

1222·2022
800
ANNI



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

dii DIPARTIMENTO
DI INGEGNERIA
INDUSTRIALE

Università degli Studi di Padova
Dipartimento di Ingegneria Industriale (DII)

Corso di Laurea Magistrale in Ingegneria Aerospaziale

Tesi di Laurea Magistrale

Analysis of HYPSONS diffraction grating and detector

Relatore: Prof. Giampiero Naletto

Laureando: Andrea Meneguzzo

Matricola: 1234055

ANNO ACCADEMICO 2021 - 2022

ABSTRACT

In order to obtain information related to the morphology (3D) and composition (i.e. spectral data) of a surface, two instruments are required: a stereocamera and a spectrograph. However, the use of two separate instruments is not trivial, as they must be cross-calibrated with each other.

The *Hyperspectral Stereo Observing System HYPSONS*, is designed to overcome this and other obstacles, as it will be an instrument comprising both a stereocamera and a spectrograph. *HYPSONS* will therefore be able to provide 4D information simultaneously: spatial (3D) and spectral (1D).

This prototype is currently being developed at the LUXOR laboratory of the Institute of Photonics and Nanotechnology in Padua.

SOMMARIO

Per poter avere informazioni legate alla morfologia (3D) e alla composizione (quindi informazione spettrale) di una superficie sono necessari due strumenti: una stereocamera e uno spettrografo. Tuttavia, l'utilizzo di due strumenti separati non è banale, in quanti questi devono essere cross-calibrati tra loro.

Il *Sistema di Osservazione Stereo Iperspettrale HYPPOS*, nasce per superare questo ed altri ostacoli, in quanto sarà uno strumento che comprenderà sia una stereocamera che uno spettrografo. *HYPPOS* sarà quindi in grado di fornire contemporaneamente informazioni 4D: spaziali (3D) e spettrali (1D). Tale prototipo è attualmente in fase di sviluppo presso il laboratorio LUXOR dell'Istituto di Fotonica e Nanotecnologie di Padova.

CONTENTS

Chapter 1 - INTRODUCTION	pag.1
Chapter 2 – WHAT IS LIGHT?	pag.2
Chapter 3 – HYPSONS	pag.4
3.1 What is HYPSONS	pag.4
3.2 HYPSONS optical configuration	pag.5
3.3 On going activities	pag.6
3.4 Some information	pag.7
Chapter 4 – DIFFRACTION GRATINGS	pag.8
4.1 What is Diffraction?	pag.8
4.2 Diffraction Gratings and Prism	pag.11
4.3 Diffraction Gratings: equation and main parameters	pag.14
4.3.1 <i>Grating equation</i>	pag.14
4.3.2 <i>Angular Dispersion</i>	pag.16
4.3.3 <i>Resolving Power</i>	pag.17
4.3.4 <i>Spectral Resolution and Free Spectral Range</i>	pag.18
4.4 Processing of Diffraction Gratings	pag.19
4.4.1 <i>Mechanical Ruling</i>	pag.19
4.4.2 <i>Holographic gratings</i>	pag.20
4.4.3 <i>Replica Gratings</i>	pag.21
4.4.4 <i>Ruled vs Holographic</i>	pag.21
4.5 Type of diffraction gratings	pag.23
4.5.1 <i>Plane gratings</i>	pag.23
4.5.2 <i>Concave gratings</i>	pag.24
4.6 Grating Efficiency	pag.25
4.6.1 <i>Absolute and relative efficiency</i>	pag.25
4.6.2 <i>Littrow condition</i>	pag.26
4.6.3 <i>Blazed Gratings</i>	pag.26
4.6.3.1 <i>Interference and Blaze function</i>	pag.27
4.6.3.2 <i>Blazed holographic gratings</i>	pag.30

Chapter 5 – CHARACTERISATION OF THE GRATING	pag.31
5.1 Characterisation with AFM	pag.31
5.1.1 <i>Elements and operating principle of AFM</i>	pag.31
5.1.2 <i>Results and Discussion</i>	pag.32
5.2 Characterisation of the Diffraction Grating	pag.35
5.2.1 <i>Characterisation of the photodiodes</i>	pag.35
5.2.2 <i>Characterisation of the Diffraction Grating</i>	pag.40
5.2.2.1 <i>Zero angle of incidence</i>	pag.43
5.2.2.2 <i>Near-Littrow configuration</i>	pag.45
5.2.2.3 <i>HYPPOS like</i>	pag.47
5.2.2.4 <i>Conclusion</i>	pag.49
Chapter 6 – SILICON SEMICONDUCTOR DETECTORS	pag.51
6.1 Material's energy band	pag.51
6.2 Photoelectric effect	pag.53
6.3 Charge Coupled Device (CCD)	pag.56
6.3.1 <i>CCD architecture and principle of work</i>	pag.56
6.3.2 <i>CCD illumination mode</i>	pag.60
6.4 Active Pixels Sensor Complementary Metal-Oxide-Semiconductor (APS-CMOS)	pag.62
6.4.1 <i>APS CMOS architecture and principle of work</i>	pag.62
6.4.2 <i>CMOS readout mode</i>	pag.64
6.5 Difference between CCD and CMOS image sensor	pag.65
6.6 APS CMOS in space	pag.66
6.7 Noise	pag.66
Chapter 7 – CHARACTERISATION OF THE DETECTOR	pag.67
7.1 Case 1	pag.68
7.1.1 <i>Whole surface analysis</i>	pag.69
7.1.2 <i>Divided surface analysis</i>	pag.71
7.1.3 <i>Analysis with integrating sphere</i>	pag.74
7.2 Case 2	pag.79
7.3 Conclusions	pag.82
RINGRAZIAMENTI	pag.83
BIBLIOGRAPHY	pag.84

CHAPTER 1

INTRODUCTION

This thesis is concerned with the analysis of the *HYP*SOS diffraction grating and detector.

The thesis is divided into three parts.

The first part briefly describes light, its nature and what is involved in its study. Also in this first part, *HYP*SOS is briefly described in terms of its optical design and potential.

The second part deals with the theory of diffraction gratings and the *characterisation of the diffraction grating* in terms of morphology and absolute efficiency.

The third part of the thesis deals with silicon semiconductors and concludes with *linearity analyses performed on the CMOS sensor*.

CHAPTER 2

WHAT IS LIGHT?

This section will briefly describe what *light* is and what is involved in studying it.

Light is a form of energy: it can be produced for example by heating a metal. In fact when a metal is heated its color change, it becomes orange and in the end white. So *light* emitted by a body depends on the body temperature, which is a form of microscopic kinetic energy.

Light can propagate through reflection or refraction, depending on the type of interface present between two different media.

Reflection is at the basis of the colours of a given body/substance perception. For example, the green colour of grass is due to the fact that this 'body' absorbs all other colours and reflects only green.

Refraction, on the other hand, explains (as will be seen below) how a light ray can be bent when it encounters a medium with a different refractive index.

But *light*, is a particle or a wave? It took two centuries to answer this question and the answer is: both. It is a particle because it is able to carry energy, interacts with surfaces and disperses; it has a wave-like behaviour because when two or more waves meet, interference phenomena arise.

Another question which has to be answered is: where exactly does *light* originate?

Light is simply a manifestation of the propagation of electromagnetic radiation, generated from oscillating electrical charges. This explains why heating a metal causes it to emit light! By supplying energy to the material, in this case to the atoms that make it up, they 'vibrate', i.e. they oscillate thanks to their electrical charge. This vibration generates an electromagnetic wave, which among other effects is also responsible for the manifestation of visible light.

The understanding of what has just been described was made possible thanks to Clerk Maxwell, who described first the laws of electromagnetism: through these he explained how electric and magnetic fields propagate in space and interact each other, forming the electromagnetic field, which is the basis of electrical, magnetic and optical phenomena.

His equations in the vacuum are the follows:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (2.1)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.2)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.3)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (2.4)$$

where:

- ' ρ ' is the charge density;
- ' \mathbf{E} ', ' \mathbf{B} ' are the electric and magnetic fields respectively;
- ' \mathbf{j} ' is the current density;
- ' ϵ_0 ', ' μ_0 ' are the dielectric constant and the magnetic permeability.

Another important detail to consider is that these waves travel through the vacuum at the speed of light 'c':

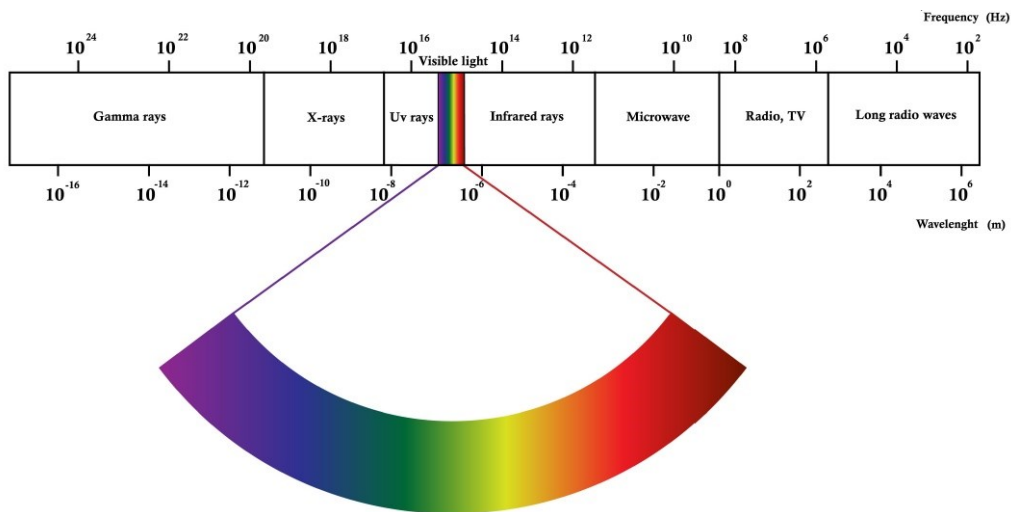
$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \quad (2.5)$$

When we talk about a colour, for example green, actually we are referring to a particular wavelength (λ) and in turn to a particular frequency (f) due to the relation:

$$v = \lambda f \quad (2.6)$$

where 'v' is the velocity of the waves.

The human eye can distinguish colours, i.e. a certain range of wavelengths from approximately [380-780] nanometers. However, this does not mean that nothing exists outside this range; on the contrary, this range is called the '*Visible Spectrum*' and is only part of the '*Electromagnetic Spectrum*' (Figure 2.1):



(Figure 2.1: *Electromagnetic Spectrum*)

At the end of this introduction, one last question remains to be answered: why do we study *light* and how? Though the spectroscopic analysis, which is a measurement of the amount of energy carried by the wave that can be converted in the wavelength analysis, we can obtain information about the composition of the source, planets, their temperature, their evolution and so on (for example, Black Holes can be studied by means of X-ray radiation).

CHAPTER 3

HYPSONS:

This chapter briefly describes the *Hyperspectral Stereo Observing System*, namely *HYPSONS*, the prototype being developed at the LUXOR laboratory of the Institute of Photonics and Nanotechnologies in Padua, highlighting its characteristics and potential.

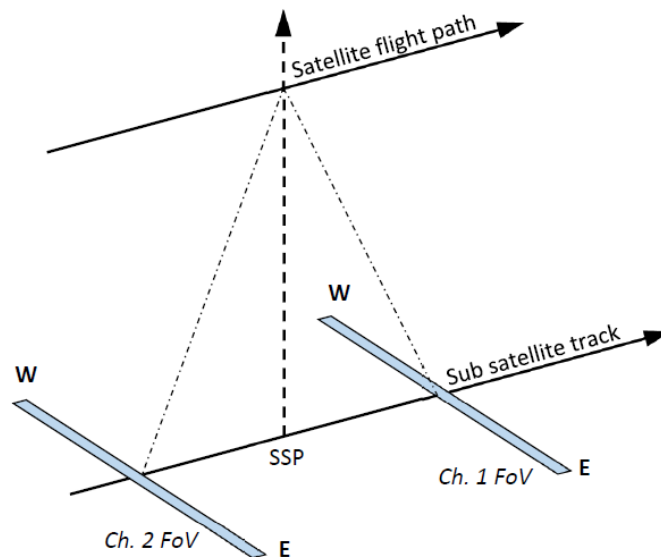
3.1: What is HYPSONS

In recent years there has been increasing interest in the study of morphology (3D), i.e. the surface structure of the planets: this is made possible by digital terrain model (DTM) of the surface, by means of the data obtained by a stereocamera on board of the satellite.

However, through the DTM technique, it is possible to obtain only the morphological information (i.e. 3D) of the object to be studied; to obtain information regarding its composition, it is necessary to use spectrographs. Unfortunately, merging the data between these two instruments is not trivial, as they must be cross-calibrated and very often the FOV is different, which makes it difficult to reconstruct a 3D image with spectral information: this therefore implies a possible degradation of the image.

It is precisely to overcome this problems that *HYPSONS* has been designed: it will enable spectral terrain models (SDTM) to be obtained. In this way, it will be possible to obtain both 3D (i.e. spatial) and spectral information, i.e. 4D information, with a single instrument.

The attitude system of *HYPSONS* will be Nadir-Pointing, i.e. one axis points towards the centre of the target of interest in order to analyse it, and its mode of operation will be push-broom: it will consist of two channels tilted by $\theta_{1,2} = \pm 20^\circ$ (Figure 3.1) with respect to Nadir along the flight direction.



(Figure 3.1: Satellite motion for stereo acquisition)

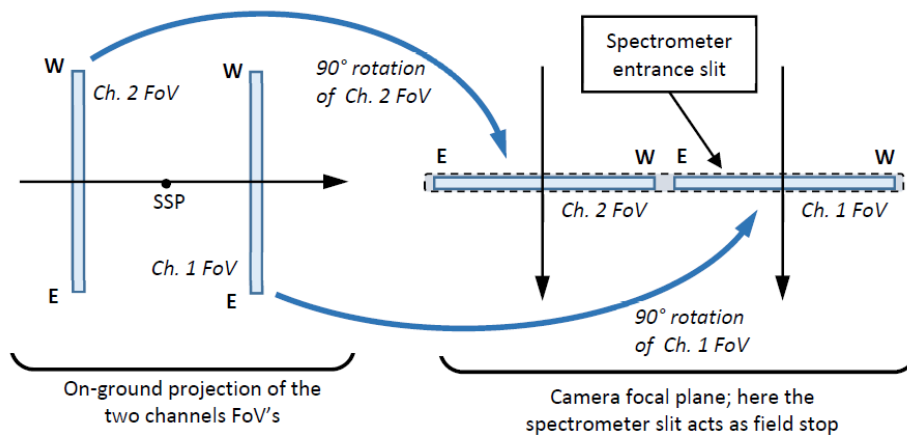
The potential of HYPSONS are enormous and it can be used in a number of sectors including space, geology and agriculture, for example.

3.2: HYPPOS optical configuration

As mentioned above, the optical configuration of *HYPPOS* consists in one single instrument which combined a spectrograph and a stereoscopic pushbroom camera.

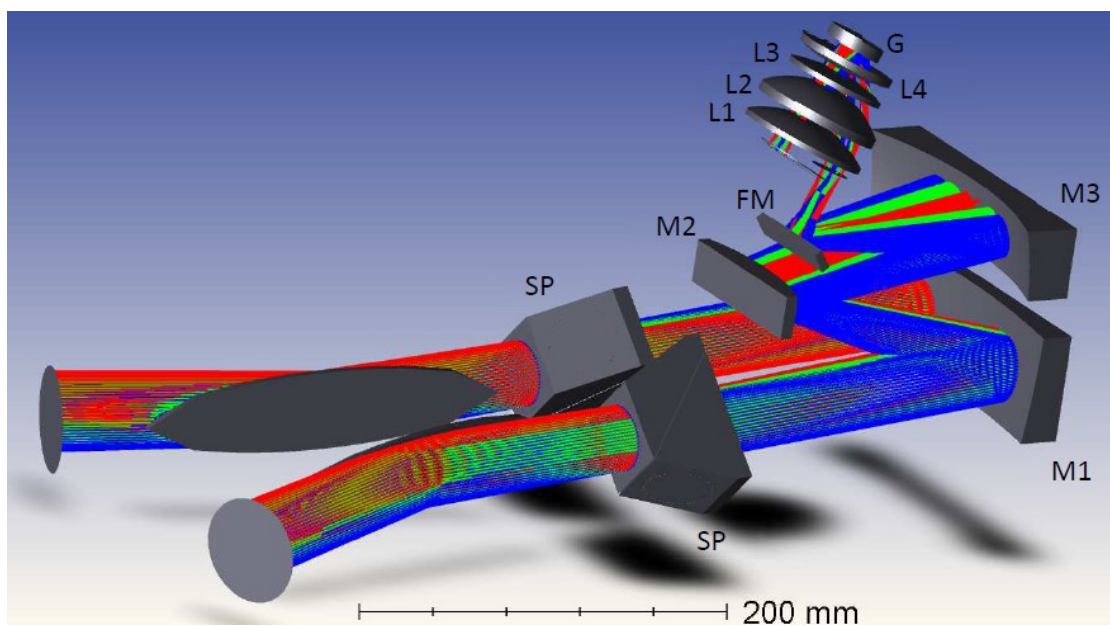
This was done by taking as a reference the STC stereocamera of SIMBIO-SYS, which is the imaging system of ESA's BepiColombo mission to Mercury: this device, in fact, consists of two channels, tilted $\theta_{1,2} = \pm 20^\circ$ with respect to the Nadir direction, which send the light collected in the same sensor.

This strategy was also adopted for *HYPPOS* in order to minimise mass and resources. As said above, what distinguishes *HYPPOS* is the presence of the spectrograph: this means that the system must be designed in such a way that the collected light beams are sent into the entrance slit of the spectrometer (Figure 3.2).



Returning to the instrument: the stereo camera consists of a three mirror anastigmat telescope (TMA), which is found to minimise spherical, astigmatism, and coma aberration.

Figure 3.3 shows the *HYPPOS* configuration:

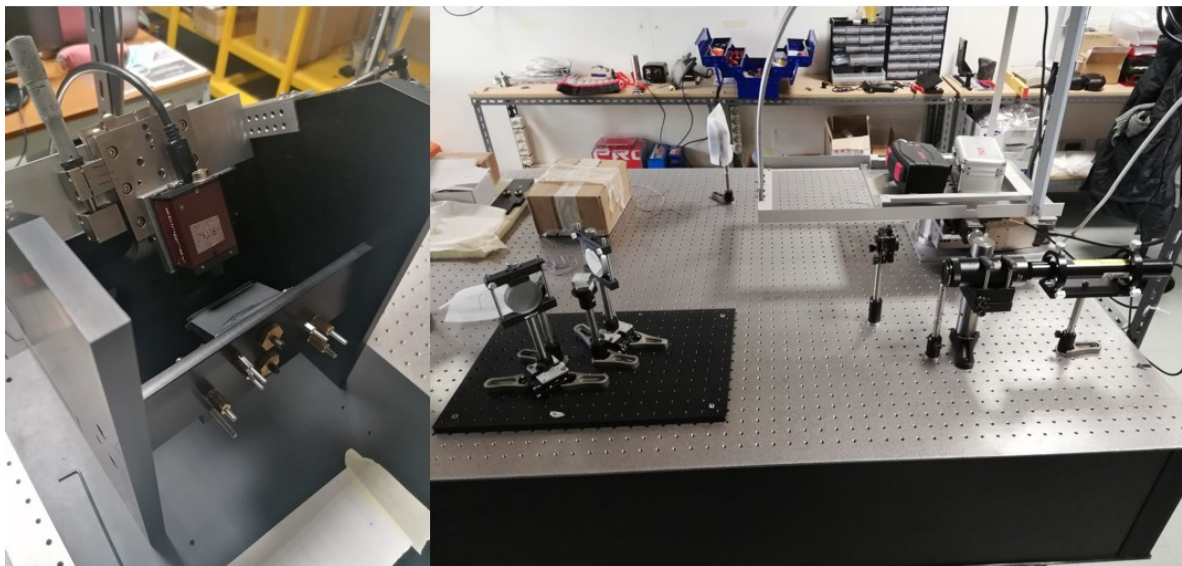


In particular, *HYPPOS* will consist of: two flat mirrors, which collect the light from the two channels; two 45° tilted Schmidt-Pechan prisms, which aim to rotate the FOV by 90° (as seen in Figure 3.2). Once passed the prisms, the two beams of light enter in the TMA consisting of three mirrors (M1, M2 and M3 respectively), and a flat folding mirror (FM), which focuses the rays on the entrance slit of the spectrometer. Once through the slit, the rays enter the spectrometer, which consists of four lenses (L₁, L₂, L₃ and L₄) and a concave diffraction grating (G). The grating diffracts the incident rays, and the diffracted beam passes through the lenses again, until it is focused on a two-dimensional sensor.

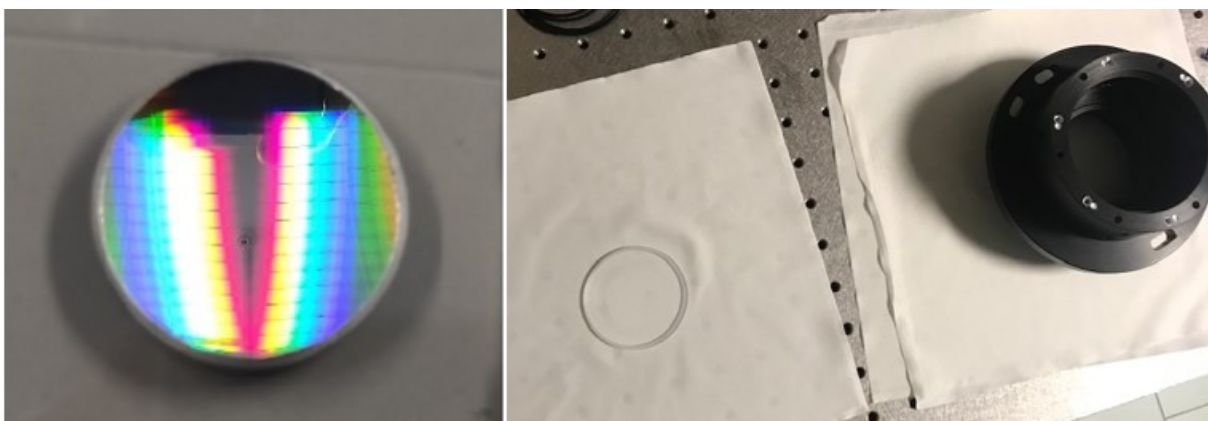
3.3: On going activities

Currently, the diffraction grating has been characterised and linearity tests have been performed on the sensor: those aspects will be described in section 5 and 7.

The next steps will be to align the telescope (Figure 3.4), assemble the spectrometer (Figure 3.5) and characterise the system, before moving on to the assembly of the final *HYPPOS* configuration shown in Figure 3.3.



(Figure 2.4: telescope and alignment system)



(Figure 2.5: spectrometer elements)

3.4: Some information

As the work of this thesis is based on the analysis of the diffraction grating and the sensor of *HYPSONOS*, the main characteristics of the two are shown below.

The following table shows the characteristics of the diffraction grating:

DIFFRACTION GRATING	
Grating type	Holographic concave spherical
Radius of curvature	83.7 mm
Ruling density	678 lines/mm

About the sensor, the following table shows some of its main characteristics:

DETECTOR	
Sensor type	Active CMOS
Shutter type	Global
Pixel size	3.45 μm
Imaging size	6840 x 4856 (22.36 mm x 16.75 mm)
Frame size	12 (streaming) / 14.4 (burst) [full frame]
ADC	12 bit
Comm.Protocol	USB3
Output	MONO8, MONO12, MONO12p

CHAPTER 4

DIFFRACTION GRATINGS:

This section will discuss diffraction gratings, starting with an understanding of diffraction and ending with the concept of efficiency.

4.1: What is *Diffraction*?

An interesting way to understand diffraction is to look at the waves of the sea as they meet an opening before ending their course when they reach the beach (Figure 4.1).



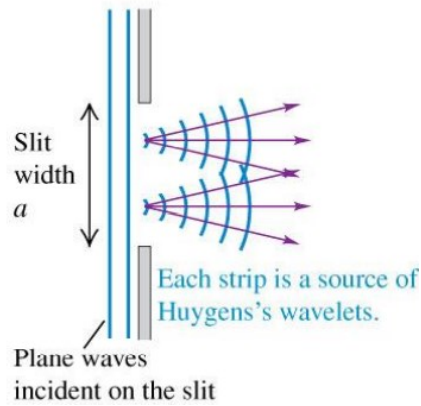
(Figure 4.1: Example of Diffraction)

When sea waves encounter an obstacle their direction of motion changes: this phenomenon is called *diffraction*.

Back to the light, *diffraction* is an interference phenomenon which takes place every time a wave encounters an aperture, which width is comparable to the wavelength of the incoming wave: so diffraction can be described with the wave model of light. Now, if the aperture width is much larger than the wavelength of the incident beam, the direction of the wave remains unchanged; instead, if the width of the aperture is of the same order of the wavelength of the light, or smaller, from the aperture the waves have different directions of propagation and in a certain point of the space, they overlap, giving rise to interference phenomena: it is as if new sources, all in phase with the original source, originated from the slit. In other words, each portion of the slit acts as a source of light wave (Huygens' principle).

It is important to notice that *diffraction* is wavelength dependent: this means that the effect of the aperture is different for the various wavelengths.

To better understand these aspect let's consider diffraction caused by a linear narrow slit (Figure 4.2):



(Figure 4.2: Diffraction by a linear slit)

As mentioned before, when the incoming light reaches the slit, it can be divided in N-equally spaced sources, where the distance between each other is 'dy'.

As said above, the direction of the light is changed, so the light propagates in every direction from the slit, and if a screen is put a certain distance (L) from the slit (where L>>a), what can be seen is a diffraction pattern, consisting of bright and dark areas (Figure 4.3).

From theory, the total phase difference of the waves emitted from the slit is:

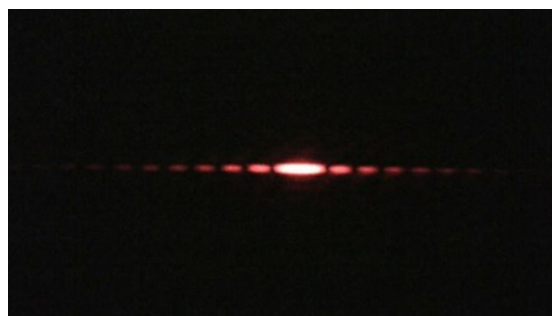
$$\Delta\phi = \frac{2*a*\sin \theta}{\lambda} \quad (4.1)$$

where: 'a' is the width of the slit, 'θ' is the angle of a point 'P' in the space and 'λ' is the wavelength of the incoming light.

From the theory of interference, it is known that constructive interference occurs when the phase difference is equal to zero or is equal to an even multiple of 'π':

$$\Delta\phi = 2m\pi, \text{ where } m = 0, \pm 1, \pm 2, \dots \quad (4.2)$$

Thanks to this, bright fringes can be seen on the screen (Figure 4.3):



(Figure 4.3: Example of bright fringes)

A further mathematical step gives:

$$a \sin \theta = m\lambda \quad (4.3)$$

From this equation, it is clear that for a different ' λ ' a different ' θ ' is obtained, so a different position of the bright fringes is reached: this represents the capability of the system to disperse the light, if this is polychromatic.

