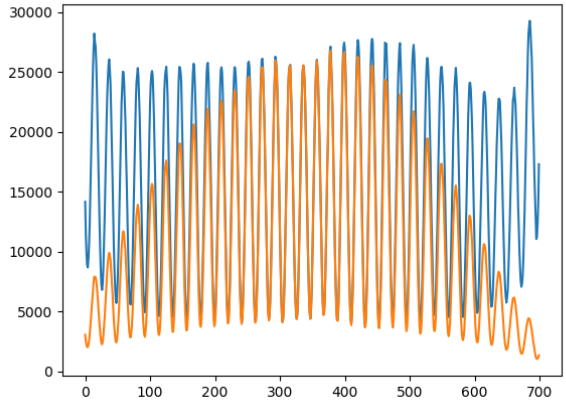


(a) Data of one row fitted with a curve

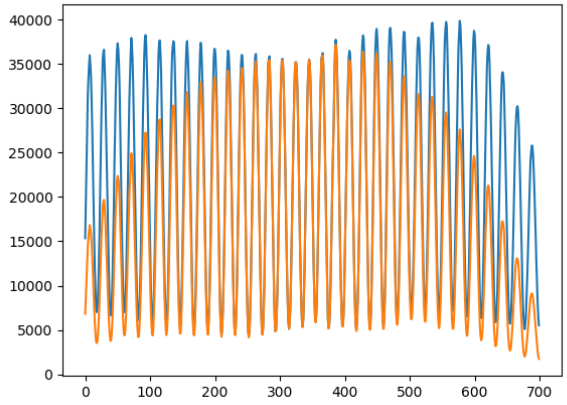


(b) Curvefitting every row of one Laser Image (x: Pixel, y: Fittingcurve)

Figure 6: Curvefitting in least square sense



(a) Parabolic fittingcurve with oscillation at the edge



(b) Hyperbolicfitting curve

Figure 7: Using the sensitivity function to counter vignetting

respectively, there is still some vignetting left at the bottom and top of the image. This results from the absence of a maximum to raise the rest of the peaks according to it. But this is not a problem as we are more interested in the middle of the image.

5 Results

By applying the described methods of tilting and curve fitting to our exemplary synthetic detector image in Figure 3 and doing a Discrete Fourier Transformation on the middle row of the resulting image, the spectrum is visualized in Figure 9. Due to the decreased horizontal pixel number a transformation of the x-axis to a wavelength scale without an acceptable error is not possible. The same holds for the result on the detector measurement of the calibrated laser light after applying tilting, pincushion, saturation of the pixel and curve fitting algorithms, see Figure 10.
For the measurements in the orbit the detected image differs from the ones of monochromatic light by the area of interesting data. In Figure 11 the detected image is presented by sizing the interval of the considered detected intensities down, otherwise the interferogram would be flat. Only in the middle of the image are data stored we are interested in, so for our analysis we looked at the 10 rows

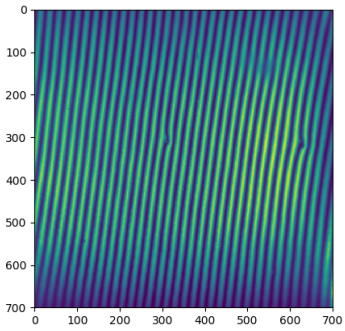


Figure 8: Countering Saturation and Vignetting on 763.109nm laser

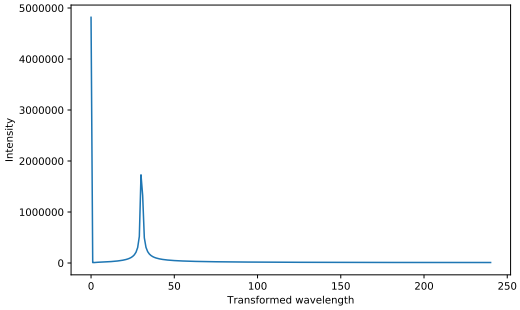


Figure 9: Result of DFT on synthetic tilted and flattened image of wavelength 763.109nm

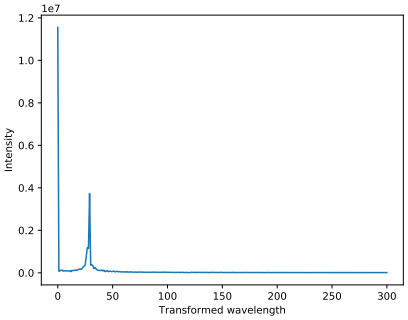


Figure 10: Result of DFT on laser calibrated measurement of wavelength 763.109nm

in the middle of the picture. First we applied the curve fitting on the median of those rows and after that used the DFT to get the composition of the light spectrum and intensities. In comparison to that, the same plot contains the result of the DFT only on the median of the rows.
The graphs do not differ much, but the spectrum of the version with applied curve fitting has some higher peaks and the lower intensity wavelengths of the smaller wavelengths have a smaller intensity.

References

[1] Michael Deiml. *Development of a Small Satellite Remote Sensing Payload for Passive Limb Sounding of the Atmospheric Oxygen Emission*. PhD thesis, Bergische Universität Wuppertal, 2018. http://elpub.bib.uni-wuppertal.de/servlets/DocumentServlet?id=7530.

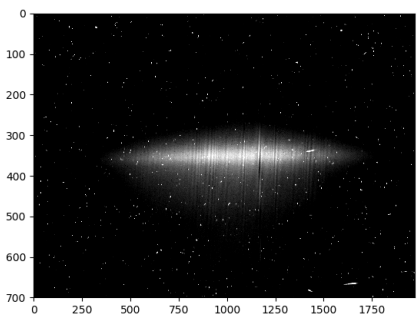


Figure 11: Cutout of the detected orbit image in a certain interval of intensities

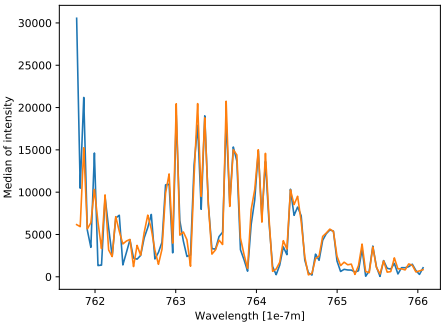


Figure 12: Comparison of the DFT on the median with applied curve-fitting (orange) and without (blue)