Instead if the phase difference is equal to an odd multiple of ' π ', destructive interference occurs:

$$\Delta\phi = (2m + 1)\pi, \text{ where } m = 0, \pm 1, \pm 2, \dots \quad (4.4)$$

In this case no fringes are present on the screen.

Now, if instead of a single slit, there were several slits at the same distance from each other (comparable to the wavelength of the incident light), a device called a *diffraction grating* would be created.

4.2: Diffraction Gratings and Prism

In the previous section it was understood that a diffraction grating is a dispersive element consisting of N-slits separated at a certain distance from each other with the same order of magnitude as the light to be studied.

The physical characteristic of diffraction grating is defined through the spatial modulation of the refractive index; precisely, when an electromagnetic wave arrives on the surface of the grating, its electric field amplitude and phase will be modified in certain manner, due to the variation of the refractive index in the region near of the grating. From this, the phenomenon of diffraction happens and from the surface of the grating, different sets of diffracted waves are generated by constructive interference (as seen from Figure 4.3).

But what is and when *dispersion* occur? *Dispersion* is simply a phenomena which separates the light in different spectral components, characterized each one by a certain wavelength; this is due to the different phase velocity, the wavelength and on the material intercepting the light.

More precisely, dispersion occurs every time there is something which changes the direction of the incident light, in such a way which is strongly dependent on the wavelength.

The phase velocity of a wave (v), is defined as the ratio between the speed of light in vacuum (c) and the refractive index of the material (n) in which light propagates:

$$v = \frac{c}{n} = \frac{\lambda f}{n} \quad (4.5)$$

where ' n ' is a number equal or greater than 1.

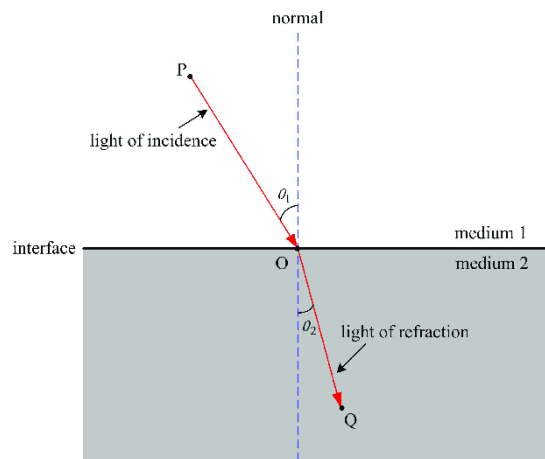
In general the refractive index is a function of the wavelength of the incoming light (according with the Cauchy equation):

$$n(\lambda) = A + B/\lambda^2 \quad (4.6)$$

where ' A ' and ' B ' are constant for a specific material.

This means that the phase velocity is a function of the wavelength; consequently different wavelengths making up white light have different phase velocities.

To get a clearer idea of what *dispersion* is, let's now consider white light, which hits a prism, characterized by a refractive index ' n_p '. From Snell's equation is well known that when a luminous rays intercepts a medium which divides two zone having different refractive index, the phenomenon of refraction occurs (Figure 4.4):



(Figure 4.4: Refraction phenomena)

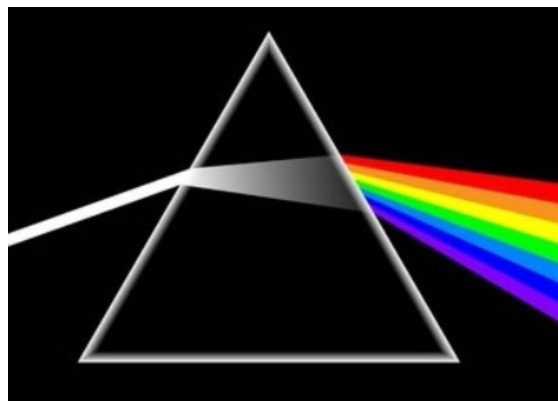
Considering the figure above, knowing the angle of incident ' θ_1 ', the refractive index ' n_1 ' and the refractive index ' n_2 ', it is possible to obtain the direction of the refractive beam ' θ_2 ' respect to the vertical, by means of the *Snell law* refraction, just by applying the relation below:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (4.7)$$

Instead the relation between the two phase velocities is given by the following formula:

$$\frac{v_1}{v_2} = \frac{n_2}{n_1} \quad (4.8)$$

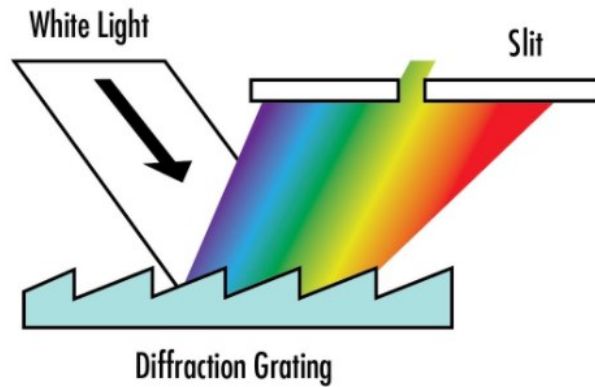
Returning now to the beam of white light and the prism: from the discussion above and remembering that the refractive index is a function of wavelength, it is clear that the various wavelengths that make-up white light will be refracted at different angles; so at the exit of the prism we will see the various wavelengths, in other word, the rainbow (Figure 4.5).



(Figure 4.5: Refraction by a prism)

Let us now return to diffraction gratings and consider reflective gratings in particular.

From the discuss done before, it is now clear that if monochromatic light hits the surface of the grating, this will be diffracted in its different orders, but remains monochromatic; instead if polychromatic light hits the grating, this will be diffracted in its various wavelength in different directions (Figure 4.6).



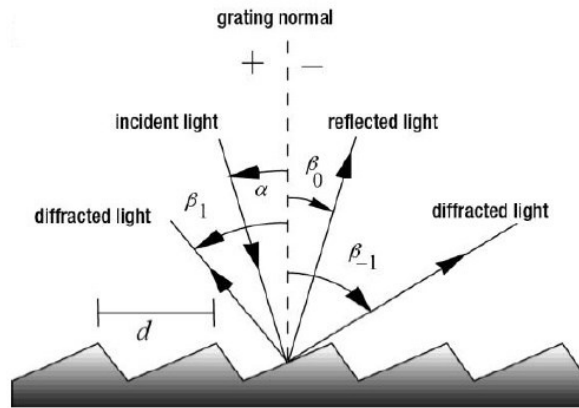
(Figure 4.6: reflective diffraction grating)

To conclude: the difference between prism and grating spectrum is that in prisms the spectrum is created due to the scattering of light, i.e. it passes through surfaces with different refractive index, whereas in the grating, the light is diffracted: i.e., dispersion in the prism is due to the phenomenon of refraction, whereas in the case of gratings it is due to the fact that each grooves acts as a source of secondary waves that interfere at a certain point in space.

4.3: Diffraction Gratings: equation and main parameters:

4.3.1: Grating equation:

To understand how a *diffraction grating* operates, let's consider Figure 4.7:



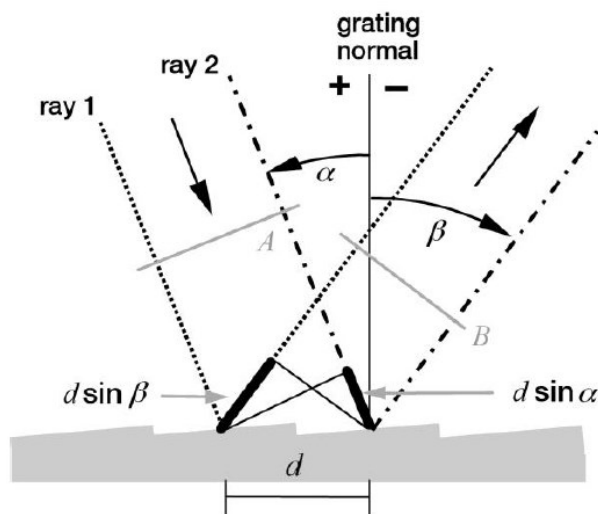
(Figure 4.7: reflective diffraction grating)

where:

- ' d ' is the grooves spacing in millimeters ;
- ' α ' is the angle of the incident beam;
- ' β_n ' are the various diffracted angles for different orders.

By convention, both incidence and diffracted angles, are measured with respect to the grating's normal: this means that they have the same sign if they stay on the same side of the normal.

Let's now consider the following picture (Figure 4.8), where wavefronts are used:



(Figure 4.8: diffraction grating with wavefronts)

The geometrical path difference of the light from adjacent grooves is:

$$\Delta = d \sin \alpha + d \sin \beta \quad (4.9)$$

where ' $\sin \beta$ ' < 0, due to the sign convention present in the figure above.

From the principle of the interference, it is known that, constructive interference occurs, only when the different of path of light is equal to a integer number of the wavelength under study: this means that the various maxima orders can be seen:

$$m\lambda = d \sin \alpha + d \sin \beta \quad (4.10)$$

where ' m ' is the diffraction order and is equal to ' $\pm 2, \pm 1, 0$, etc..'

The previous equation is the so called *grating equation*.

A particular case happens when ' $m = 0$ ': in this situation the grating behaves like a mirror, so if polychromatic light affects the grating, the wavelengths will not be separated. When the diffraction order is equal to zero, this mean that the angle of incidence and of diffraction are equal (in absolute value):

$$\alpha = -\beta$$

and this is called specular reflection.

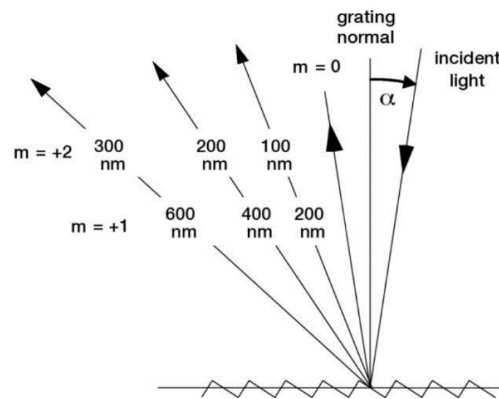
From the grating equation it is easy to see which spectral orders can exist, known ' λ, d '. To understand this, let's rewrite the grating equation in a different way:

$$\frac{m\lambda}{d} = \sin \alpha + \sin \beta \quad (4.11)$$

the maximum and minimum value that the function can assume is ± 2 , since $-1 \leq \sin \gamma \leq 1$, so the condition to evaluate if a determinate order exists by the following inequality:

$$-2 \leq \frac{m\lambda}{d} \leq 2$$

One of the main problems with gratings is the overlapping of the spectral lines (Figure 4.9).



(Figure 4.9: example of spectral overlap)

To understand this, consider white light as a source. The $\lambda=800$ nm diffracted in the first order ($m = 1$) will overlap with the $\lambda=400$ nm of the second order ($m = 2$) and this is also true for other wavelengths. This then leads to a loss of spectral information, because the sensor is not able to distinguish the two, or more, wavelength as the two spectral lines overlap: a technique to mitigate this problem is to use 'order filters'.

Now some properties of gratings will be introduced: angular dispersion, resolving power, spectral resolution and free spectral range.

4.3.2: Angular dispersion:

The angular dispersion is the measurement of the separation between diffracted wavelengths: its represents the spectral range per unit of length.

To obtain the angular dispersion (D) of a grating, simply take the grating equation, isolate the diffracted angle (β) and differentiate the expression with respect to the diffraction angle (assuming the incident angle is constant), so:

$$D = \frac{d\beta}{d\lambda} = \frac{m}{d \cdot \cos \beta} \quad (4.12)$$

if the grooves density increase, the angular dispersion increase.

4.3.3: Resolving Power:

The resolving power (R) is defined as the capability of the grating to separate two almost equal wavelengths, where $\lambda_1 \sim \lambda_2 \sim \lambda$ and $\Delta\lambda = \lambda_2 - \lambda_1$ represents the limit of resolution.

By the way, the resolving power can be written as:

$$R = \frac{\lambda}{\Delta\lambda} \quad (4.13)$$

If the incident radiation illuminates N -grooves, the resolving power at 'm-th' order can be written as:

$$R = \frac{\lambda}{\Delta\lambda} = Nm \quad (4.14)$$

by observing the last formula, the resolving power increases as the diffraction order increases and as the number of illuminated grooves increases.

Let's now make two hypothesis:

1. the groove spacing ' d ' is uniform all over the grating surface;
2. the substrate of the grating is planar;

by replacing the grating equation in equation 4.14, we obtain:

$$R = \frac{Nd(\sin \alpha + \sin \beta)}{\lambda} \quad (4.15)$$

Now, applying the two previous hypothesis, the last equation become:

$$R = \frac{W(\sin \alpha + \sin \beta)}{\lambda} \quad (4.16)$$

Where ' W ' is the ruled width of the grating.

If we take another step, we see that the maximum value that the numerator can take is exactly ' $2W$ ', so the maximum resolving power becomes:

$$R_{max} = \frac{2W}{\lambda} \quad (4.17)$$

Observing the penultimate formula, it emerges that the resolving power is a function not only of the angle of incidence and diffraction, but also of the quality of the grating surface, the uniformity of the grooves spacing and the quality of the substrate.

To better understand this concept, consider a grating of 60 mm wide with 700 grooves: the resolving power at the first order of diffraction ($m=1$) is: $Nm = 198000$, which implies a resolution of 0.003 nm at a wavelength of 500 nm.

In the end, considering the equation 4.14: if $m=0 \rightarrow R = 0$, so no wavelength can be distinguished (this was just demonstrated before).

4.3.4: Spectral Resolution and Free Spectral Range:

The *spectral resolution* is the minimum wavelength difference ' $\Delta\lambda$ ' between two wavelengths that can be resolved without any ambiguity by the system, while the range of wavelengths for a certain order, where the phenomena of the spectral overlap does not occur is called the *free spectral range* (F_λ):

$$F_\lambda = \frac{\lambda}{m} \quad (4.18)$$

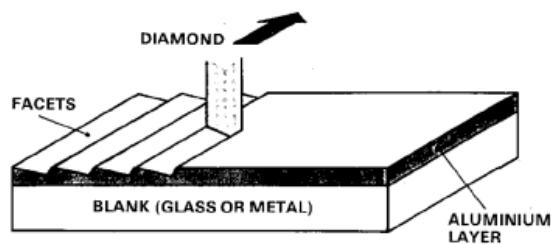
4.4: Processing of diffraction gratings

Let's discuss now on how about diffraction grating can be made. There two ways to realize this device which are: 'mechanically ruled' or via interference, the so called 'holographic gratings'.

4.4.1: Mechanical Ruling

With the *mechanical ruling* process (Figure 4.10), the grating are realized by burnishing the grooves individually with a tool realized in diamond material, on a thin coating evaporated metal layer which has been applied on the surface of the grating.

To realize the best grating ruling, in terms of both ruling equispacing and parallelism, a high precision interferometric machines is used.



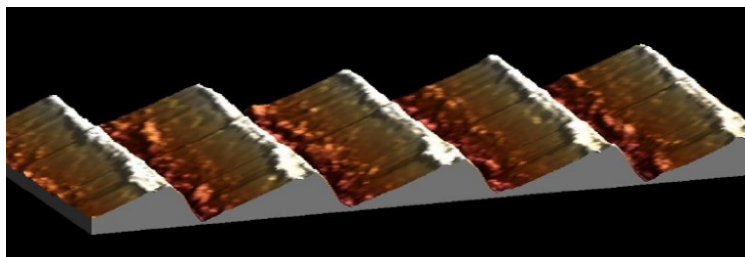
(Figure 4.10: Mechanical Ruling of a grating)

With this process, particular attention must be paid to the diamond tooth: it will define the shape of the various grooves, thereby impacting the efficiency and quality of the grating.

The device that guides the tooth back and forth during the ruling process must maintain its position within a few nanometers throughout the entire machining process, which can take weeks.

However, attention must also be paid to the vibration transmitted to the tooth, temperature and pressure variations in the working chamber: typically, the maximum permissible temperature variations are around 0.01°C. The main consequence of a temperature and pressure change is a variation in the groove spacing: if this value change, also the performances of the grating will change, in according to the grating formula.

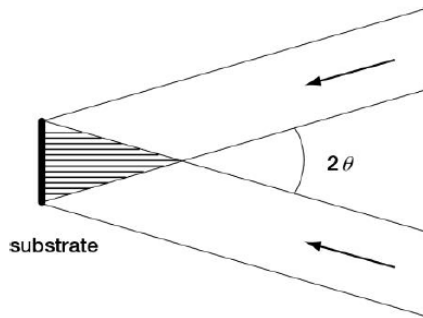
Once the grating is finished, it is called *master grating* (Figure 4.11), and it will be tested to verify its spectral resolution, efficiency and so its surface's quality. After these checks, if the grating satisfies all the specification it can be used to make the various *replica gratings*.



(Figure 4.11: Master ruled grating)

4.4.2: Holographic gratings

Holographic gratings instead are made by a layer of photosensitive material on a polished substrate and then it is exposed to interference fringes which are created by two coherent monochromatic optical plane waves of equal intensity and same wavelength ' λ '. From the intersection of these two plane waves, a standing wave pattern is formed in the region of intersection as can be seen in Figure 4.12:



(Figure 4.12: holographic process)

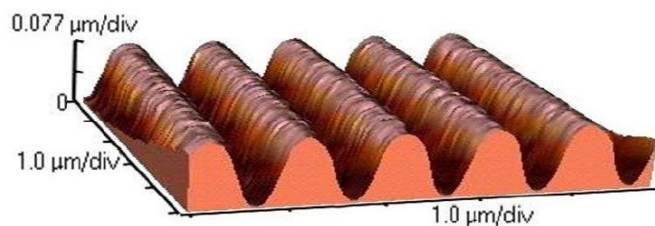
From the figure above, from the intersection of the two laser beams, an interference pattern composed of straight equally spaced planes fringes (bright and dark) arises. The combined intensity change in a sinusoidal way with the position as the interference pattern is scanned along a line.

The centers of adjacent fringes are separated from each other by a certain distance ' d ' given by:

$$d = \frac{\lambda}{2 \sin \theta} \quad (4.19)$$

where ' θ ' is the half angle which separates the two wavelengths (as it can be seen from Figure 4.12).

After this process, a chemical treatment applied on the photosensitive layer dissolved the exposed area and created the grooves, so a surface relief pattern is now formed with a sinusoidal profile (Figure 4.13).



(Figure 4.13: example of holographic grating)

This is the way in which a *master grating* is produced and then, by just applying a coating on the grating, *replica grating* can be done.

The profile of the grooves depends on the geometry of the substrate: if the substrate is planar, the grooves will be straight, equally spaced and parallel, instead if the substrate is concave, the grooves will be curved and non equally spaced.

It is also possible to use other wavefront surfaces, such as spherical wavefront. In this case the fringe pattern will not be equally spaced, so non-rectilinear and non-equally spaced grooves will be present on the surface of the grating. This technique is used to compensate some aberration.

The process just described, can be called *double beam interference* since two wavefront of the same monochromatic light and intensity are used.

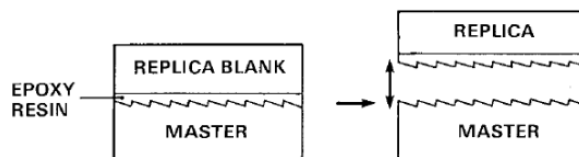
As in the previous process, certain precautions must be taken here too:

- the monochromatic light source must remain stable during all the process;
- ambient light must be turn-off, so that fringe pattern is maximal;
- temperature changes must be avoid, otherwise the local index of refraction in the beam of the interferometer will change.

Through these two main processes it has become clear which *master gratings* are made. It now remains to be understood how *replica gratings* are made from them.

4.4.3: Replica gratings:

To make a *replica* Figure 4.14, an epoxy resin is deposited in the interface between the master and the 'replica blank'. Once the resin has dried, the two are separated from each other: the resin remains attached to the replica, thus giving it the shape of the master.

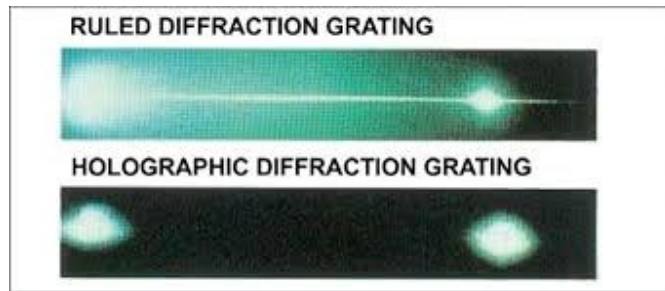


(Figure 4.14: replica grating)

4.4.4: Ruled vs Holographic:

A question now arises: which of the two types of grating is best? When looking at the time and cost aspects, the holographic type is the best candidate for both factors.

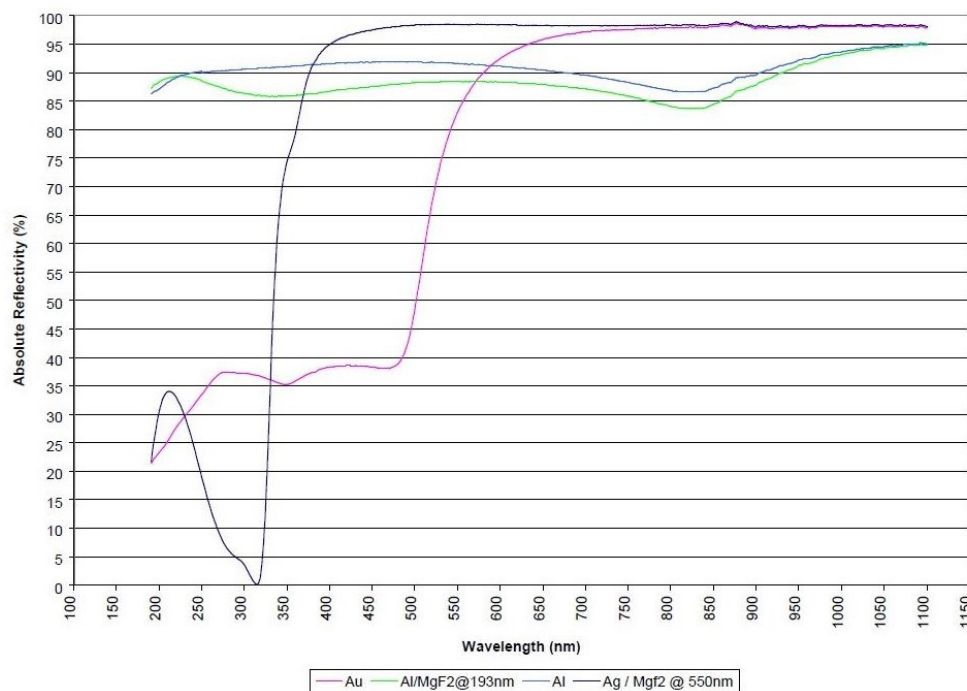
But if performance is taken into account then things change. In fact on the one hand the grooves may have a triangular or similar shape and thus provide a higher efficiency (when compared for the same wavelength and diffraction order), while on the other hand they have a sinusoidal pattern: they have a lower efficiency but reduce the scattering phenomenon and do not suffer from ghost lines (Figure 4.15).



(Figure 4.15: ghost line)

However, in addition to the shape of the grooves, the distribution of the grooves along the surface of the grating also affects the efficiency. For example, if concave gratings are considered, in the case of ruled grating the angle of the various facets varies along the surface; consequently the blaze wavelength will also vary. In the case of holographic gratings this phenomenon is much smaller (this concepts will be clear later). Ultimately, there is no ideal candidate, as each one has its advantages and disadvantages, but holographic ones are typically the most widely used.

For what concerns the material of the gratings: typically it is realized in glass; instead for the coating of the gratings tipycally aluminum, magnesium fluoride (MgF_2) and gold are used. Figure 4.16 shows the reflectivity curve of those three coating:



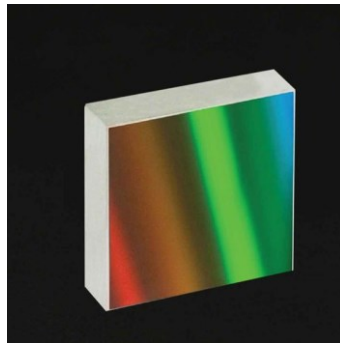
(Figure 4.16: reflectivity curve)

4.5: Type of diffraction gratings

Different type of gratings exist, such as: plane gratings and concave gratings. The main differences of those different type of gratings are described below.

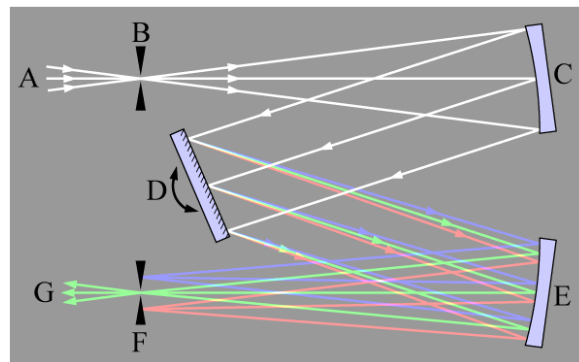
4.5.1: Plane grating

Simply, a plane grating (Figure 4.15), is a grating which its surface is flat. Those types of gratings are used when collimated light hits the surface, where the wavelengths are then dispersed but not focused: so they act like mirrors.



(Figure 4.15: example of plane grating)

Plane gratings are used, for example, in the so called *Czerny-Turned monochromator* (Figure 4.17): it must be said, that this configuration, where plane gratings are used, is not the only one.



(Figure 4.17: *Czerny-Turned monochromator*)

Where:

- A: is the incoming light source;
- B: is the entrance slit;
- C: is the collimated mirror;
- D: is the plane reflective grating;
- E: is the focusing mirror;
- F: is the exit slit;
- G: is the single channel detector.

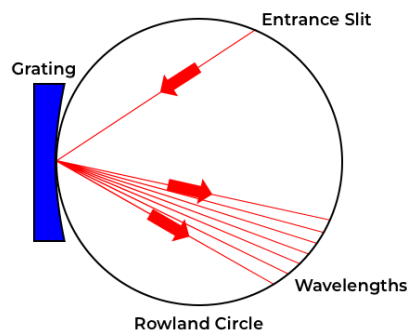
From the previous figure it is clear that a single wavelength can be studied. In order to study the other wavelength, the grating must be tilted (as it can be seen in Figure 4.17): so by tilting the grating with an appropriate manner, the study of the entire spectrum of interest can be done. What has just been explain is the principle of work of a monochromator.

4.5.2: Concave grating:

The main advantage of concave gratings with respect to plane gratings is the capability of focusing. Two main types of substrate shapes can be identified:

- toroidal gratings, whose surface is defined by two radii of curvature (precisely a tangential curvature and saggital curvature);
- gratings whose surface is a portion of a sphere of a certain radius: in these case, the saggital and tangential curvatures are equals.

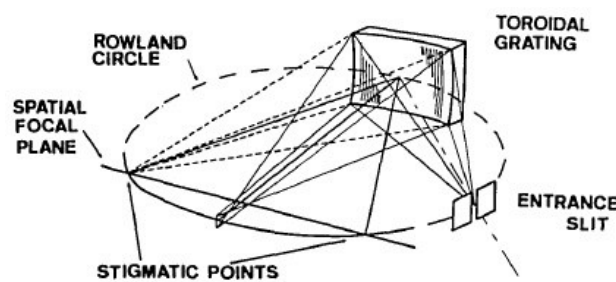
The simplest way of using concave gratings is through the so-called Rowland mounting. In this situation, the entrance slit, the vertex of the grating and projected spectrum, lie all on a circle of diameter R equal to the grating radius of curvature (Figure 4.17):



(Figure 4.17: Rowland mounting)

However, although the advantage of this configuration is that it focuses the diffracted beam from the grating along Rowland's circle, this configuration suffers from astigmatism.

In order to have a stigmatic line, the solution is to use toroidal gratings (Figure 4.18):



(Figure 4.18: Rowland mounting with toroidal grating)

Again, Rowland's mounting is only one of several configurations in which concave gratings can be used.

4.6: Grating Efficiency

This section will discuss the efficiency of a diffraction grating and its relationship to the type of grating processing.

4.6.1: Absolute and relative efficiency

For a reflection grating, the efficiency is defined as the ratio between the power of the light beam that the grating diffracts and the power of the incident light beam (at the same wavelength): this ratio defines the absolute efficiency of the grating for a given wavelength and diffraction order:

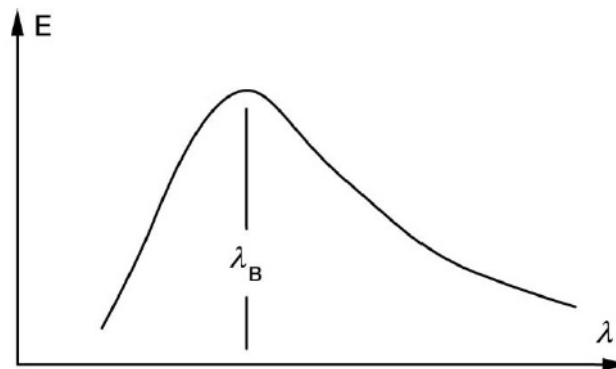
$$\eta(\lambda, m) = \frac{P_{d,\lambda}}{P_{i,\lambda}} \quad (4.20)$$

where:

- ' $P_{d,\lambda}$ ' is the power that the grating diffracts for a given wavelength;
- ' $P_{i,\lambda}$ ' is the power affecting the grating for a given wavelength;

In addition, the relative efficiency of the grating is also defined as a comparison of the energy diffracted into the desired order with the energy diffracted by a plane mirror that has been treated with the same coating as the grating.

Figure 4.19 shows an example of the absolute efficiency of a grating:



(Figure 4.19: example of grating absolute efficiency)

Typically, efficiency plots are referenced to a specific diffraction order. If this is not specified, the order represented is $m = 1$.

As can be seen from the graph, the efficiency has a peak at a certain wavelength, which is called the *blaze wavelength* (λ_B).

The efficiency of a grating depends on many factors, including: the material of the grating and its coating, the shape of the grooves and also the ratio ' λ/d '.

4.6.2: Littrow condition

This is a particular configuration where the angle of incidence and the angle of diffraction (for a given order and wavelength) are superimposed: in other words, the two beams have the same direction but opposite sense.

Given that: $\alpha = \beta$, from the grating equation results in:

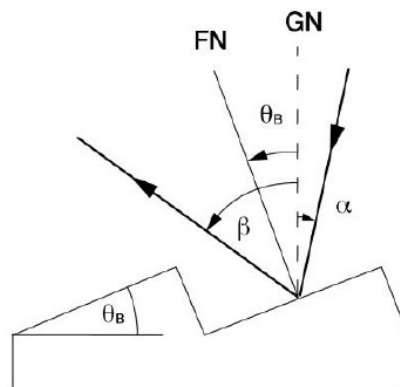
$$m\lambda = 2d \sin \alpha \quad (4.21)$$

4.6.3: Blazed Gratings

Typically, the highest efficiencies are obtained with gratings whose groove profile is triangular: from this it is clear that ruled gratings have higher efficiencies.

In this type of grating (in which the grooves have a triangular profile), ' θ_B ' is defined as the groove angle (see Figure 4.20). In this case, ' θ_B ' is the 'blaze angle'. λ_B , on the other hand, defines the blaze wavelength, where the maximum efficiency of the grating is reached.

Figure 4.20 shows a ruled blazed grating:



(Figure 4.20: ruled blazed grating)

Where: 'GN' is normal respect the plane of the grating, while 'FN' is the facet normal.

Gratings characterised by a profile of triangular grooves are called gratings with *sawtooth* profile.

Now a question arises: why do you need to change the profile of the grooves to increase efficiency, and what happens if the blaze angle is changed? For the second answer, simply ' λ_B ' changes, so the peak efficiency shifts with wavelength (this will be clear in the next sections). As for the first question, Lord Rayleigh deserves credit for the answer. He discovered that the diffracted energy distribution could be modified by changing the geometry of the grooves.

To better understand this aspect, the concept of *Blaze function* will be discussed below.

4.6.3.1: Interference and Blaze function:

As mentioned before, constructive interference occurs when the difference path of light is a multiple of ' π ':

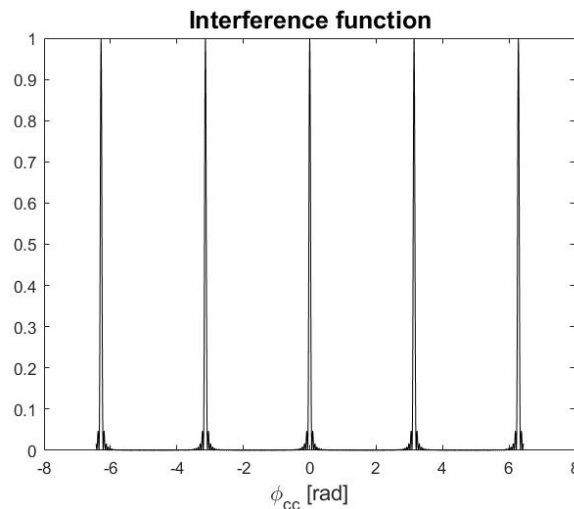
$$\phi_{cc} = \frac{\pi d(\sin \alpha + \sin \beta)}{\lambda} = m\pi \quad (4.22)$$

The '*Interference function*' is simply defined as the pattern of the normalized intensity for a certain wavelength:

$$I_i(\phi_{cc}) = \frac{\sin(L*\phi_{cc}/d)^2}{\sin(\phi_{cc})^2} \quad (4.23)$$

where the ' L/d ' ratio is the total number of grooves on the grating's surface and is an integer. Given a wavelength ' λ ', a range of interference phases occurs, with a peak of intensity when constructive interference occurs and a minimum when destructive interference occurs (' π ' is not multiplied by an integer number).

The shape of interference function can be seen with the following graph (Figure 4.21):



(Figure 4.21: example of Interference function)

The pattern of the '*Interference function*' is periodic and symmetric for all the orders ' m '. This function is modulated by another function, called '*Blaze function*', which occurs due to the interference phases between the half-distance of each facet. So considering the space between the center and the edge of a facet of the grating, the relation which gives the *phase modulation* is:

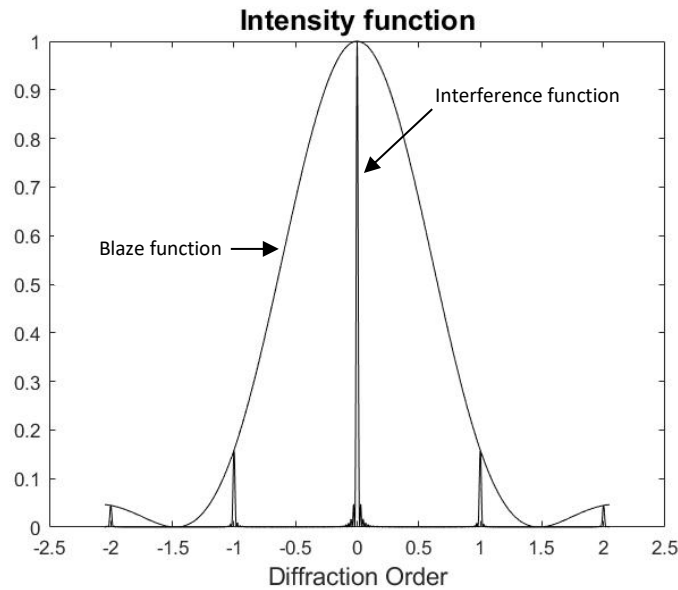
$$\phi_{ce} = \frac{\pi s(\sin \alpha + \sin \beta)}{\lambda} \quad (4.24)$$

where 's' is the grooves' width.

As before, the expression of the *Blaze function* is given by the following function:

$$I_i(\phi_{ce}) = \frac{\sin(\phi_{ce})^2}{\phi_{ce}^2} \quad (4.25)$$

Figure 4.22 clearly illustrates how the *Blaze function* modulates the *Interference function*: their product defines the so called *Intensity function*.



(Figure 4.22: Intensity function)

The graph represents the intensity pattern for a single wavelength, which is the product of the two functions.

Looking at the graph, it can be seen that the peak of the *Blaze function* occurs at ' $\phi_{ce} = 0$ ', which corresponds to zero diffraction order ($m=0$), if plane-parallel facets are considered (like a rectangular shape): this means that the grating has the peak of efficiency along the zero-order.

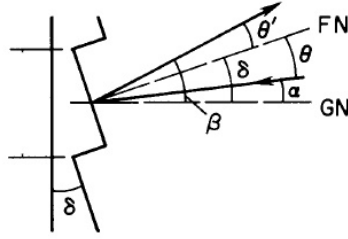
However, having the efficiency peak at zero order is not convenient, especially if polychromatic light is indexed on the grating: since the resolving power is zero at this order, it follows that spectral information are lost: in fact at the zero order, there is no dispersion.

As mentioned before, the problem was solved thanks to Lord Rayleigh: grating must be '*blazed*', so the facets must be made at a certain angle ' δ '.

Thanks to this technique the peak of the *Blaze function* shifts along the x-axis: in this way, the peak of efficiency is centered in the desired diffraction order ' m '.

So by tilting the facets with a certain angle, the phase difference from the center to the edge of the facets varies as the follows:

$$\phi_{ce} = \frac{\pi s (\cos \delta) (\sin \theta - \sin \theta_r)}{\lambda} \quad (4.26)$$



(Figure 4.23: blazed diffraction grating)

where as it can be seen from Figure 4.23:

- ' δ ' is the tilt of the facet;
- ' θ' ' is the angle between the facet normal 'FN' and the incident angle ' α ', and is equal to ' $\alpha - \delta$ ';
- ' θ ' is the angle between the facet normal 'FN' and the diffracted angle ' β ', and is equal to ' $\beta - \delta$ '.

Again, the peak of the *Blaze function*, is reached when ' $\phi_{ce} = 0'$ ': in this situation, this happen when:

$$\theta = \theta'; \quad (4.27)$$

This is called *blaze condition* and it means:

$$\delta - \beta = \alpha - \delta \rightarrow \alpha + \beta = 2\delta \quad (4.28)$$

Based on the observations made, the grating equation is modified as follows:

$$\frac{m\lambda}{d} = \sin(\theta' + \delta) + \sin(\theta + \delta) \quad (4.29)$$

And applying the '*blaze condition*' in which the normal to the facet bisects the angle between the incident and diffracted rays, it follows that the grating equation becomes:

$$\frac{m\lambda_B}{d} = 2 \sin(\delta) \cos(\alpha - \delta) \quad (4.30)$$

where ' λ_B ' is the *blazed wavelength* for the m-oder.

For $\lambda = \lambda_B$ and $\alpha = \beta$, the *blaze Littrow condition* is reached.

From these analyses, it is clear what was mentioned earlier: if the *blaze angle* varies, the *blaze wavelength* also varies, so the peak efficiency shifts.

It can therefore be concluded that the useful effect of *blazing* a grating is to ensure that the efficiency is maximised under specular reflection conditions on the ruling facets.

4.6.3.2: Blazed holographic gratings

As mentioned above, in order to create a blazed grating using the ruling machine, it is sufficient to change the shape of the diamond tooth which rules the surface. As far as holographic gratings are concerned, things are different.

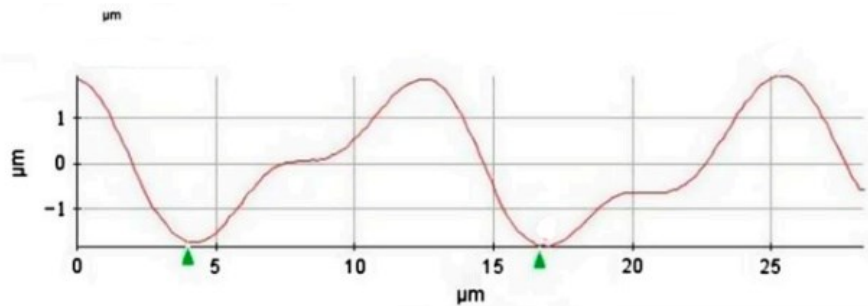
In order to realize a holographic blazed gratings, the technique used is called '*ion beam etching*' (IBE). It is a combination of holographic lithography and IBE.

Briefly resuming the holographic process: a layer of photoresist material is applied to the surface substrate; this is then shaped through an interference pattern thus creating a sinusoidal profile on the grating.

Then the grating is etched with an ion beam (it can be used Argon or CF_4) so to give a pseudo-triangular shapes to the grooves: the ion-beam removes material at some part of the profile, while other parts are hardly etched (Figure 4.24).

With this method 4° - 27° of blazed angle can be achieved.

The main advantages of realized blazed gratings by this way are: uniform grooves' pattern, absence of ghost line.



(Figure 4.24: blazed holographic grating)

CHAPTER 5

CHARACTERISATION OF THE GRATING:

This chapter deals with part of the thesis work, namely the *characterisation of the concave grating of HYPSONS*. The characterisation of the grating was carried out in two ways: a preliminary characterisation through the Atomic Force Microscope (AFM) to evaluate the surface quality of the grating, while the second concerned the analysis of the absolute efficiency of the grating, carried out through the *Johnson-Onaka monochromator*. Both analyses were performed at the LUXOR laboratory of the Institute of Photonics and Nanotechnologies in Padua.

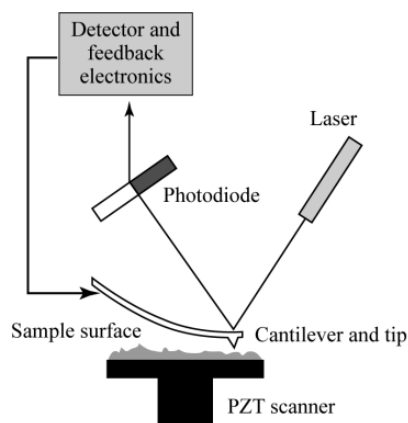
5.1: Characterisation with AFM

5.1.1: Elements and operating principle of AFM

Let's start with the analysis carried out by the AFM: this device allows to obtain images of the sample under consideration in order to study its morphology, so to verify the quality of the surface, i.e. the regularity between the various grooves in terms of height and distance, as a defect in these can cause a drop in efficiency as seen above.

The experimental set-up consists of the following elements as can be seen in the Figure 5.1:

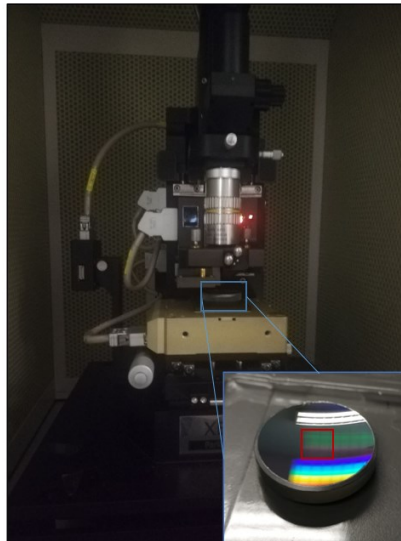
- laser;
- cantiliver and tip;
- sample for analysis;
- photodiode;
- electronics;
- there is also a feedback mechanisms (not shown in Figure 5.1), which maintains the force constant between the sample and the tip;



(Figure 5.1: AFM set-up)

In this case, the operating principle of the AFM is based on the inter-atomic interaction between the AFM tip and the sample surface: a laser beam strikes the back of the cantiliver and the reflections induced by a change in frequency and amplitude of the cantiliver tip are sent to a photodiode. As the tip scans the sample, it moves up and down following the profile of the surface; the change in its position (vertically), changes the intensity of the light that will strike the photodiode: it is through this difference in intensity that the image

can be reconstructed; in fact, from the photodiode, the signal is converted into an electrical impulse and sent to a control logic that converts the signal into a photo of the part of the sample being analysed: Figure 5.2 shows the *AFM* in the LUXOR laboratory in Padua which has been used to test the grating.



(Figure 5.2: *AFM* in the LUXOR)

As mentioned above, measurements were made in a 'non-contact' mode, where the cantiliver tip could 'float' between 5-14 nm relative to the grating surface: this method uses Van der Waal forces between the tip and the surface to provide topographic images.

5.1.2: *Results and Discussion:*

The following figures show the images obtained with the *AFM*.

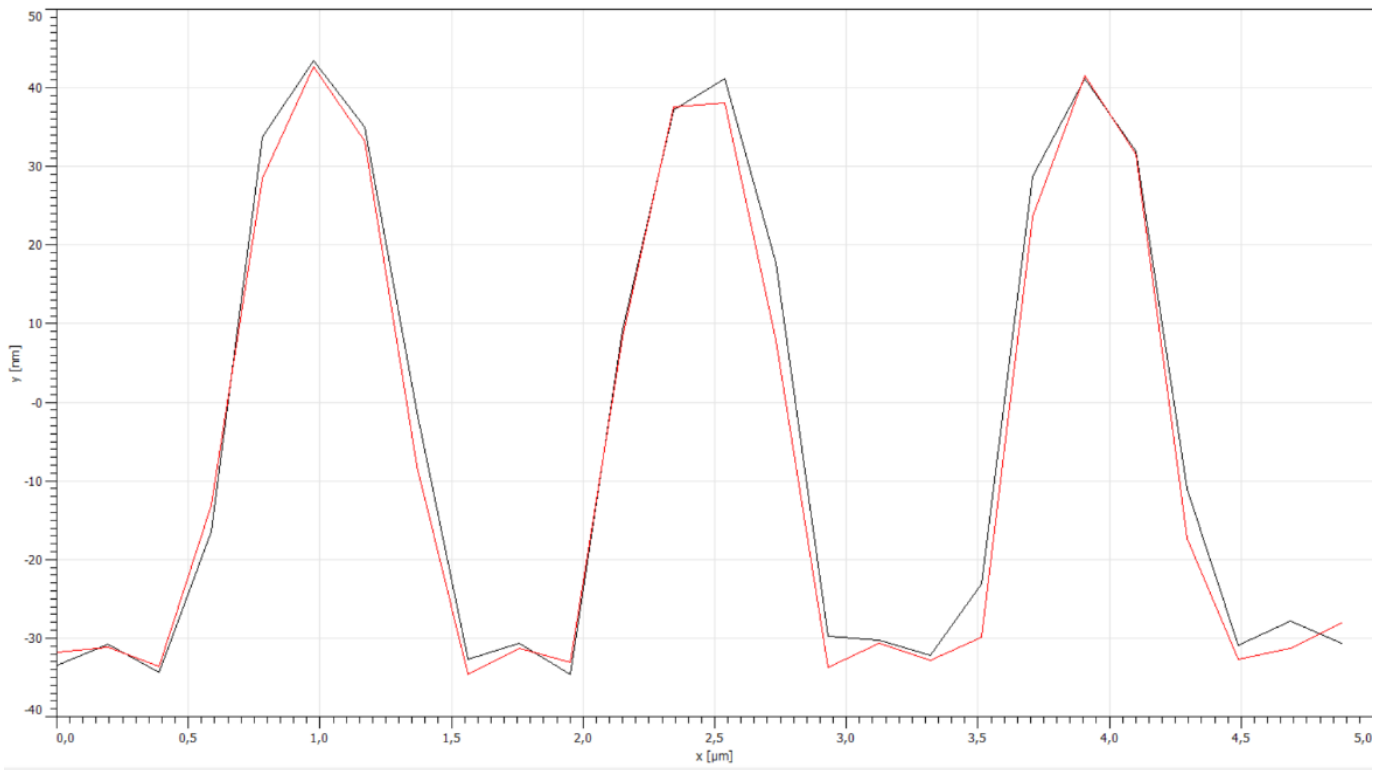
What can be seen is that the portion of the grating surface analysed is almost uniform (Figure 5.3A and 5.3B), so the surface quality is good.

In the first plot the values of the graphed curves are 5 microns apart in height (along the 'y' axis'). The figure aims to highlight the level of uniformity of the various grooves in different areas of the grating.

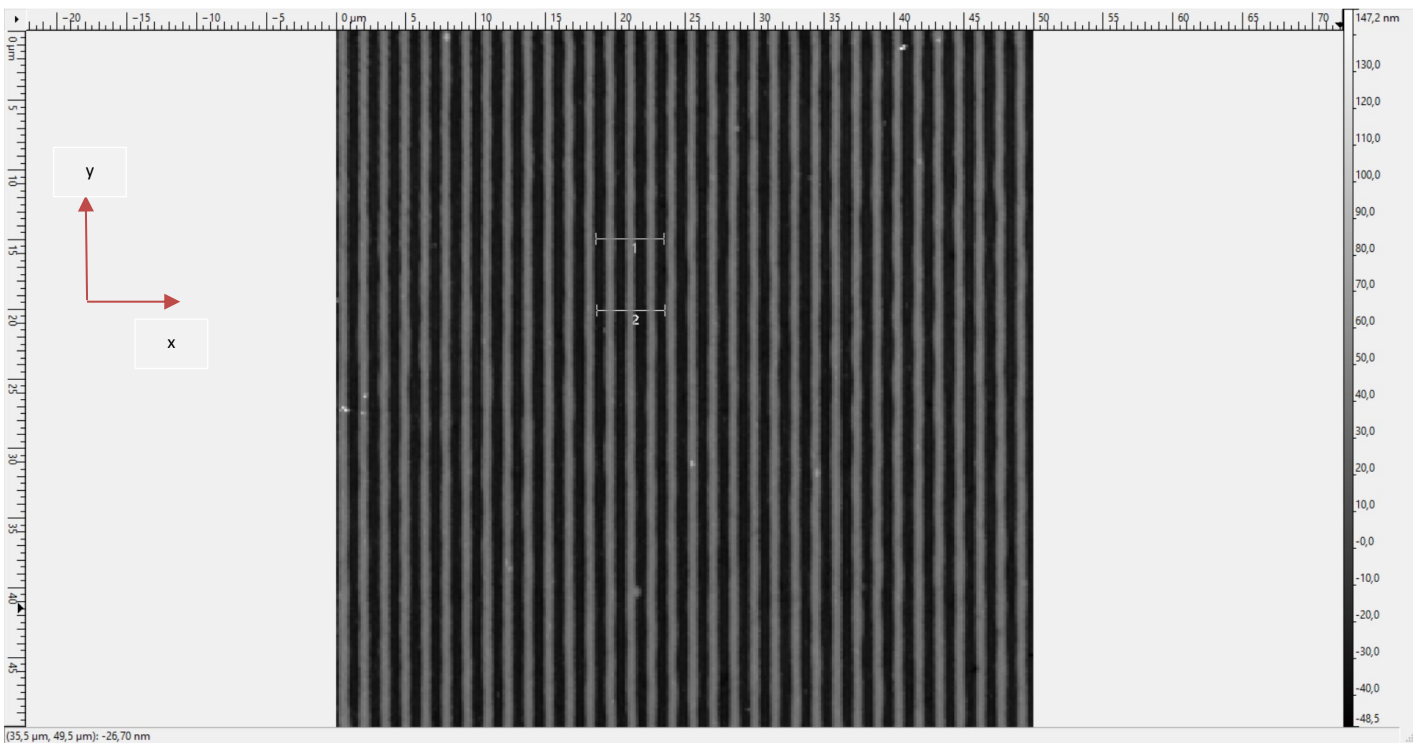
The second graph shows the portion of the grating which have been analysed, while the third graph (Figure 5.4) shows the distance measurements in different areas of the grid. It can be seen that the various quantities (x,y) differ little from each other.

From this preliminary analysis it is expected that the grating has no deficit in efficiency and that the efficiency is the same (at the same order) whether it is illuminated from one side or the other.

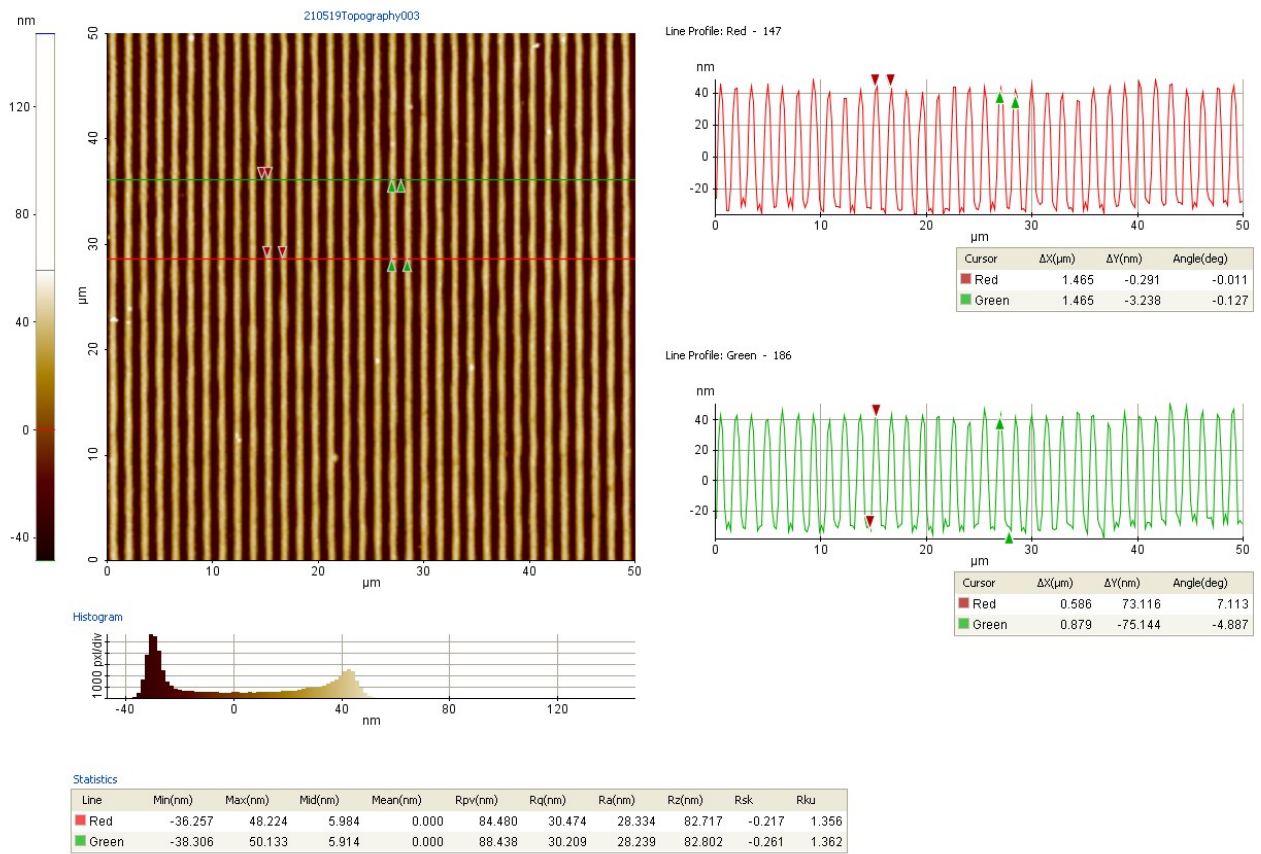
The information regarding the morphology of the grating was obtained through the *Guidoin* software.



(Figure 5.3A: grooves' profile)



(Figure 5.3B: portion of the surface analysed)



(Figure 5.4: surface quality)

5.2: Characterisation of the diffraction grating

In this section the analysis of the grating absolute efficiency for the first order of diffraction is described. To test the efficiency of the grating two inversely polarized photodiodes have been used inside the *Johnson-Onaka monochromator*.

The grating has been tested in three different configurations:

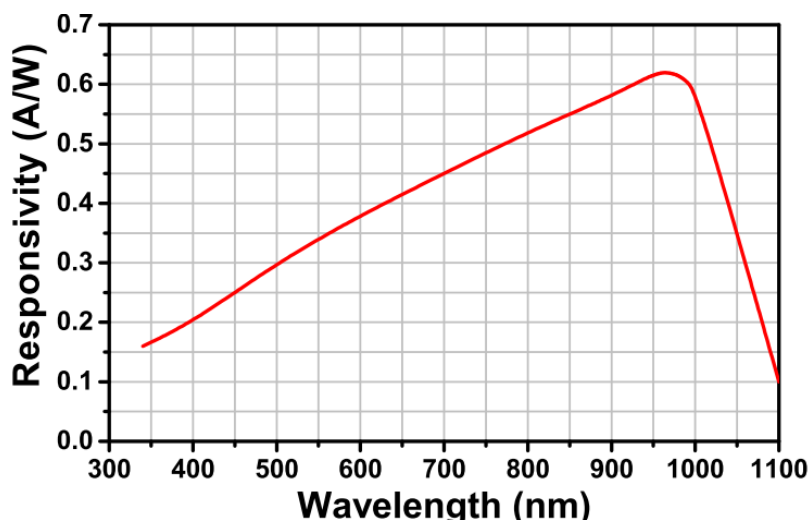
- with the angle of incidence equal to zero: $\theta_i = 0^\circ$;
- Near-Littrow configuration to validate the datasheet calibration curve;
- with the angle of incidence: $\theta_i = 40^\circ$ (HYPSON like);

Before the study of the grating, the characterisation of the photodiodes must be done, because it will be these two devices that will allow the measurement of the direct and reflected beam: the aim is to check that the photodiodes have a linear behaviour as the light intensity changes: these two devices will work in reverse bias mode, so a bias voltage has been imposed: the system will work in photoconductive mode.

5.2.1: Characterisation of the photodiodes

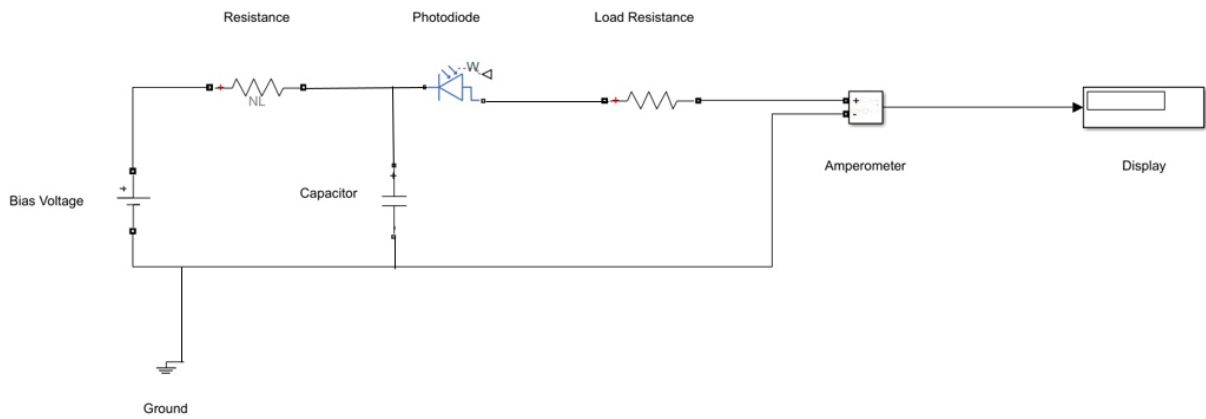
The main characteristics of the photodiodes purchased through *THORLABS* are shown in the following table, while Figure 5.5 under the table shows the 'Responsivity Curve':

Wavelength range	340-1100 nm
Peak Wavelength	960 nm
Responsivity	0.62 A/W
Active area	100 mm ²
Dark Current (5V)	200 pA
Max Bias (Reverse) Voltage	5V



(Figure 5.5: Responsivity curve)

In order to characterise the two components, an electric circuit has been built following the directions of THORLABS (Figure 5.6A):



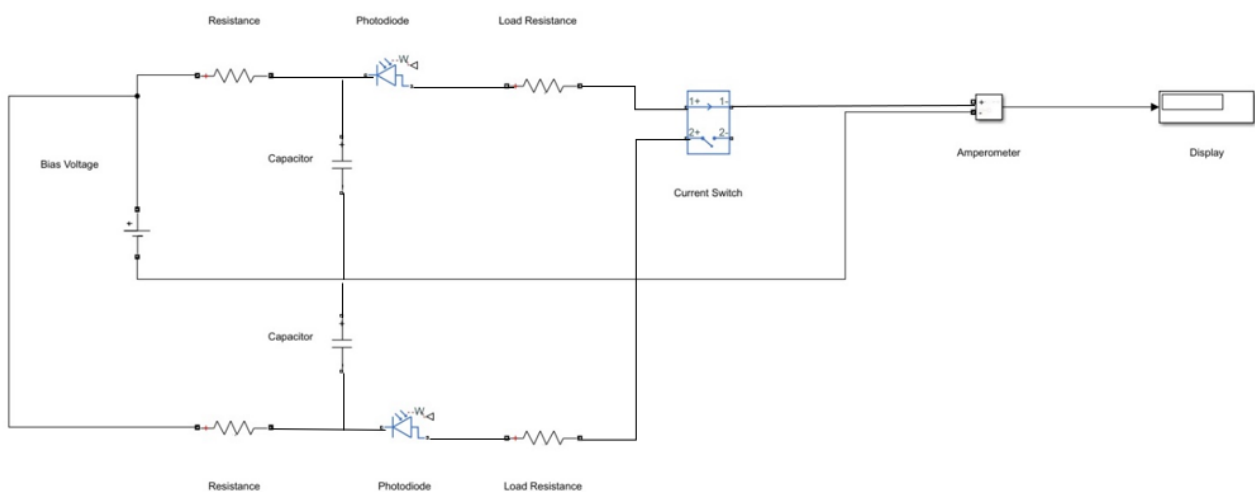
(Figure 5.6A: electric circuit)

Where:

Resistance	1 k Ω
Capacitor	100 nF
Load Resistance	47 Ω

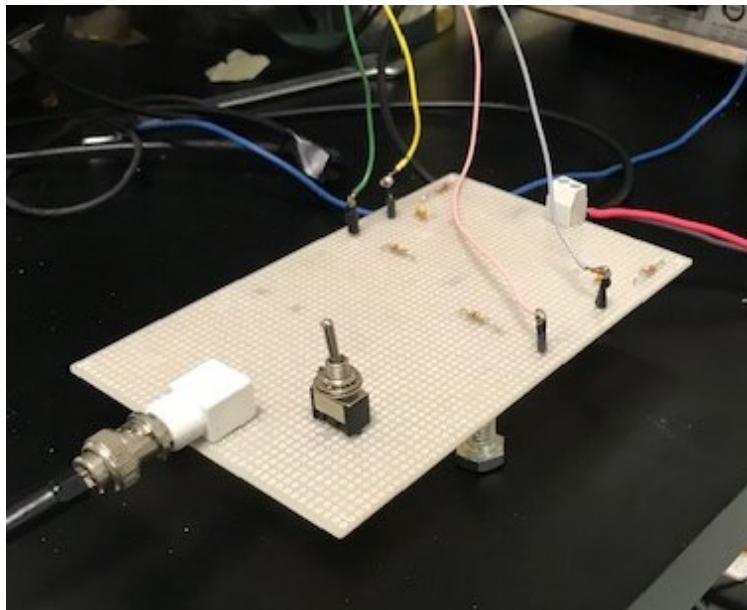
The previous circuit has been modified from the datasheet to allow current readings and not voltage readings.

Finally, since the previous circuit was for a single photodiode, and two photodiodes were needed for the efficiency measurement, the final circuit is as follows (Figure 5.6B):



(Figure 5.6B: complete electric circuit)

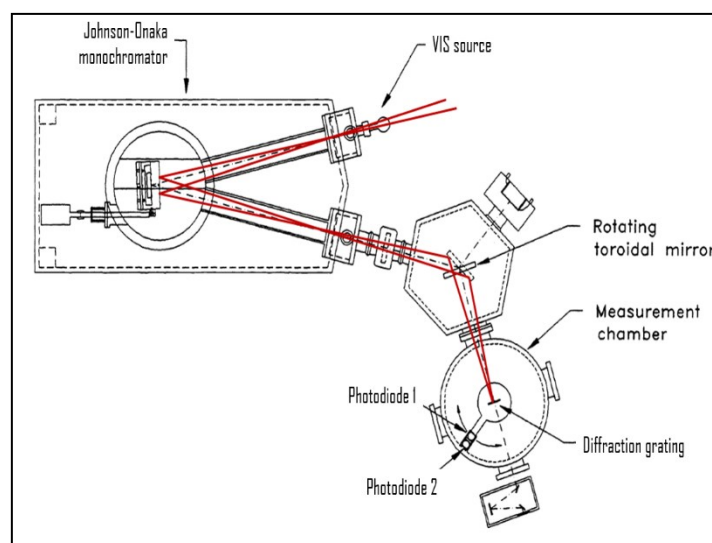
The Figure 5.7 shows the circuit realised with tin soldering:



(Figure 5.7: electric circuit)

As mentioned before, the two photodiodes have been tested inside the *Johnson-Onaka monochromator* (Figure 5.8), which consists, as can be seen from the figure, of:

- an entrance slit;
- toroidal grating of 600 *grooves/millimeters*;
- exit slit;
- toroidal mirror;



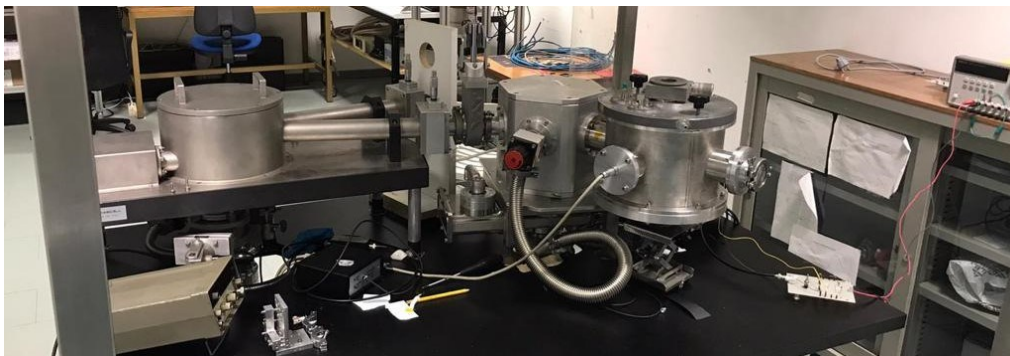
(Figure 5.8: Johnson Onaka monochromator)

In this situation the monochromator is set to $\lambda = 0$ nm, so it behaves like a mirror and white light reaches the two photodiodes.

Inside the measurement chamber there are, in this case, only the two photodiodes; they are placed on a rotator which allows a full 360° rotation. For this scope only 180° of rotation is need (see previous figure). The following figures (Figure 5.9-5.10) show the experimental set-up:



(Figure 5.9: experimental set-up)



(Figure 5.10: experimental set-up)

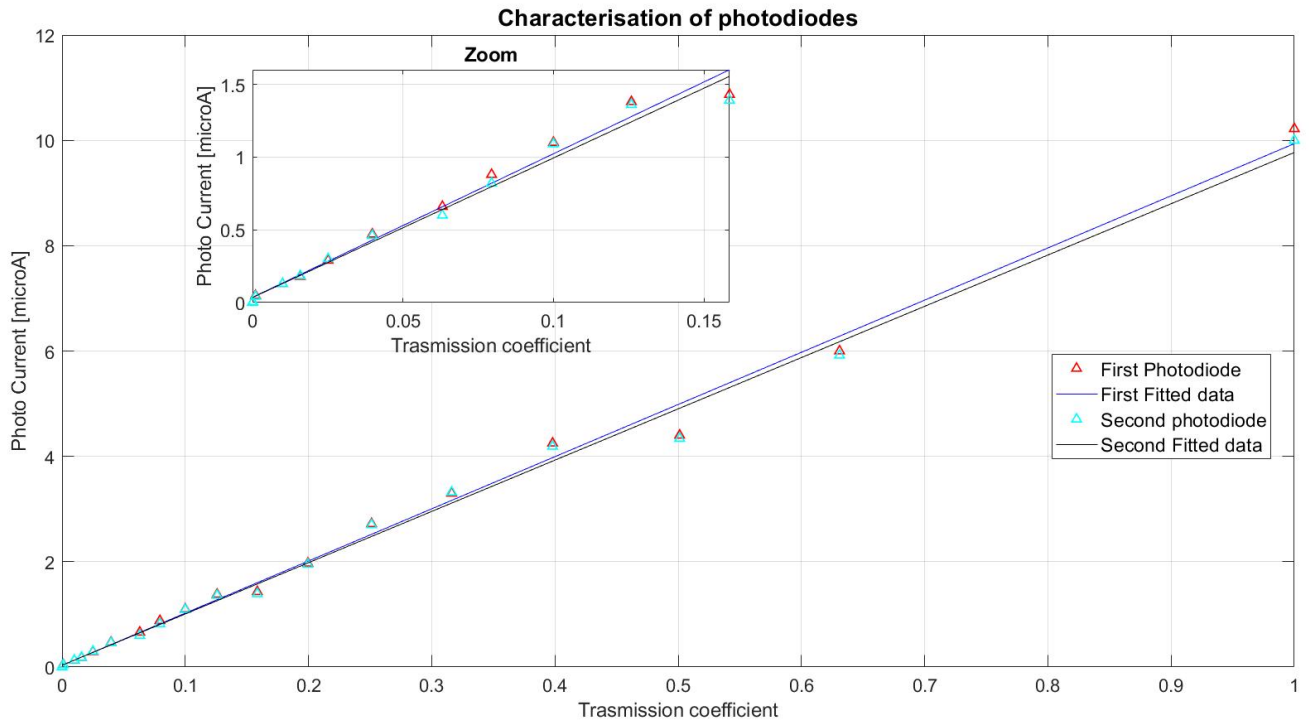
To vary the intensity of the source, in order to study the linearity of photodiodes, neutral filters were used: it has been started with an OD = 0 and ended with an OD = 6, where 'OD' stands for *optical density of the filter*.

To understand the actual transmission index, it is sufficient to apply the following relation:

$$T = 10^{-OD} ;$$

where 'T' stands for transmission index.

Figure 5.11 shows the behaviour of the photodiodes as the light intensity changes. The voltage bias has been set to 3.6 V.



(Figure 5.11: photodiodes linearity)

The following table shows the equation of the two fitted curves:

photodiode	Equation
first	$y = 9.9*x + 0.032$
second	$y = 9.7*x + 0.034$

The graph shows a good linearity of the photodiodes especially at low intensities.

The *Non-Linearity (NL[%])* of the photodiodes was also calculated, i.e. the maximum deviation between the output of the photodiodes and the linear fitting of each one using the following relationship:

$$NL[\%] = 100 \frac{|E_{max}|}{FS},$$

where:

- ' E_{max} ' is the maximum deviation between photodiode output and linear fit;
- ' FS ' is the ' γ ' value of the linear fit.

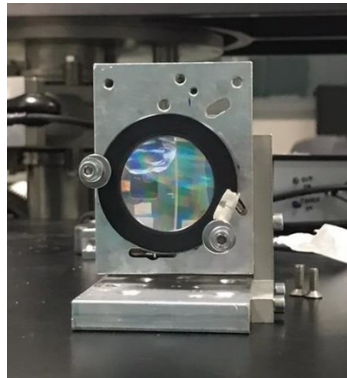
The table below shows the *Non-Linearity* of the two photodiodes:

Photodiode	NL[%]
First Photodiode	11.94
Second Photodiode	11.47

Once the photodiodes have been characterised, the next step was to assemble and analyse the grating.

5.2.2: Characterisation of the Diffraction Grating

As mentioned above, the first step was to fix the grating in a holder (see Figure 5.12); the assembly was done in such a way that the spectrum was projected in a horizontal direction and that the orders were all at the same height so as not to have any 'height difference' problems during the measurement phase with the two photodiodes.



(Figure 5.12: grating on the support)

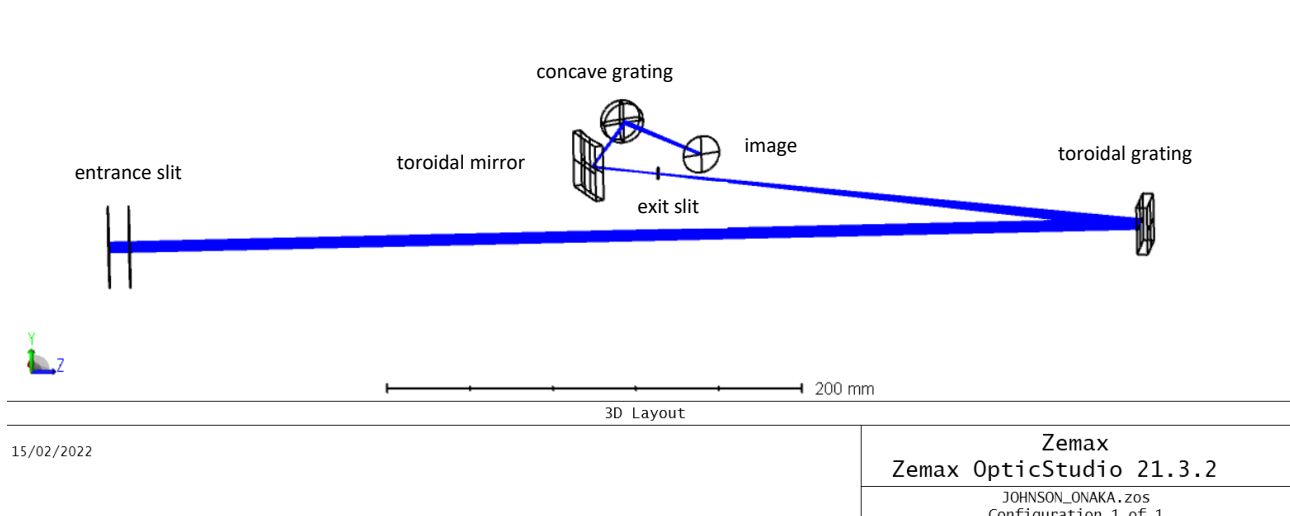
After the previous procedure, the diffraction grating with its support has been mounted inside the facility, precisely in the 'Measurement Chamber', where it rests on a platform that allows both translation and rotation.

In order to understand if the system was well aligned with the source, it was necessary to verify that the beam reflected by the grating in normal incidence was superimposed to the incident beam: an auto-collimation was therefore carried out.

Figure 5.14B shows what has just been described.

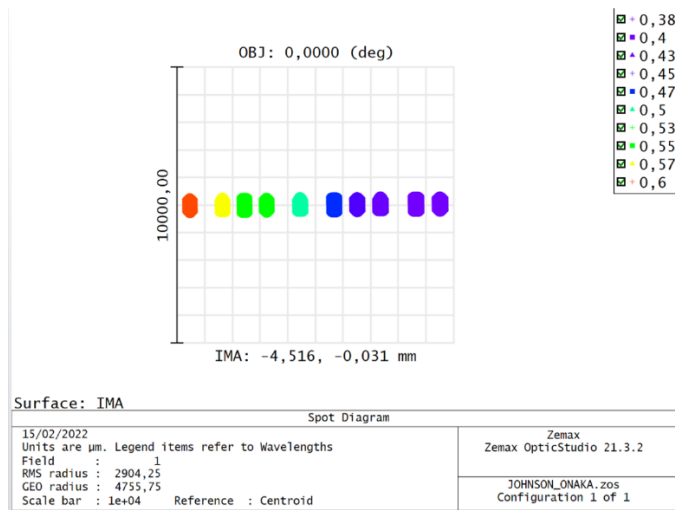
To get an idea of what would happen inside the 'Measurement Chamber', a simulation of the *Johnson-Onaka* was developed using *Zemax* software.

Figure 5.13 shows the *Johnson-Onaka* simulation performed with *Zemax*:



(Figure 5.13: Johnson Onaka simulated with Zemax)

The next two figures (5.14A and 5.14B) show the comparison between the spectrum projected in the 'Measurement Chamber' and the spectrum obtained by *Zemax*:



(Figure 5.14A: spectrum simulated via Zemax)

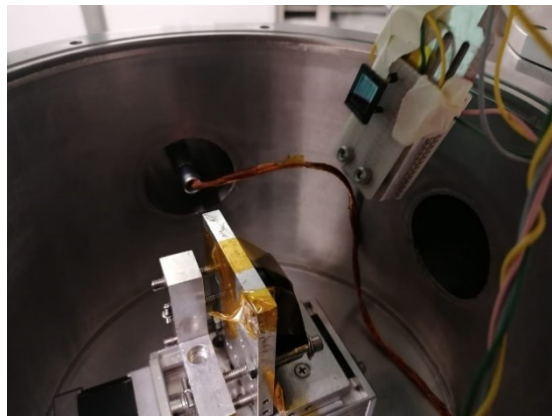


(Figure 5.14B: spectrum in the Measurement Chamber)

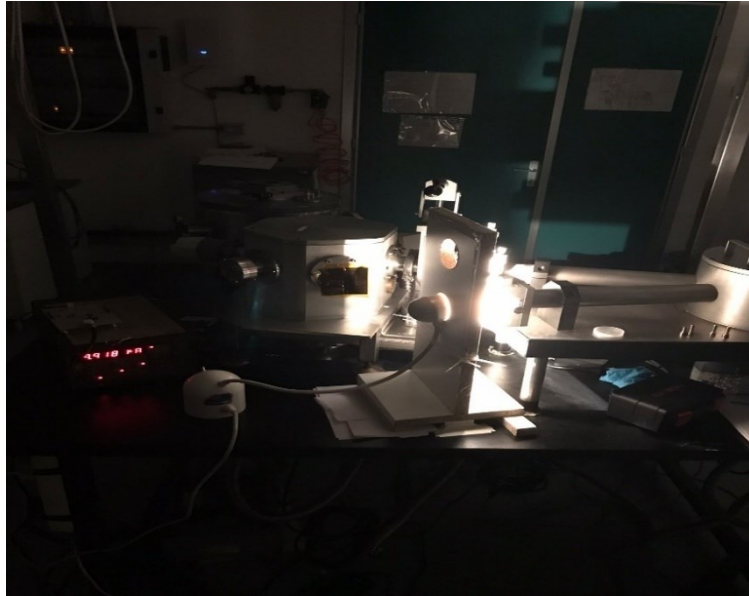
In the previous figure (5.14B) it can be notice the first ($m = \pm 1$) order of diffraction.

After this was done, the grating was moved back a few millimetres so that the second photodiode was arranged along Rowland's circle (as explained above).

The following figures (5.15 and 5.16) show the grating and photodiodes inside the 'Measurement Chamber' and the Measurement Setup:



(Figure5. 15: grating and photodiodes in the Measurement Chamber)



(Figure 5.16: Measurement Set-up)

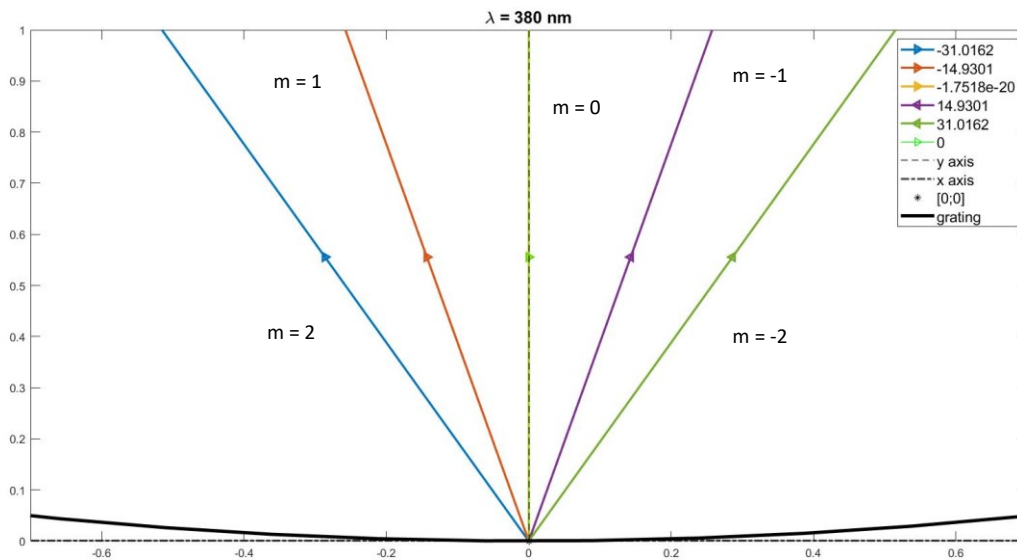
With everything set up, the next step was to study the efficiency of the grating in various configurations.

5.2.2.1: Characterisation of the Diffraction Grating: zero angle of incidence

In this first situation, the efficiency of the grating was studied in the configuration where the angle between the normal of the grating and the source is zero: $\theta_{\text{incidence}} = 0^\circ$.

MATLAB software was used to identify the angles diffracted by the grating.

Figure 5.17 illustrates the simulation carried out in *MATLAB*: in the graph, the $m = \{\pm 2, \pm 1, 0\}$ orders, for $\lambda = 380$ nm can be seen, as well as the incident beam:



(Figure 17: grating simulation)

Once the diffraction angles were obtained from the grating equation, the rotator on which the photodiodes lie was rotated by an angle equal to that previously obtained: this was to ensure that the second photodiode was in correspondence with the diffracted beam.

At this stage the monochromator is no longer set at $\lambda = 0$, so white light no longer reaches the photodiodes, but is rotated so that the appropriate wavelength impacts the grating.

The procedure performed to determine the efficiency of the grating in this configuration is as follows:

- rotation of the toroidal grating to select the appropriate wavelength;
- dark current measurement of both photodiodes;
- incident beam measurement;
- diffracted beam measurement;
- compute of the efficiency;

Consequently, the efficiency formula mentioned above becomes the following:

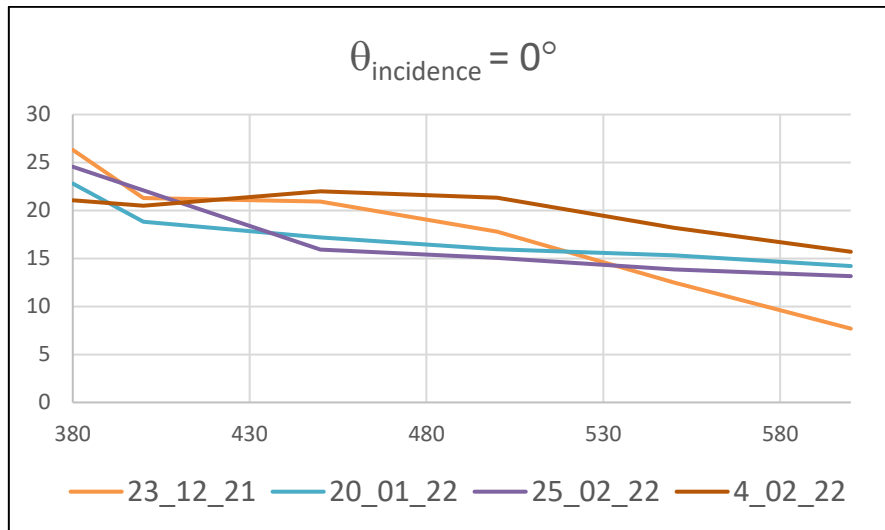
$$\eta(\lambda, m) = \frac{I_d^m(\lambda) - I_{\text{dark}_1}}{I_i(\lambda) - I_{\text{dark}_2}}$$

where:

- I_{dark_1} is the background current of to the photodiode which measures the light intensity diffracted by the grating;
- I_{dark_2} is the background current to the photodiode which measures the incident light;

- $I_d^m(\lambda)$ is the current generated by the beam diffracted by the grating for the first order of diffraction and for a given wavelength;
- $I_i(\lambda)$ is the current generated by the incident beam for a given wavelength.

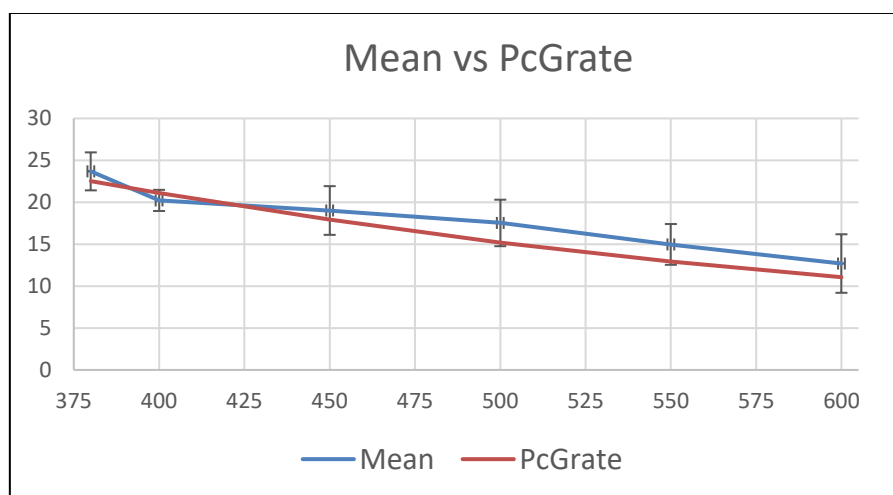
The following graphs (5.18 and 5.19) show the efficiencies measured on different days:



(Figure 5.18: grating efficiency in different days)

The graph below (Figure 5.19) shows the average of the measurements compared with an efficiency curve simulated with the *PcGrate* software.

The purpose of using *PcGrate* was to interpret the results and determine if the measurement was reasonable: within the software, to try to make the simulation as truthful as possible, the profile of the grating simulated is that obtained through the *Atomic Force Microscope*:



(Figure 5.19: Measurement vs Simulated)

Observing the graph above, it can be noticed that the curve obtained by the simulation is inside the error bars of the curve obtained by averaging the measurements shown in Figure 5.18: so there is a good correspondence between what is measured and what is simulated.

The maximum deviation of the two curves is about 15.3%.

5.2.2.2: Characterisation of the Diffraction Grating: Near-Littrow configuration

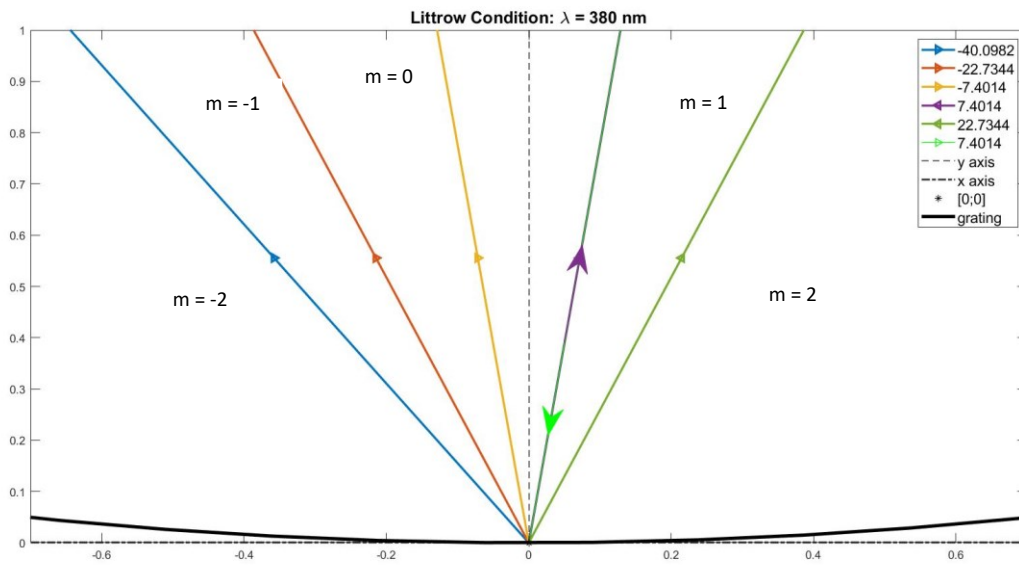
It is now analysed the grating in the Littrow configuration in order to compare this measurement with calibration curve as in the datasheet. Actually, the grating was studied in a *Near-Littrow* configuration because, due to mechanical constrains, measurement in the Littrow configuration was impossible. What was done was to rotate the grating each time by an angle ' θ ' equal to the following relationship,

$$\theta = \sin^{-1} \frac{m\lambda}{2d} + \text{cost} ,$$

as obtained from the Littrow condition, in which the angle of incidence and the angle of diffraction are equal (as seen before).

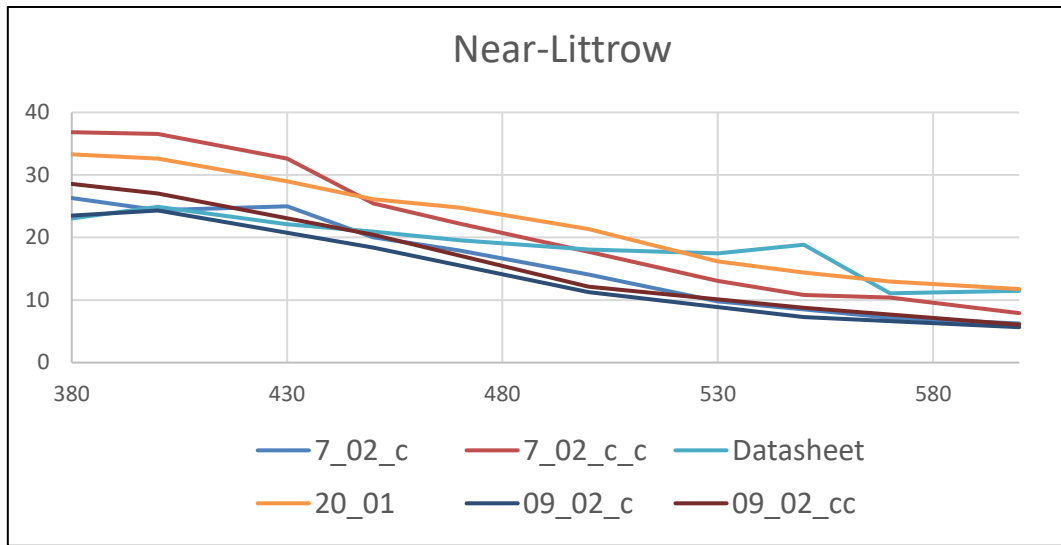
A constant angle of 6° was added to the rotation of the grating in order to obtain a Near-Littrow configuration; similarly, the photodiode is also rotated by the same angular amount.

The angles of the Littrow configuration have always been derived from *MATLAB* (Figure 5.20), where it can be seen that the first-order angle of diffraction and the incident angle are superimposed:



(Figure 5.20: Near-Littrow simulation in MATLAB)

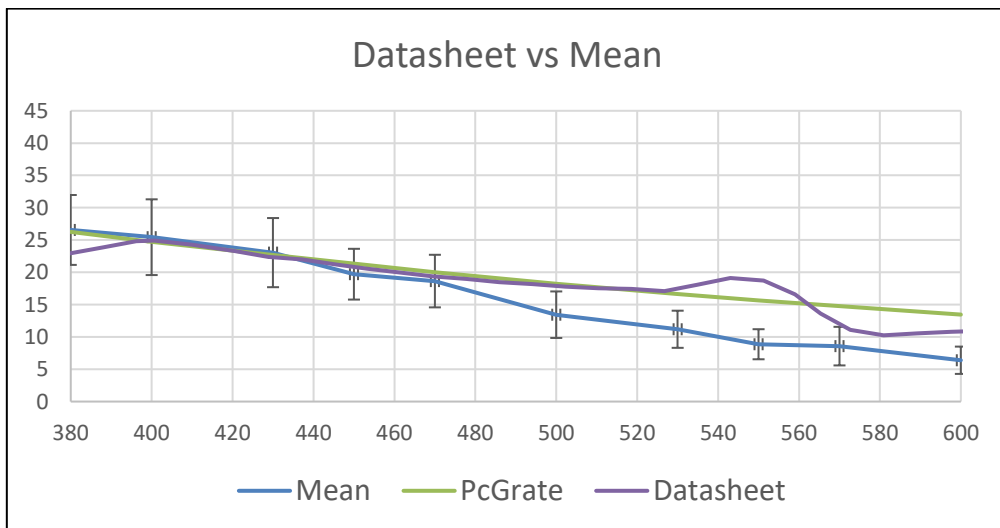
The following plots (5.21 and 5.22) show the efficiencies measured on different days:



(Figure 5.21: grating efficiency in different days)

In the legend of the previous graph, the '_c' means that the grating has been rotated clockwise with respect to the incident beam, while '_c_c' means counter-clockwise: the aim was to verify that by changing the sign of the angle of incidence the performance of the grating remained almost unchanged. This is dictated by the fact that the grating is holographic with an approximately sinusoidal groove profile (Figure 5.3A).

The next plot shows the mean of the measurements versus the datasheet with the error-bars:



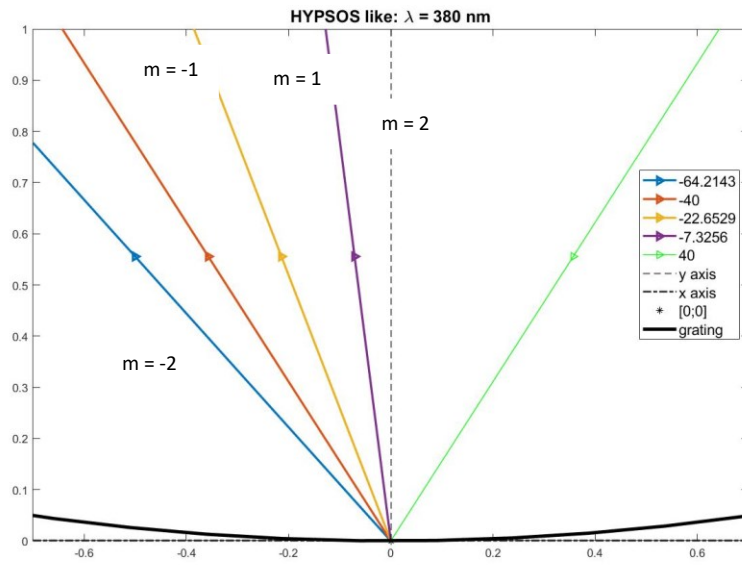
(Figure 5.22: Measurement vs Simulated)

In this situation we are not always within the error bars. In particular, it can be seen that in the datasheet curve between $\lambda=[540-550]$ nm there is an increase in efficiency, which never occurs in the measurements shown in the present graph. The grating may therefore have some defects, as in the range between 540-560 nm the efficiency is always below the nominal one.

Similarly, *PcGrate* does not take this change in trend into account either.

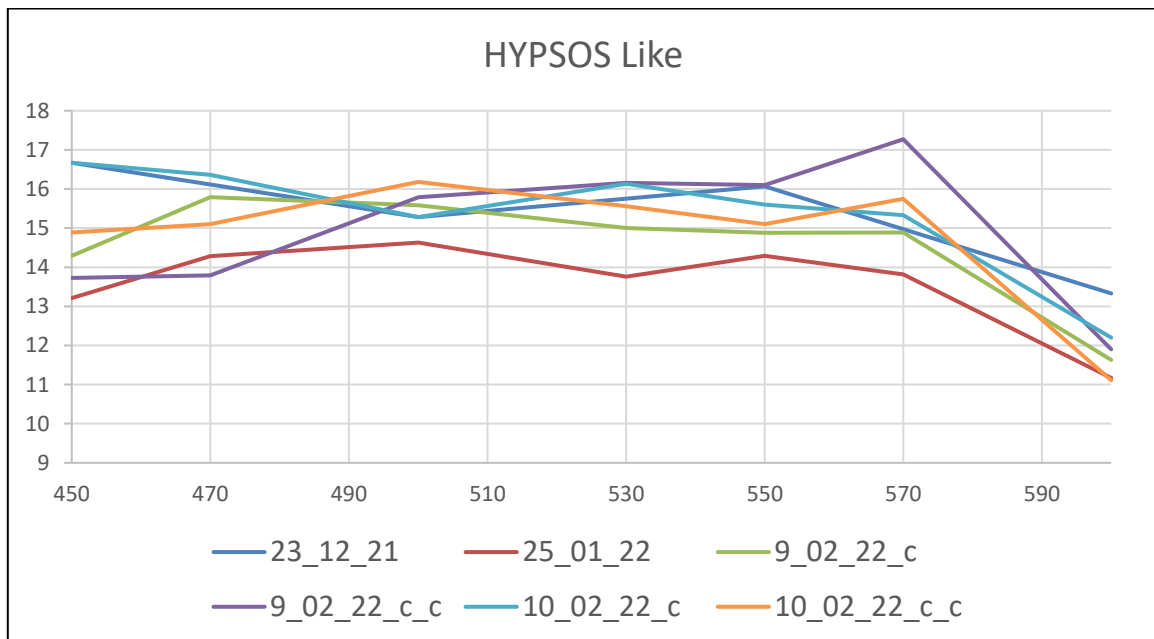
5.2.2.3: Characterisation of the Diffraction Grating: HYPSSOS like

In this last section, the analysis of the efficiency of the grating in the *HYPSSOS* operating configuration, i.e. with a source angle of about 40° (verified by ray tracing), will be discussed. The diffraction angles were also obtained using *MATLAB* (Figure 5.23):



(Figure 5.23: HYPSSOS like configuration simulated in *MATLAB*)

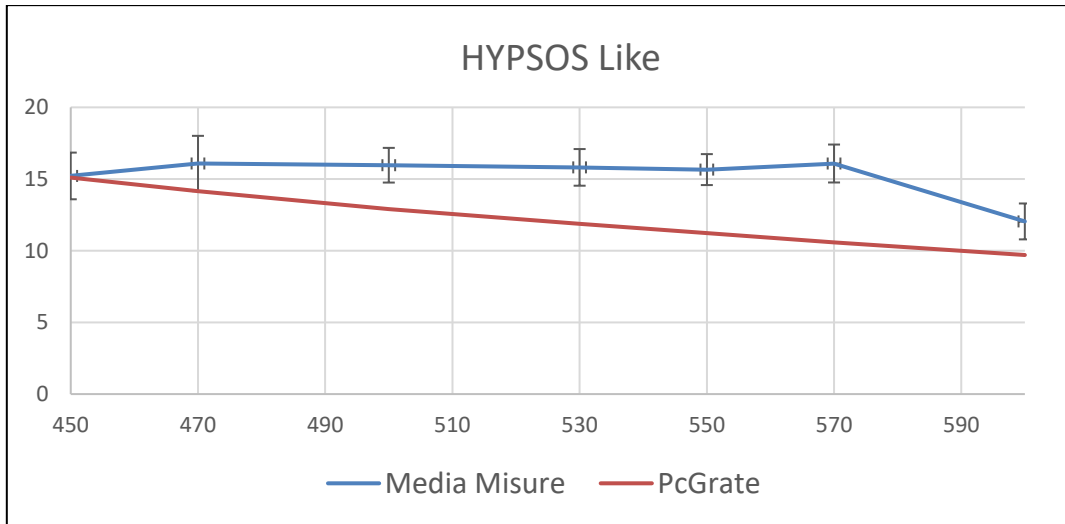
The following plots (Figure 5.24 and 5.25) show the efficiencies measured on different days:



(Figure 5.24: grating efficiency in different days)

As before, in the legend of the previous graph, the '_c' means that the grating has been rotated clockwise with respect to the incident beam, while '_c_c' means counter-clockwise: it can be seen that the measurements taken on 10_02_22 differ little from each other, except for those of 9_02_22, where there is a marked deviation at $\lambda = 570\text{nm}$: This may be due to not having waited long enough before taking the measurement, i.e. the photodiode had not yet stabilised.

The next plot (Figure 5.25) shows the mean of the measurements with the error bars versus the results obtained simulating this configuration with *PcGrate*:



(Figure 5.25: Measurement vs Simulated)

In this case the simulation differs significantly from the measurement, being largely out of the error bars, which means that the simulation done with *PcGrate* is not so reliable.

In contrast to the two previous configurations, here the range of wavelengths analysed is from 450 to 600 nm; this is because at λ less than 450, the measured current values were close to the dark current, providing a signal to noise close to 1, so not reliable enough.

5.2.2.4: Characterisation of the Diffraction Grating: conclusion

Starting from the configuration with the angle of the source null with respect to the grating normal, it was noted that all measurements fall within the error bars.

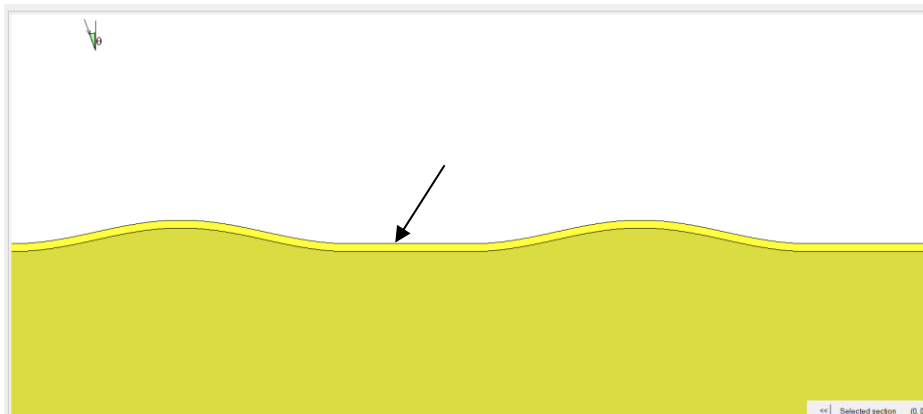
As far as the Near-Littrow configuration is concerned, however, not all the measurements fall within the error bars, and the measurements taken, never show an increase in efficiency in the area between $\lambda = [540; 550]$ nm, as one would expect from the datasheet.

Moreover, even in this configuration there is a deviation between the curve simulated with *PcGrate* and the average curve of the various measurements.

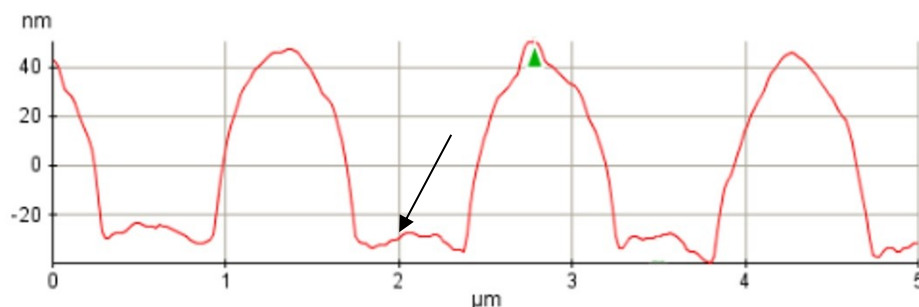
On the other hand, comparing the *PcGrate* curve with that of the datasheet, it can be seen that the former does not show an increase in efficiency in the area between $\lambda = [540; 550]$ nm, but always has a decreasing trend: this means that the simulation conducted with *PcGrate* is not so reliable.

As far as the HYPPOS like configuration is concerned, the situation here is completely different; in fact, the simulated results are completely outside of the error bars: this means that the simulation conducted with *PcGrate* is not reliable.

The substantial difference between what is measured and what is simulated therefore lies in the profile inserted into *PcGrate*: despite the fact that the grating has as its parameters those obtained from *AFM*, the profile of the grooves is merely sinusoidal with a flat area between one groove and another, while the real one is not entirely: just look at the figures below (Figure 5.26A and 5.26B, black arrow):



(Figure 5.26A: PcGrate Simulation)



(Figure 5.26B: grooves' profile taken with AFM)

Looking at the three cases, it is easy to see that the measurements taken on different days slightly differ from each other: this is mainly due to the fact that the source used is not calibrated but is fluctuating, so the

intensity of the light emitted varies over time, thus causing those variations in efficiency even between two days and the fact that the environmental conditions were different.

Then the fact of using two photodiodes is not so convenient because they have different dark-current and also slightly different measured current values (as can be seen in the Figure 5.11): in fact, for the same position and lighting conditions (as far as possible) they output slightly different current values. Moreover it may be that during the rotation of these, their previously fixed position varied slightly, due to the tension induced by the cables: mainly, the choice of using both is dictated by space issues.

Finally, although the range of the visible spectrum extends down to around $\lambda=780$ nm, the wavelengths studied go down to $\lambda=600$ nm: this is due to the fact that the monochromator is optimised for working in the UV region. Despite this, from the performed measurements, it is expected that in the operational configuration, the efficiency of the grating will be in the range of 10% - 16%.

CHAPTER 6

SILICON SEMICONDUCTOR DETECTORS

In order to better understand the working principle of semiconductors, it is best to take a step back and explain the theory of energy bands for the various types of materials: conductors, semiconductors and insulators.

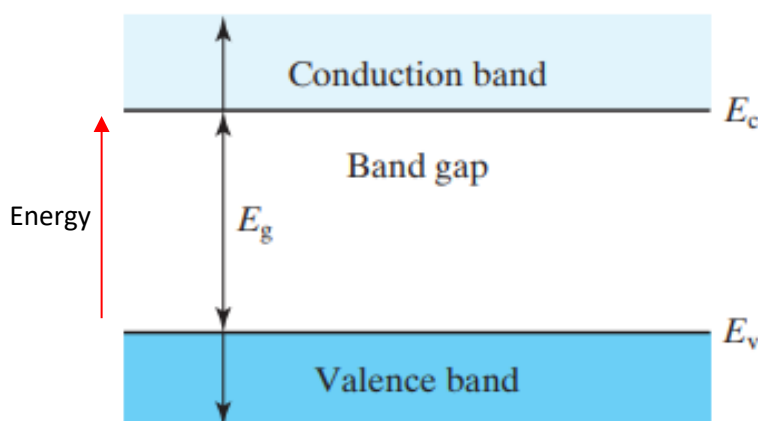
6.1: Materials's energy band

Typically, electrons tend to occupy the lower levels of energy bands: the lower the energy, the richer the electron content. Bands at higher energies, on the other hand, will be 'poor' in electrons, and therefore almost empty. Between the electron-rich and electron-poor bands, there are two particular bands: the first is the valence band, which is the lowest band of the permitted states and is completely full of electrons, especially when the temperature tends to zero Kelvin; this is the band with the highest energy (in absolute value). The second band is the conduction band, which is almost empty of electrons and represents the lowest energy state in absolute value.

The 'jump' between the valence band and the conduction band is called the band gap (Figures 6.1) and it is equal to:

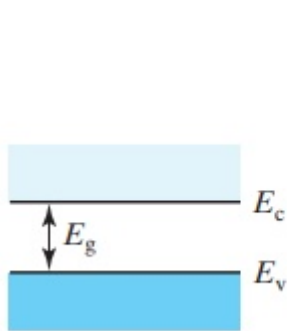
$$E_g = E_c - E_v \quad (6.1)$$

By supplying energy to the material, e.g. by exchanging heat or introducing light, the electrons can absorb this energy, until the energy gap between the two bands is equalled or exceeded. When this happens, the electron leaves the valence band and migrates to the conduction band, where it will be free to move.

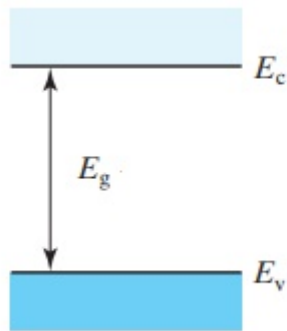


(Figure 6.1: energy band diagram)

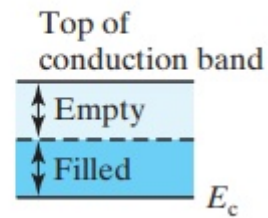
Depending on the energy gap between the two bands, three classes of materials can be identified, which are: semiconductors, insulators and conductors (Figures 6.2A, 6.2B and 6.2C):



(Figure 6.1A: semiconductor)



(Figure 6.1B: insulator)



(Figure 6.1C: conductor)

Starting with semiconductor materials, these are characterised by having a small energy gap between the two bands; in these materials the absorption of energy, if sufficient, allows the electron to pass from the valence band to the conduction band.

Insulating materials are those with a much higher energy gap than the previous ones; in these circumstances, the passage of electrons in the conduction band is not possible. A material is typically considered insulating when the energy gap is equal or greater than 4eV.

Conductive materials, on the other hand, are those in which the two energy bands are superimposed, i.e. the electrons are free to move within the material.

6.2: Photoelectric Effect

In the second chapter it was briefly mentioned that light has an ambiguous behaviour: it has both wave and particle one.

As we saw in Chapter 3, it has a wave-like behaviour because it gives rise to phenomena such as diffraction, while it has a corpuscular one because when it interacts with a material, this one releases electrons.

The connection between the two aspects is given by the following formulas:

$$E = hf \quad (6.2A)$$

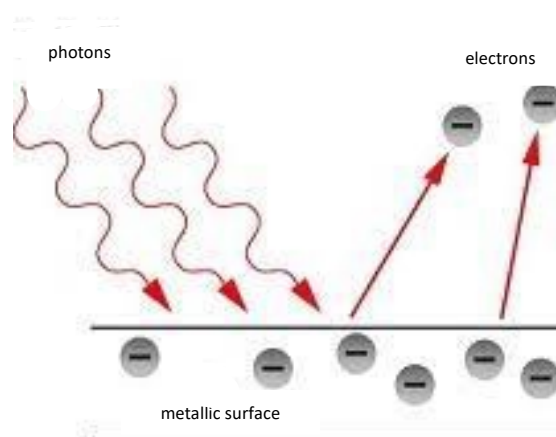
$$\mathbf{p} = \frac{h}{\lambda} \quad (6.2B)$$

Where:

- ' E ' is the energy of the radiation;
- ' \mathbf{p} ' is the momentum, which it is a vector with orientation equal to the direction of propagation of the radiation;
- ' h ' is Planck's constant;
- ' f ' is the frequency of the emitted radiation;
- ' λ ' is the wavelength of the incoming light.

Equation (6.2A) is known as *equation of the photon's energy*: the electromagnetic radiation can only be emitted as discrete energy corpuscles (quanta), which are proportional to the frequency ' f ' of the radiation. The analysis of the corpuscular behaviour of light makes it possible to understand the *photoelectric effect*, which is the basis for understanding *photo-emission detectors*.

The *photoelectric effect* can be described simply as the emission of electrons from a metal surface when it is struck by an electromagnetic radiation of a certain frequency or wavelength (Figure 6.3):



(Figure 6.3: photoelectric effect)

Without going into too much detail, the complete explanation of the photoelectric effect is due to Einstein: he assumed that the energy of the incident radiation was quantized into discrete packets, called photons. Now, if the energy of the incoming radiation is sufficiently high, an electron can be ejected from the surface with a certain kinetic energy (E_k): equation (6.3) provides a better understanding of this aspect:

$$E_K = \frac{1}{2}mv_{max}^2 = hf - W_0 \quad (6.3)$$

Where:

- 'W₀' is the extraction work;
- 'v_{max}' is the maximum speed that the electron can have;
- 'm' is the mass of the electron.

It can be demonstrated that, assuming E_K = 0, the minimum extraction work is given by equation (6.4):

$$W_0 = hf_0 \quad (6.4)$$

where 'f₀' is a different frequency for each metal and constitutes the photoelectric threshold. These means that 'f₀' is the minimum frequency at which electrons can be extracted.

Figure 6.4 shows some values of the work function for differents semiconductor materials:

Material	Work function (eV)
Indium (In)	4.08
Gallium (Ga)	4.35
Silicon (Si)	4.95
Germanium (Ge)	5.15

(Figure 6.4: values of work function. 1eV = 1.609*10⁻¹⁹ J)

What has just been described is known as the *external photoelectric effect*, since an electron, if properly excited, can leave the metal surface with a certain kinetic energy defined by the equation (6.3).

The *external photoelectric effect* can be summarised in four points:

1. *Absorption*: not all photons generate photo-electrons; part of the light is absorbed and reflected.
2. *Excitation*: in the metals electrons are generated; instead in semi-conductor material, electron-hole pairs are generated;
3. *Electronic Transport*: the excited electrons can interact with other electrons; if their energy is above the threshold they can escape;
4. *Extraction*: the electrons with sufficient kinetic energy can leave the surface.

From points 1 and 2, it is clear that not all photons interacting with the material release electrons.

Based on this, *quantum efficiency* is defined as the ratio of generated photoelectrons to incident photons:

$$\eta = 100 * \frac{n_e}{n_{ph}} \quad (6.4)$$

Where:

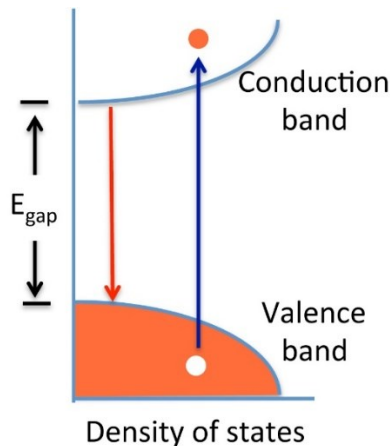
- ' n_e ' is the number of photoelectrons;
- ' n_{ph} ' is the number of the incident photons.

Now let's consider the Rutherford-Bohr model of an atom of Silicon, in which electrons have discrete energy level separated by energy gaps: electrons can acquire sufficient energy, for example by the absorbing of the incoming radiation, so they can escape from the forces that bind them to the nucleus; this is the so called *internal photoelectric effect*. These electrons will be 'free' to move in the Silicon crystal.

With ' E_{gap} ' is defined the energy gap, which is the minimal required energy to make possible the *internal photoelectric effect* and it is less than the work function energy defined in the *external photoelectric effect*: for Silicon, at a temperature of 300 K, the energy gap is 1.1 eV.

The electrons deficiency left in the covalent band may be replaced by adjacent electrons; this movement gave birth of a 'new particle' called *hole*. So when the *internal photoelectric effect* is considered, in which $E_{ph} > E_{gap}$, the creation of a electron-hole pair happens.

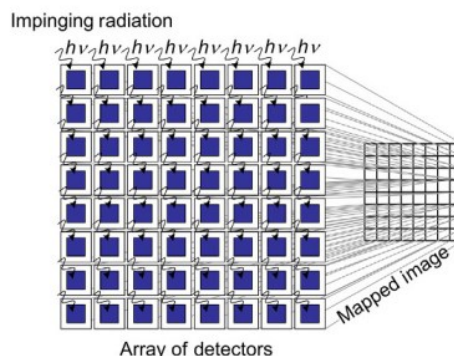
What has just been described explains the behaviour of semiconductors (Figure 6.5):



(Figure 6.5: electron-hole pair generation)

The *photoelectric effect* is used in detectors: by measuring the charge extracted (called *photocharge*), it is possible to determine the level of incident radiation.

Now, if an array of photosensitive elements is illuminated and the information is mapped on a two-dimensional surface, an image will appear (Figure 6.6):



(Figure 6.6: example of imaging principle)

6.3: Charge Coupled Device (CCD)

In this section, *CCDs* will be briefly discussed as these devices have dominated the industrial imaging application since their creation.

As mentioned above, silicon detectors are sensitive to photons of wavelength where electrons can be created via *internal photoelectric effect*. This means that an electron of the valence band, could reach the conduction band if the photon energy is greater than the bandgap of silicon:

$$E_{\lambda} = \frac{hc}{\lambda} > E_{gap} \quad (6.5)$$

Where:

- ' E_{λ} ' is the photon energy of wavelength λ ;

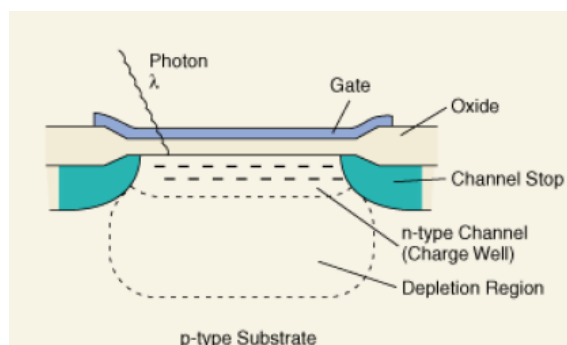
The energy bandgap for silicon ($E_{bandgap}$) is equal to 1.1 eV, so by using equation (6.2A) it is easy to find the cutoff wavelength (λ_{cutoff}) for silicon in which no transmission is possible:

$$\lambda_{cutoff} = \frac{hc}{1.1 \text{ eV}} = 1.12 \mu\text{m} \quad (6.6)$$

6.3.1: CCD architecture and principle of work

A *CCD* (Figure 6.6) consist of an array of pixels (picture elements) that produce potential wells from applied clock signals to store and then transport charge packets, which are electrons generated by the photoelectric effect.

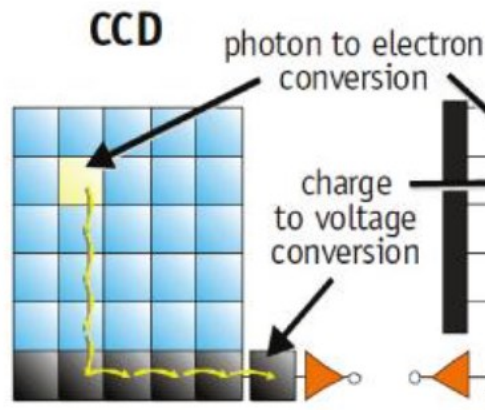
Basically, the structure of a *CCD* is a MOS capacitor (Figure 6.7), which consists of three elements: a metal called gate, a SiO_2 insulator (acting as an oxide) and a silicon substrate.



(Figure 6.7: MOS structure)

When light strikes the MOS capacitor, electron-hole pairs are formed: to prevent recombination, the holes are directed into the substrate and removed just applying a voltage to the gate. What remains are therefore electrons.

In order to move the charge, from one capacitor to another, the electrode potentials must vary: in doing this, the charge moves to the output amplifier, which behaves like a charge to voltage converter; then the electronics convert this output voltage measured per each pixels, into a two-dimensional image (Figure 6.8).

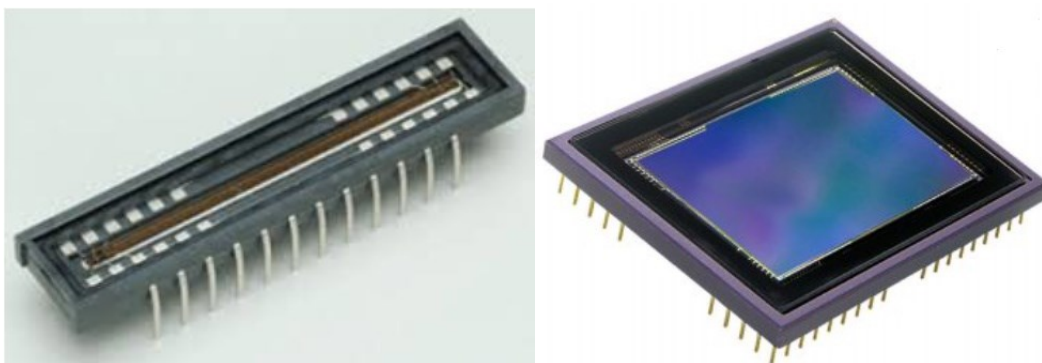


(Figure 6.8: CCD)

The principle of work of the *CCD* can be summarised in four points:

1. generation of the charge via the *internal photoelectric effect*;
2. collection of the charge in the MOS structure by the application of a voltage;
3. charger transfer, by changing the electrode's potential;
4. signal extraction and conversion.

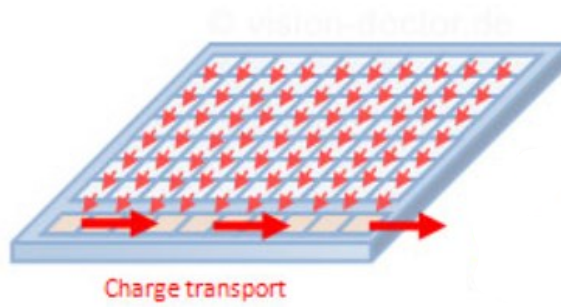
CCDs detector can be divided into two types: linear and area arrays (Figure 6.9). The main difference is that in the area array the surface is a 2D matrix of pixels. Linear arrays are less expensive and has a very fast read-out; on the other hand, area arrays are much more common for scientific applications.



(Figure 6.9: CCD linear and area arrays)

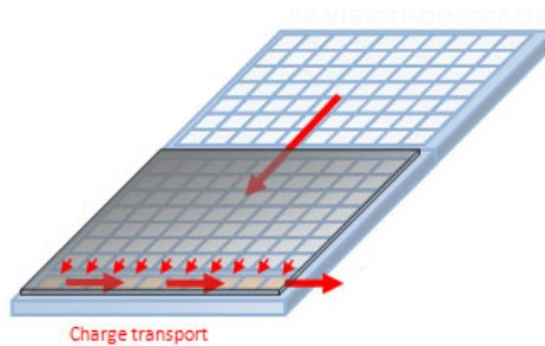
Considering *CCD* area arrays, these are divided into three groups:

1. *full-frame CCD*: which uses all the pixels to collect light. All the charges are then transferred at the bottom (Figure 6.10): the process is 'slow' and some of the charge may be lost; however the charge transfer efficiency is about 99.999%;



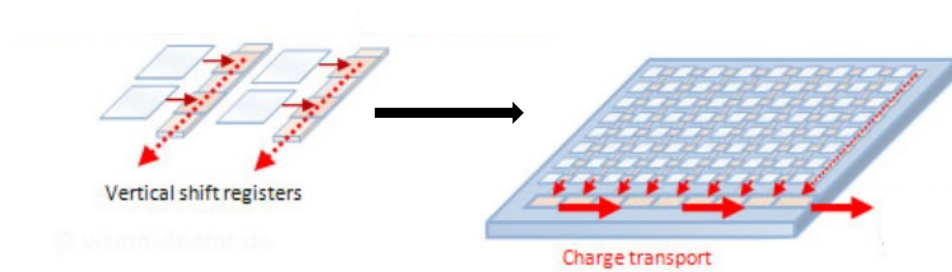
(Figure 6.10: full-frame *CCD*)

2. *frame transfer CCD* (Figure 6.11): the *CCD*'s area is divided into two identical zones, which are the image area and the memory area, which it is covered by an opaque mask. The transfer speed is higher than full frame;



(Figure 6.11: frame transfer *CCD*)

3. *interline transfer CCD* (Figure 6.12): on the side of each pixel column there is an opaque shift register, so a very rapid transfer can be made. However those type of *CCDs* are more complex to realize and need additional electronics, the fill factor (defined as the ratio of the sensitive area to the light over the entire pixel area) is lower with respect to the two previous types.

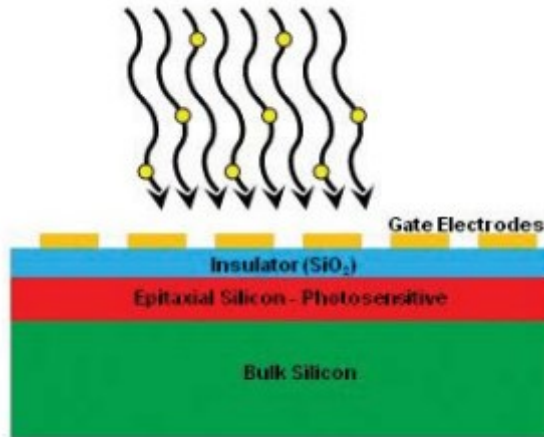


(Figure 6.12: interline transfer *CCD*)

6.3.2: CCD illumination mode

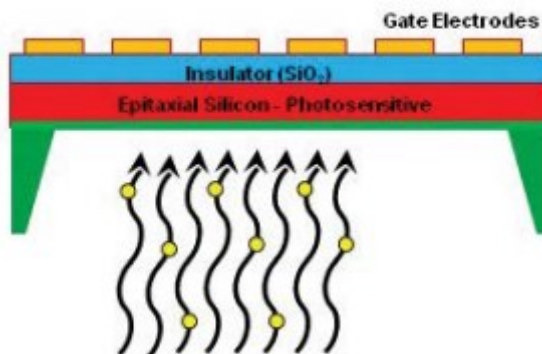
There are two ways in which *CCDs* can be illuminated: they can be illuminated from the front or from the rear.

Starting from the front illuminated one (Figure 6.13), they are cheaper than the back illuminated one; however those type of *CCDs* are not so useful at wavelengths < 400 nm, because the radiation is absorbed by the front side gates.



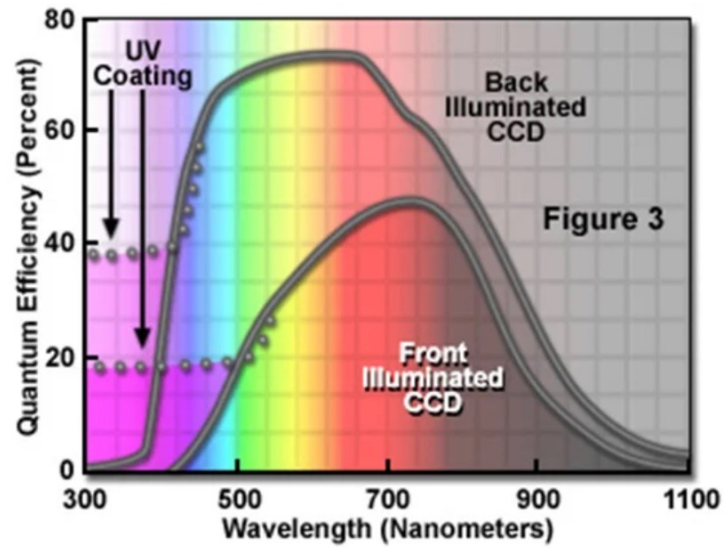
(Figure 6.13: front illuminated *CCD*)

Instead the back illuminated *CCDs* (Figure 6.14) are much expensive, but have efficiency higher respect to the front illuminated.



(Figure 6.14: back illuminated *CCD*)

Figure 6.15 shows an example on how the quantum efficiency varies with the two type of *CCDs*:



(Figure 6.15: difference in quantum efficiency)

To conclude: *CCDs* are linear passive devices; the charge generated by the arrival of photons is transferred from the pixels to the output amplifier and their readout mode is performed pixel by pixel.

6.4: Active Pixels Sensor Complementary Metal-Oxide-Semiconductor (APS-CMOS)

As in the *CCDs* detector, also *APS CMOS* are realized in silicon and principle of photocharge production is based on the *internal photoelectric effect*.

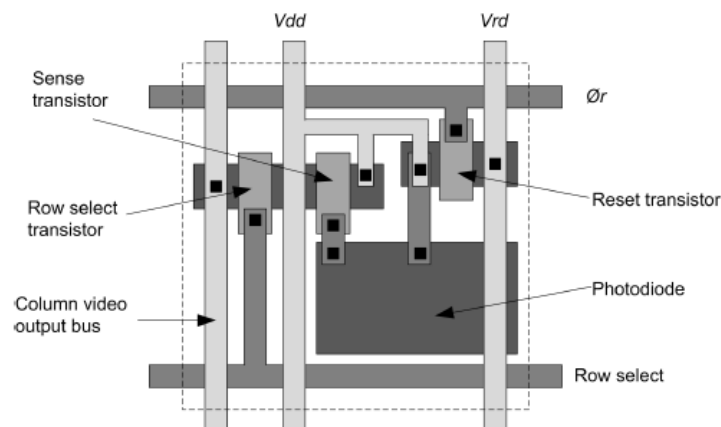
6.4.1: *APS CMOS* architecture and principle of work

CMOS detectors take advantage of the possibility of integrating transistor in the silicon circuit: this presents the opportunity to integrate a large array of pixels, each one with its own photodiode and readout transistor: precisely, an *APS CMOS* consist of a photodiode, a switch and a pixel level amplifier, which result in fast read out and low noise.

In its simplest form, an active *CMOS* pixel architecture, consists of a photodiode and three transistor (3T) (Figure 6.16):

- a photodiode reset transistor M_{RS} ;
- a sense transistor M_{SEL} ;
- a row select transistor M_{RS} ;

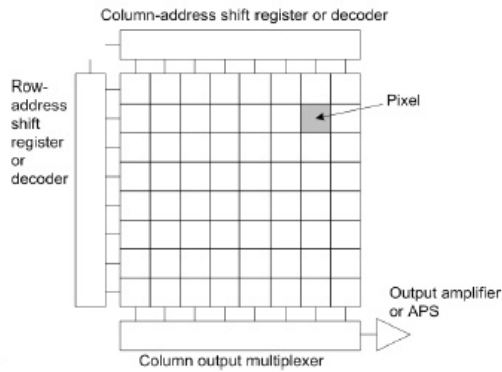
The functions of the transistors are to: pre-charge the photodiode, sense the signal on the photodiode and select the row.



(Figure 6.16: *CMOS* pixel layout)

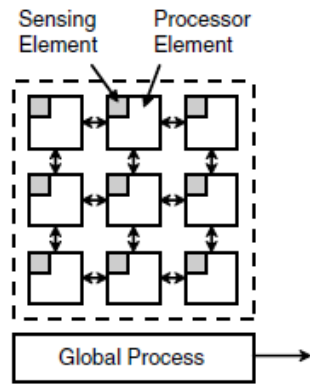
There are also *APS CMOS* types with 4 or 5 transistors: the advantage of these architectures are high sensitivity, low noise but low fill factor; instead the disadvantages of the three transistor pixel are low signal to noise ratio (S/N) and suffer from thermal noise, which is the dominant source of readout noise.

A pixel array will be x-y addressed by shifting registres as can be seen in Figure 6.17:



(Figure 6.17: architecture of the array)

The APS pixel array is accessed one row at a time by enabling all the row-select transistors within a single row of pixels. At the bottom of this array, the pixels within the row are selected and readout column by column. What has been described is the '*column parallel readout and processing architecture*'. Another type of architecture is the '*frame simultaneous readout*', which is used in high speed imaging processing. In these type of architecture, a processor element is located in each pixel and performs image processing in parallel (Figure 6.18). The main disadvantages are the complexity of the scheme and a low fill factor.

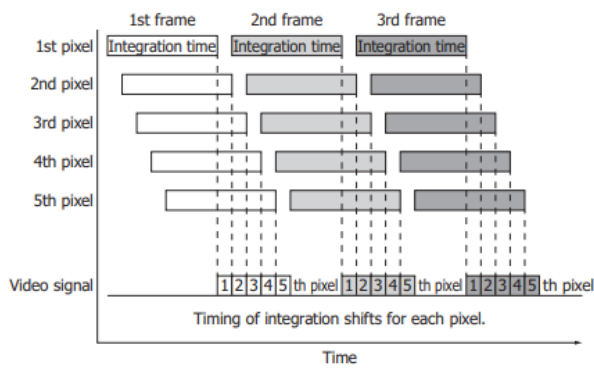


(Figure 6.18: frame simultaneous readout)

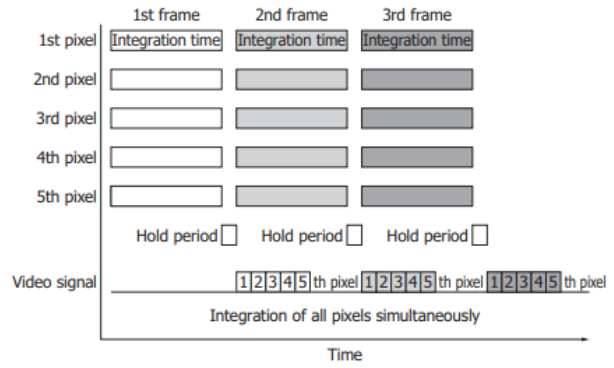
What immediately emerges is that compared to CCDs, in APS CMOS the charge generated is immediately amplified on each pixel before being transferred to the multiplexing readout line and they have a lower fill factor.

6.4.2: CMOS readout mode

In terms of operating methods, *CMOS* operates in two ways: *Rolling shutter* and *Global shutter*. With the *Rolling shutter*, the integration timing is shifted for each pixel (Figure 6.19A), while with the *Global shutter* the integration is done at the same time for each pixels (Figure 6.19B):

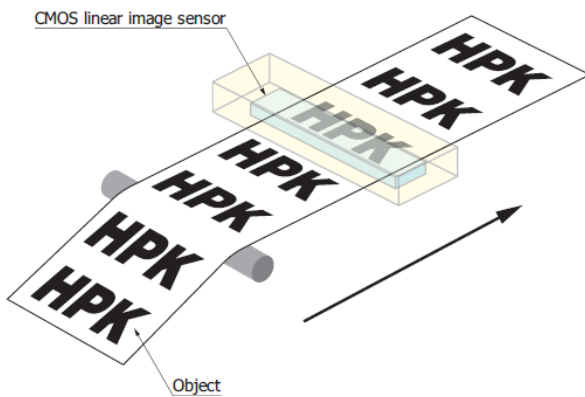


(Figure 6.19A: Rolling shutter)



(Figure 6.19B: Global shutter)

Simply, *Global shutter* is used for imaging fast moving objects and detecting light pulses, instead *Rolling shutter* is used for detection of constant light. If the *Rolling shutter* mode is used to acquire an object in movement, the image will be distorted (Figure 6.20):

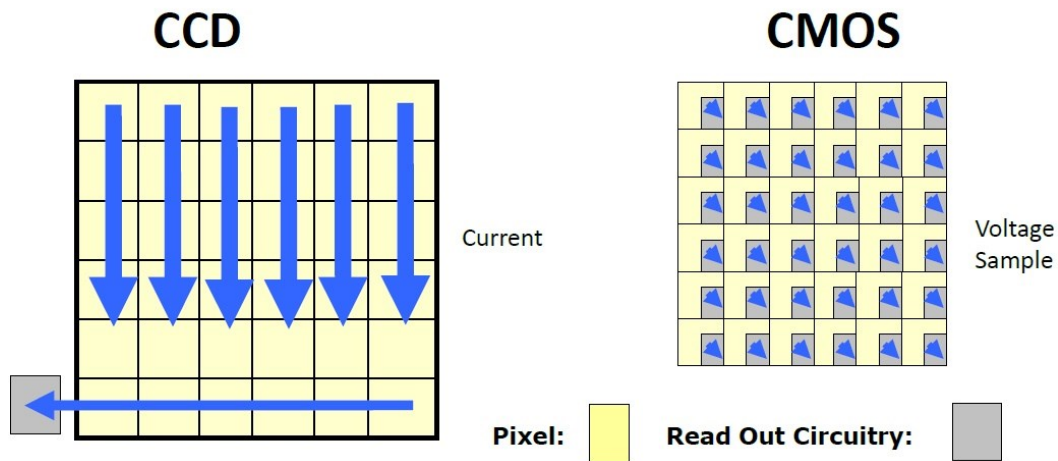


(Figure 6.20: example of distortion)

For what concerned the illumination mode: *CMOS* sensors can be either front or back illuminated, with the characteristics mentioned above.

6.5: Difference between CCD and CMOS image sensor

Principally, the main differences between *CCD* and *CMOS* are the charges transfer and readout (Figure 6.21). In the case of *CMOS*: they incorporate amplifiers, A/D converters, so the charge to voltage conversion occurs in each pixel on the sensor; instead in the case of *CCD* sensors, the charged is shifted along the pixels and finally it is read outside the sensor, i.e. the charge to voltage conversion of all pixels takes place outside the sensor. It is therefore easy to see that the readout speed of *CMOS* is higher than *CCDs*.



(Figure 6.21.: CCD and CMOS main difference)

The main advantages of *CCD* are:

- higher fill factor: this is due to the fact that in *CMOS* sensors, a large portion of the area is used for the amplification and readout electronics;
- higher sensitivity, which is the ability of the photosensitive unit to collect photons and generated electrons;
- higher dynamic range, defined as the ratio of the pixel's saturation level to its signal threshold;

Consequently, the quality of the image with *CCDs* is higher.

On the other hand, the main advantages of *CMOS* are:

- lower power consumption;
- lower cost;
- lower dark current;
- on chip functionality and miniaturisation;
- high speed imaging, because the charge is read in each pixel;
- less blooming and smearing effects.

To conclude: the ideal candidate does not exist as each has advantages as well as disadvantages, so an answer to the question 'which one is better?' can be provided on the basis of a compromise between the required imaging performance and the available resources.

6.6: APS CMOS in space

As mentioned above, *CCDs* have dominated the industrial imaging for many years, but nowadays *CMOS* sensors are making great progress and are also being used for space applications.

Their use derives from their compactness, reduced mass, lower power consumption and resistance to radiation.

Today, *CMOS* are used for example in star trackers (Figure 6.22) for attitude control and in some planetary observation applications, where they are used in Cubesats.



(Figure 6.22.: star tracker)

6.7: Noise

When an image is taken with one of the two sensors, it will not be perfect but will be affected by noise; noise can be seen as a fluctuation of the signal.

Noise can be temporal or spatial. The first one, is the variation in pixel response for constant illumination.

An example of temporal noise is the *dark current*, which it is a current generated under no illumination. *Dark current* is the result of thermally generation of electrons-hole pairs, so it is temperature dependent. A technique to mitigate the amount of *dark current* is to cool down the sensor.

Spatial noise instead, is the variation of pixel response within a frame, i.e. it is a pixel to pixel variation which is steady over time; an example of this one is the *fixed pattern noise*, where particular pixels are susceptible to giving brighter intensities above the average intensity.

CHAPTER 7

CHARACTERISATION OF THE DETECTOR:

This chapter deals with the second part of the thesis work, namely the characterisation of the detector. What has been done is:

- Case 1: linearity test as exposure time varies, under the same lighting conditions (as far as possible);
- Case 2: linearity test as light intensity varies, under the same exposure time.

The aim was to verify that the detector response under various conditions was always almost linear. For both configurations the detector settings were as follows:

- Black Level = 0.1;
- Gain = 0 [dB];
- Acquisition = MONO12;
- Frame-rate = 6 [frame/s];

For both cases, what has been 'measured' at the sensor output is the pixel signal intensity, measured in DN (i.e. digital numbers).

For characterisation, the sensor was placed in its holder (Figure 7.1):

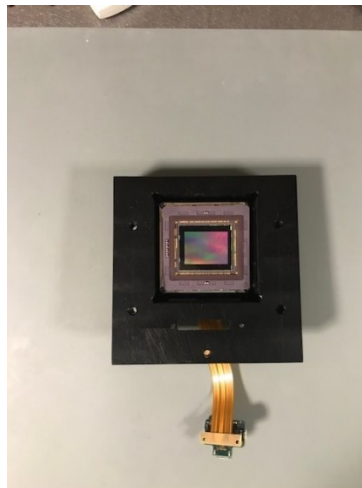


Figure 7.1: detector in its holder)

An experimental set-up was then developed to hold the detector in a fixed position (Figure 7.2):

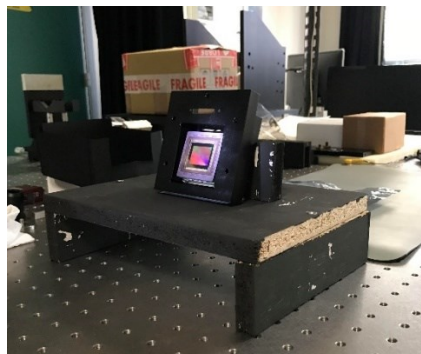


Figure 7.2: detector position)

7.1: Case 1

Starting from the condition in which the exposure time varies, the source used to illuminate the sensor was a torch. This was arranged in such a way that the height of the central spot completely affected the surface of the sensor (Figure 7.3):



Figure 7.3: complete experimental set-up)

Under this condition the detector has been studied in three different ways:

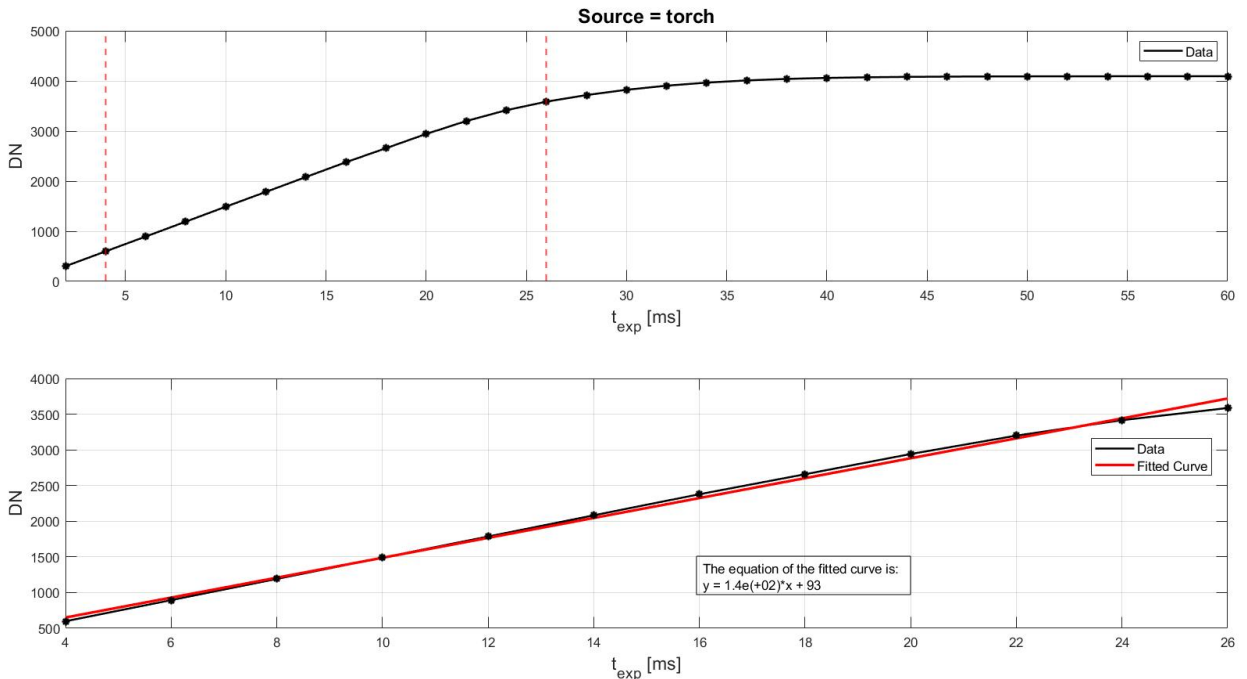
1. whole surface analysis with torch as source;
2. detector analysis with the surface divided into four zones;
3. analysis of the detector with the *integrating sphere*;

7.1.1: Whole surface analysis

In this situation the experimental setup is the same as seen in Figure 7.3.

The detector has been studied for an exposure time between 2-60 ms.

Figure 7.4 shows the obtain results:



(Figure 7.4: whole surface; linear behaviour)

What the graph shows is the average illumination level over the entire sensor's area.

In this condition the dark is not yet taken into account.

As in the case of photodiodes, the Non-Linearity [NL%] between the linear fit and the measured data was determined. So by using equation (5.1), the maximum [NL%] achieved is:

	$max\{NL\%$
Detector	8.0

Observing Figure 7.4, precisely the top graph: there are two reds vertical line. The lines are used to indicate that the linearity measurement was conducted in the area between 10% and 90% of the saturation level.

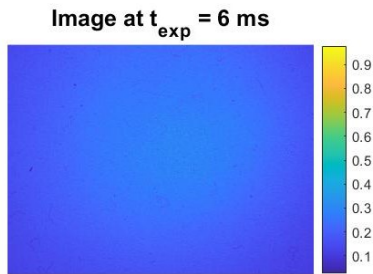
Since MONO12 is used as the channel, the linearity range to be considered goes from:

$$0.1 * 2^{12} = 409.6 \text{ [DN]}$$

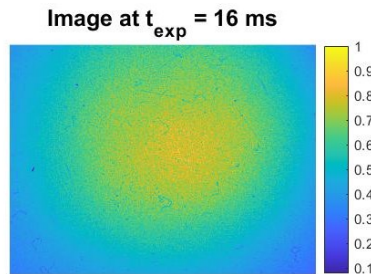
$$0.9 * 2^{12} = 3686.4 \text{ [DN]}$$

What we can see, therefore, there is a good correspondence between the linear fit and the measured data.

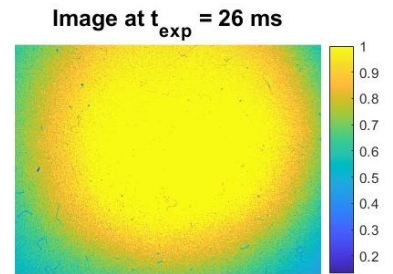
The next three figures (7.5A, 7.5B and 7.5C) show the level of illumination on the sensor, obtained at three different exposure times:



(Figure 7.5A)

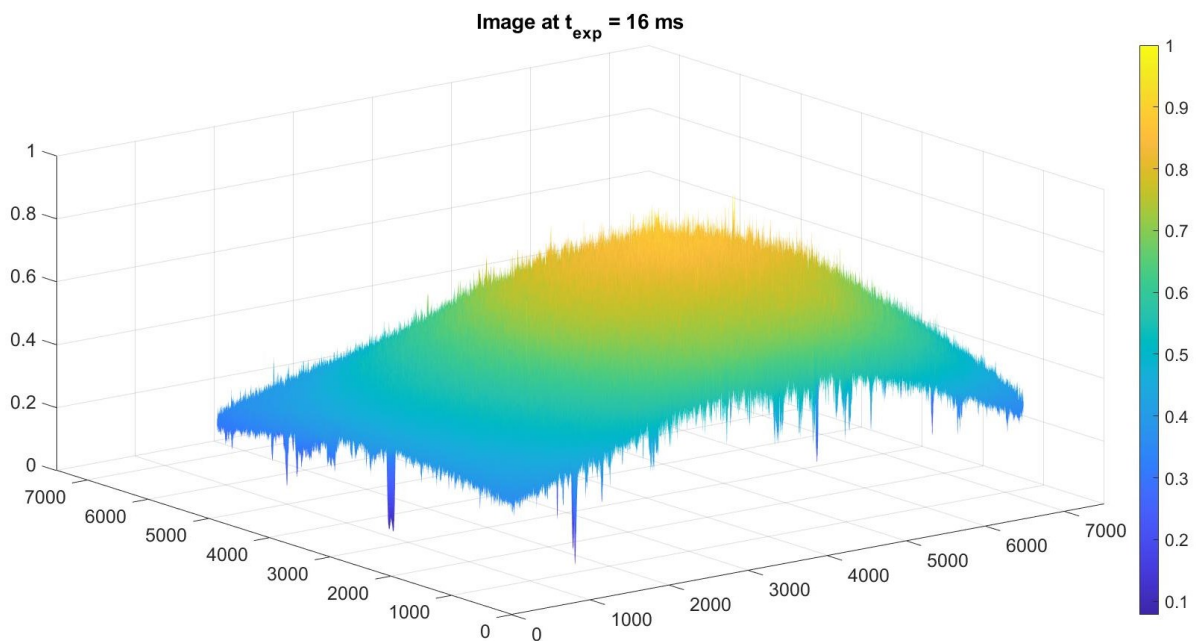


(Figure 7.5B)



(Figure 7.5C)

Looking at figure 7.5A, it is immediately apparent that the level of illumination on the sensor is not uniform, but only covers a certain portion of its surface. In figure 7.5C, on the other hand, we can see that the system starts to saturate; in fact, observing Figure 7.4, at $t_{exp} = 26$ ms, we are exiting the linearity range. Figure 7.5D (referring to figure 7.5B) better highlights the lighting condition at the sensor:



(Figure 7.5D: normalised illumination)

Looking at Figure 7.5D, there are areas where the intensity level is much lower than the others; this is due to the presence of the protective film on the surface of the detector.

7.1.2: Divided surface analysis

When studying the linear response of CMOS sensors, the analysis should be conducted pixel by pixel, in order to identify the behaviour of each pixel. However, as the HYPSON sensor is very large (6840x4856 pixels), this is a very complex task, and we preferred a simpler approach.

What was done was to divide the surface into four zones (see Figure 7.6), where the dimension of each portion of the area is:

$$A_i = 3420 * 2428 \text{ pixels}$$

and analyzed them separately, evaluating the average values over these zones.

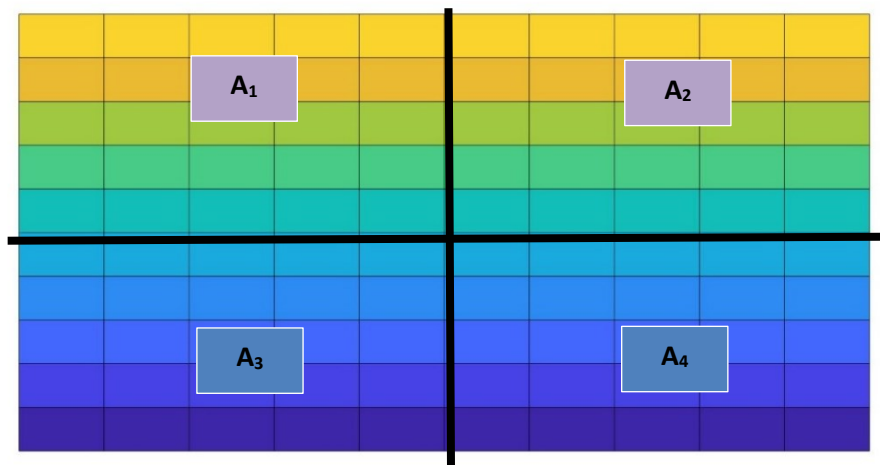
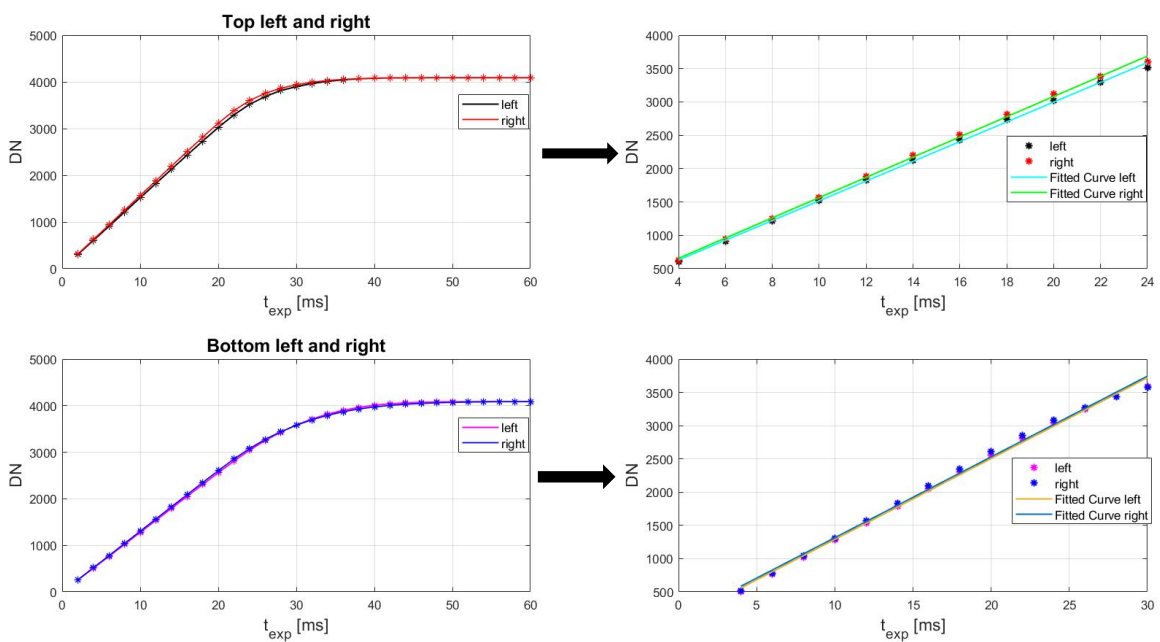


Figure 7.6: division of the detector's surface)

Figure 7.7 shows the results obtain:



(Figure 7.7: linear behaviour)

The following table shows the linear fits for the various zones:

A_i	<i>Equation</i>
A ₁	$y = 150*x + 38$
A ₂	$y = 150*x + 49$
A ₃	$y = 120*x + 73$
A ₄	$y = 120*x + 100$

Linearity was again studied in the range between 10% and 90% of the saturation level and what the graph shows is the average illumination level over the various area.

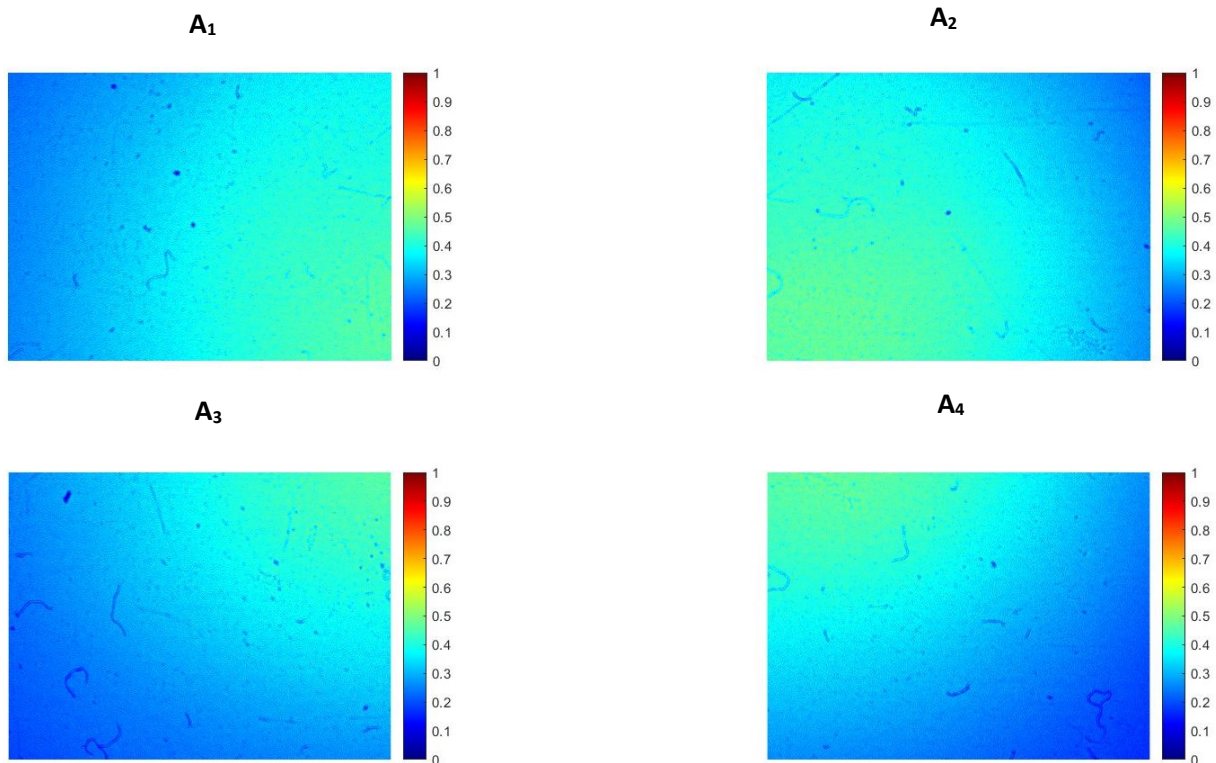
Also the Non-Linearity [NL%] has been computed again:

A_i	<i>max {NL%}</i>
A ₁	3.6
A ₂	4.5
A ₃	8.6
A ₄	11.1

In contrast to the previous case (*Whole surface analysis*), the dark signal was taken into account here; therefore the plotted value in Figure 7.7 are as follows:

$$[DN]_{plotted} = [DN]_{measured} - [DN]_{dark} \quad (7.1)$$

Also in these case, four bidimensional images of the detector has been taken (which are different from Figure 7.5), in order to understand the level of illumination in each portion of the surface (Figure 7.8). Each image has been taken at an exposure time of 10 ms:



(Figure 7.8: level of illumination)

Looking at Figures 7.7 and 7.8, it can be seen that the level of illumination on the sensor is not uniform because:

- looking at *Top left and right* plots, there is a small difference in the linear fit; in fact, the constant term in the equation is greater in the right-hand case; also, looking at Figure 7.8 (A₂), it can be seen that the illumination level is slightly higher than in area A₁;
- considering instead *Bottom left and right* plots, the linear fit equations have lower slope; this is motivated again by looking at Figure 7.8 (A₃ and A₄), where the illumination level is lower than in the other two cases.

On the basis of what has just been said, it is clear that the spot is decentralised.

Despite these slight discrepancies, also in these case, there is a good match between the measurement and the linear fit as can be seen from Figure 7.7.

7.1.3: Analysis with integrating sphere

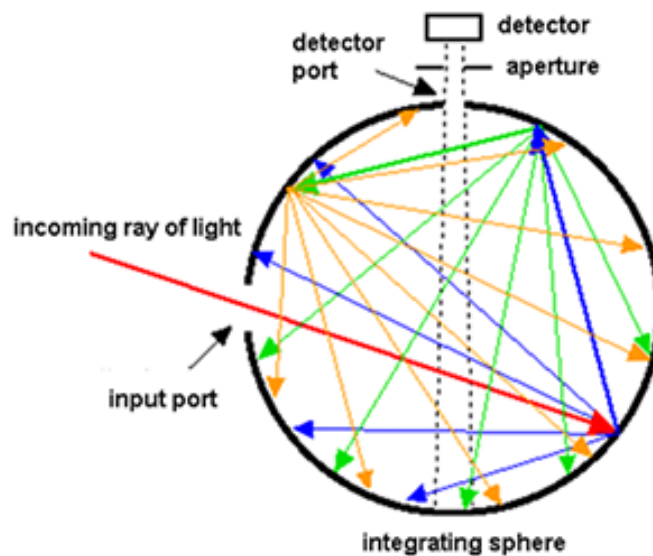
It is clear from Figure 7.5B that the illumination provided by the torch is not sufficient to cover the entire detector area uniformly. To try to overcome this problem, an integrating sphere was used, Figure 7.9:



(Figure 7.9: integrating sphere)

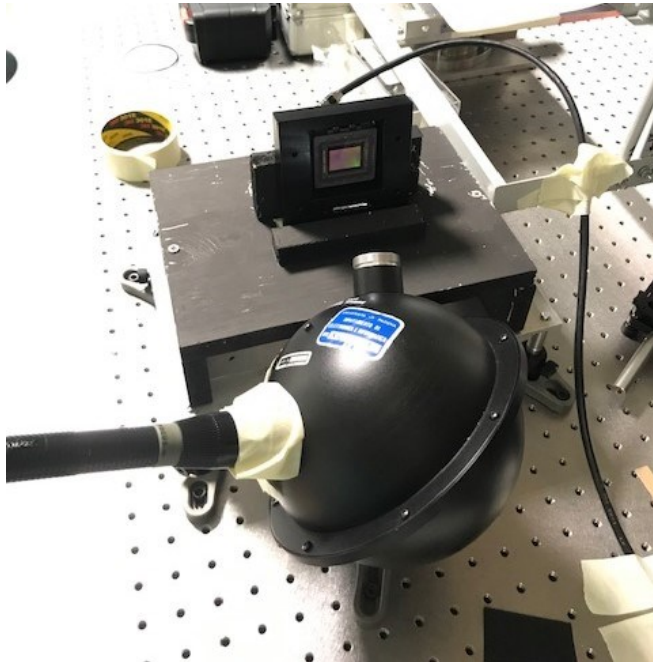
The integrating sphere is hollow inside and consists of a completely diffusing surface, which allows total reflection of the light; its main function is to spatially integrate radiant flux. The surface therefore should approach the ideal Lambertian scatter, i.e. the light which emerges from the output of the sphere is scattered in all direction with the same intensity. The aim, therefore, is to have an illumination, at the output, as more uniform as possible.

Figure 7.10 shows the working principle of the sphere.

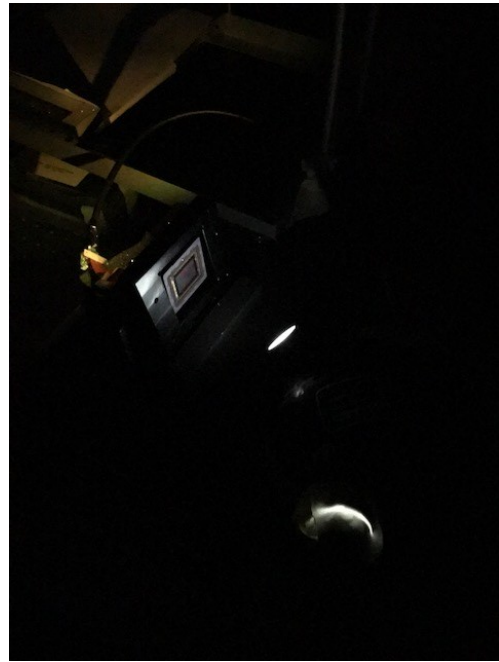


(Figure 7.10: integrating sphere, principle of work)

Figure 7.11A and 7.11B show the experimental set-up.



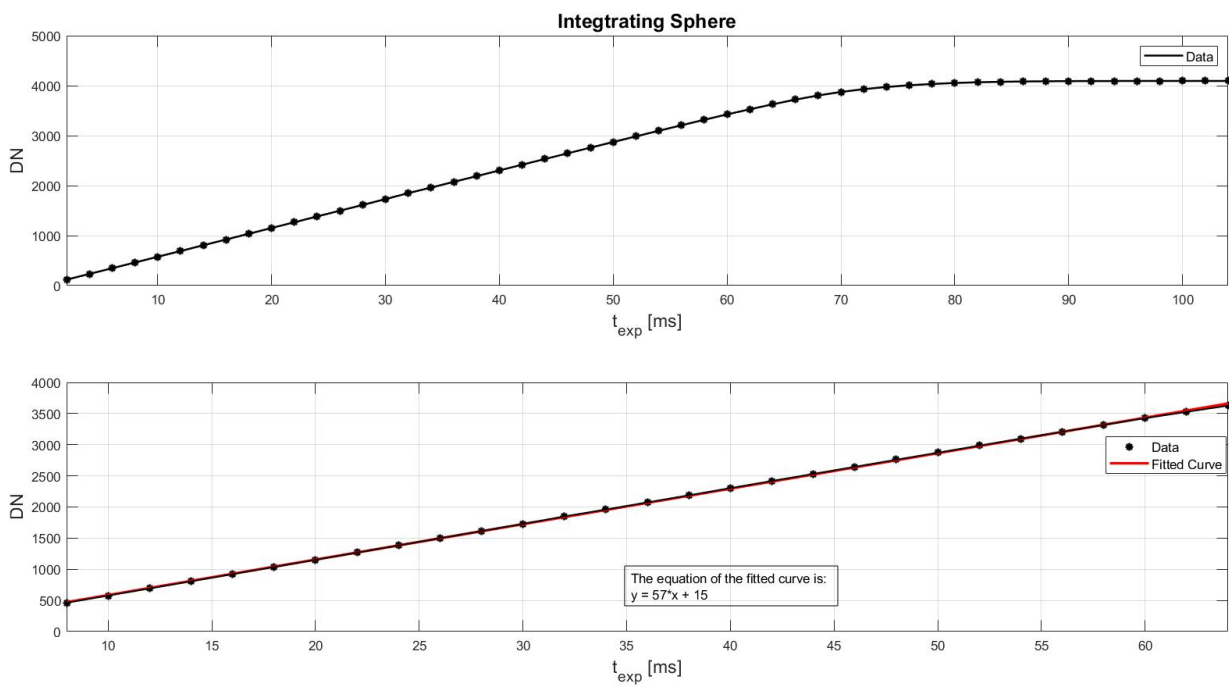
(Figure 7.11A: experimental set-up)



(Figure 7.11B: experimental set-up)

In this situation the surface has not been divided into four zone as the previous case and the dark signal has been taken into account.

Figure 7.12 shows the results obtain:



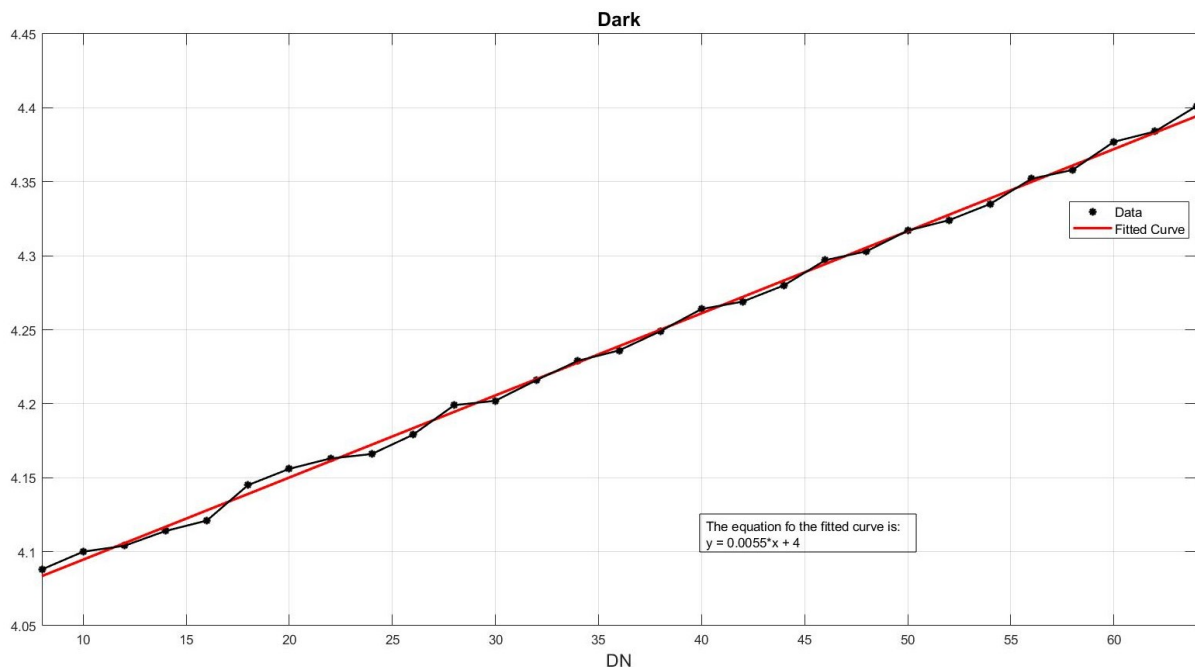
(Figure 7.12: linear behaviour)

Unlike the previous cases (Figure 7.7 and Figure 7.4), the exposure time here ranges from 2 to 104 ms; this is because the light intensity is smaller than in the previous cases, although the source used in the input of the sphere was still the torch. This is due to the fact that not all the light emitted by the source enters the sphere, but part of it spreads out of the sphere (as can be seen in Figure 7.11B).

The maximum *NL%* achieved is:

	<i>NL%</i>
<i>Detector</i>	2.1

Again, there is a good match between the measurement and the linear fit. Figure 7.13 shows the trend of the dark signal over time:



(Figure 7.13: dark signal linear behaviour)

The maximum *NL%* achieved is:

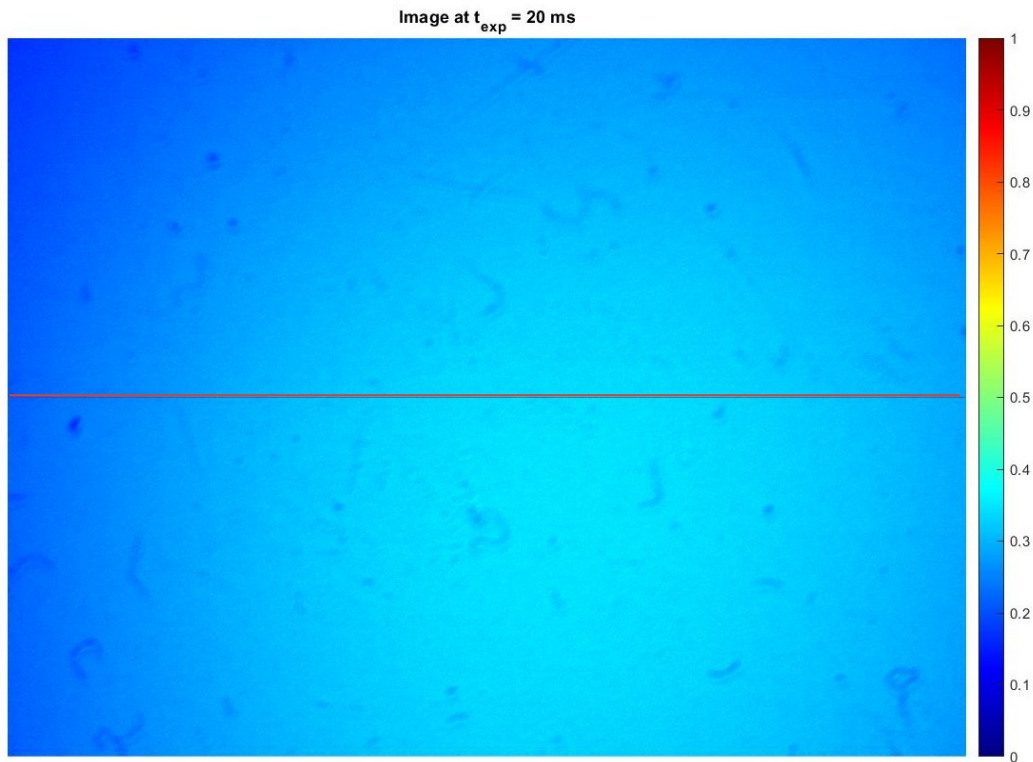
	<i>NL%</i>
<i>Detector</i>	0.2

The dark signal trend is fluctuating, rising slightly and falling. This can be due to various factors, such as:

- not having waited long enough before acquiring the signal;
- the presence of spurious light from other sources, even though the detector was completely darkened.

However, the main source of the dark signal is thermal noise, as the temperature reached by the detector was very high.

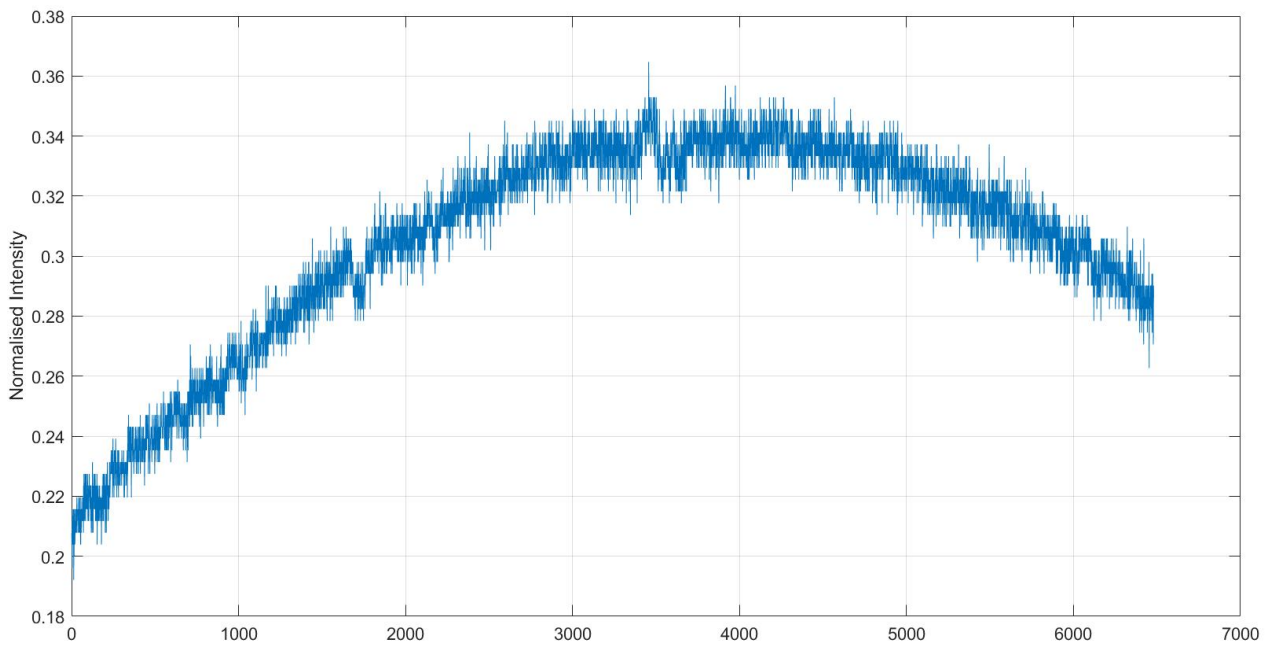
An image was also captured here to understand the level of uniformity of the light, Figure 7.14.



(Figure 7.14: level of illumination)

If you look closely at the figure above, it can be seen a central spot that is brighter than the surrounding background: this spot is due to the illumination provided by the sphere. It is clear from this, that using the sphere has largely improved the uniformity of illumination over the entire surface of the detector, with respect to the torch illumination case.

Figure 7.15 instead shows the intensity profile, normalised, along the line highlighted by the red line from Figure 7.14. Figure 7.15 highlights better how the illumination level is not uniform.



(Figure 7.15: normalised illumination)

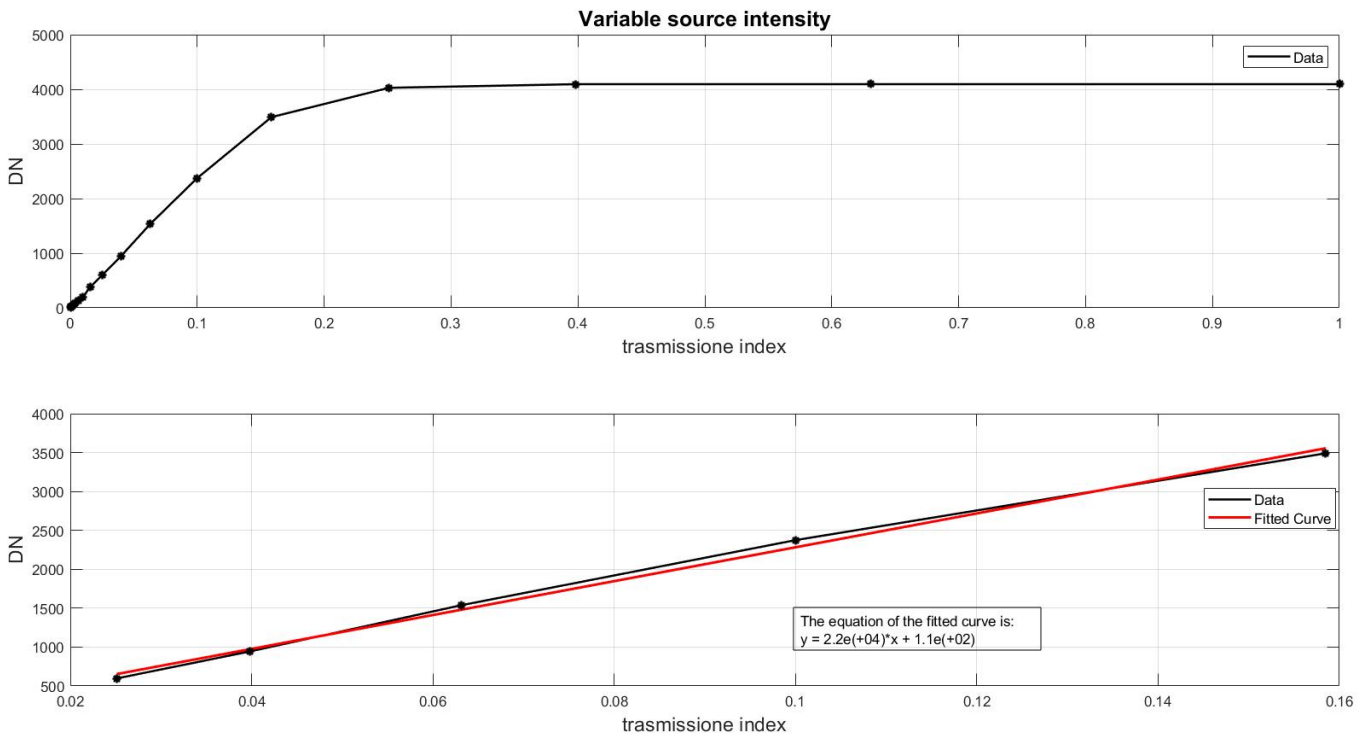
7.2: Case 2

In this situation the exposure time is fixed at 2 ms, and the light intensity varies thanks to the use of neutral filters.

As the previous three cases, the average of the signal has been considered, and the study has been done considering the entire surface of the detector.

The experiment set-up is the same as in Figure 7.3, except that several filters were placed between the source and the detector.

Figure 7.15 shows the results obtain:



(Figure 7.15: linear behaviour)

The maximum *NL%* achieved is:

	<i>NL%</i>
<i>Detector</i>	5.4

Even with this configuration there is a good correspondence between the two curves, but compared to the three previous cases, here the deviation from linearity is much more evident: this is due to the fact that there are fewer values than in the previous cases because the variation of the transmission coefficient does not scale uniformly.

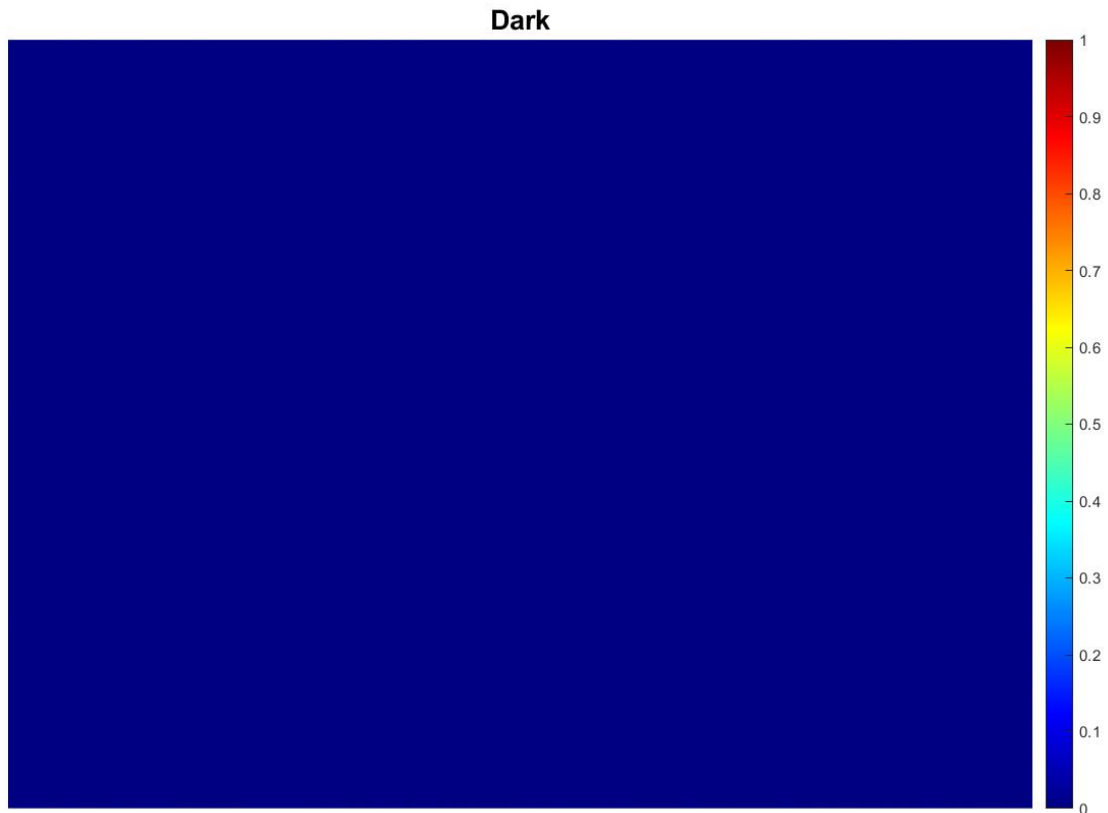
Finally, a brief analysis was carried out to identify possible hot pixels, i.e. pixels that saturate or are about to saturate even at low lighting conditions and short exposure times.

To find these, different dark signal images were taken and then a search was made to identify pixels that produced an above-average output.

Operating in this way, it is possible to identify the coordinates of the hot pixels and subsequently remove them after acquisition in the illuminated condition: this operation is called darkframe subtraction.

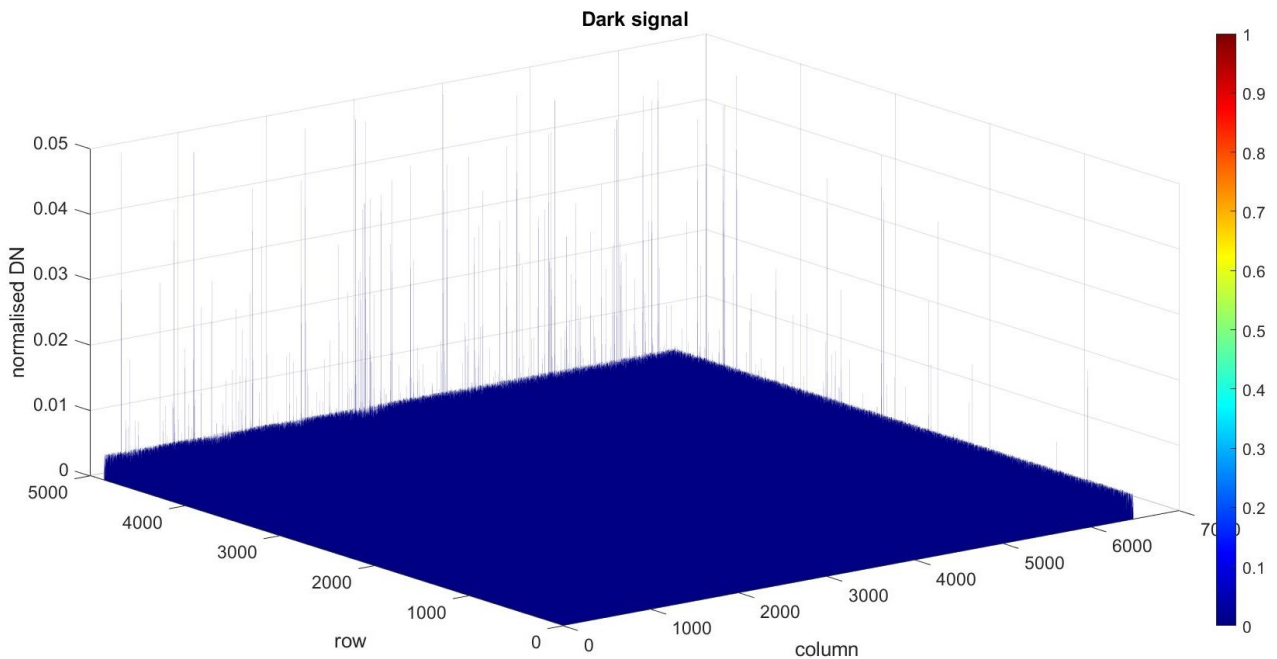
However, this operation requires that two images are taken for each exposure time, one of which in light conditions and the other in dark conditions; it is therefore a complex operation.

Figure 7.16 shows the illumination level in dark conditions for an exposure time of 20 ms:



(Figure 7.16: dark image)

Looking at the figure above, it appears that the dark signal is uniform over the entire surface of the sensor. In reality, this is not true and can be seen from Figure 7.17:



(Figure 7.16: normalised dark signal)

The graph makes it clear that the response of individual pixels is not uniform, in fact some produce much greater output than others: these are the hot pixels.

The z-axis is normalised and truncated at 5% of the saturation level to better illustrate the level of pixel response.

Once the coordinates of the hot pixels have been identified, it is sufficient to exclude these pixels in the images to be acquired in order to better analyse the images.

In addition to the correction method described above, there are two software methods for correcting hot pixels:

1. *'Replace 3x1 average'*, which substitutes the local hot pixel pixel with the average value from the left and right neighboring pixel (3x1);
2. *'Replace 3x3 median'*, which substitutes the local hot pixels with the median value calculated from the nearest neighboring in a 3 by 3 region.

Despite these measures, the analyses conducted in the previous cases were carried out without hot pixel correction.

7.3: Conclusions

By observing all the plots relative to the linear behaviour of the detector, is clear to notice that the device has a good linear behaviour (according to theory) as the light intensity varies at the same exposure time or as the exposure time varies at the same illumination condition (as far as possible), although it was not possible to illuminate the sensor area uniformly.

Finally, even the dark signal, despite a more fluctuating trend, has good linearity on the whole: the main source of the dark is still the thermal noise.

Looking at the images associated with the level of illumination on the sensor, it appears that the sensor has surface defects. In reality, these are due to the presence of the protective film, which has not yet been removed.

Finally, an analysis was carried out to search for any hot pixels, and it was seen that these produce a much higher signal than the average. Although the dark signal is relatively low, the exclusion of hot pixels will cause it to be significantly reduced.

RINGRAZIAMENTI

First of all, I would like to thank Prof. Giampiero Naletto, a polite and highly competent person, for giving me the opportunity to take part in this wonderful project and for following me throughout these months.

I would also like to thank Agostini Livio and Brotto Francesca for supporting and guiding me during the laboratory and other activities throughout my thesis period.

I would also like to thank my classmates: Anoè Lorenzo, Celadin Denis, Fortuna Simone and Magro Nicolò with whom I had the pleasure and honour of sharing doubts, intense hours of study and great laughter during these six years of University.

Finally, I would like to thank my family and my girlfriend Kristina, for always believing in me and for the support given to me during all these years.

In primis desidero ringraziare il Prof. Giampiero Naletto, persona educata e di grande competenza, per avermi dato l'opportunità di prendere parte a questo meraviglioso progetto e per avermi seguito in tutti questi mesi. Ringrazio anche Agostini Livio e Brotto Francesca per avermi supportato e guidato durante le attività di laboratorio e non, per tutto il periodo di tesi.

Ringrazio anche i miei compagni di corso: Anoè Lorenzo, Celadin Denis, Fortuna Simone e Magro Nicolò con cui ho avuto il piacere e l'onore di condividere dubbi, intense ore di studio e grandi risate durante questi sei anni di Università.

Desidero infine ringraziare i miei familiari e la mia ragazza Kristina, per aver sempre creduto in me e per il sostegno datomi in tutti questi anni.

BIBLIOGRAPHY

- [1] Christopher T. Hill e Leon Max Lederman *'Fisica quantistica per poeti'*.
- [2] Giampiero Naletto, Livio Agostini, Francesca Brotto, Gabriele Cremonese, Matteo Faccioni, Luigi Lessio, Cristina Re, Massimiliano Tordi, Carlo Bettanini, Fabrizio Capaccioni, Maria Teresa Capria6 Stefano Debei,, Ennio Giovine, Lucia Marinangeli, Francesco Mattioli, Marco Pertile, Amedeo Petrella, Giuseppe Salemi, Anna Chiara Tangari, Michele Zusi, *'Laboratory characterization of HYPPOS, a novel 4D remote sensing Instrument'*.
- [3] P.Mazzoldi, M.Nigro e C.Voci *'FISICA'*, Vol. II.
- [4] Christopher Palmer *'DIFFRACTION GRATINGS HANDBOOK'*, eighth edition.
- [5] Chris Churchill *'Spectrographs'*, chapter 6.
- [6] Jian Gao, Peng Chen, Lei Wu, Bingjun Yu e Linmao Qian *'A review on fabrication of blazed Gratings'*.
- [7] R.Casini e P.G.Nelson *'On the Intensity Distribution Function of Blazed Reflective Diffraction Grating'*
- [8] Jinzhong Ling, Dawei Zhang, Yuanshen Huang, Qi Wang e Songlin Zhuang *'Fabrication of holographic blazed grating with Fourier synthesis exposure'*.
- [9] Photonics-Enabled Technologies *'Photonics in Nanotechnology Measurements: A Study of Atomic Force Microscopy'*.
- [10] Livio Agostini *'HYPPOS: Detector and diffraction grating characterization'*.
- [12] Daniele Durini, *'High Performance Silicon Imaging: Fundamentals and Applications of CMOS and CCD sensors'*.
- [12] Nick Waltham *'CCD and CMOS sensors'*.
- [13] Chenming Hu *'Electrons and Holes in Semiconductor'*.
- [14] Chenming Hu *'MOS Capacitor'*.
- [15] HAMAMATSU *'CMOS linear image sensors'*.
- [16] Nicolas Blanc *'CCD versus CMOS – has CCD imaging come to an end?'*.
- [17] Lihua Zhang, Yongjun Jin, Lin Lin, Jijun Li e Yungang Du *'The Comparison of CCD and CMOS Image Sensors'*.
- [18] Utsav Jain *'Characterisation of CMOS Image Sensor'*.
- [19] PIKE *'Integrating Spheres – Introduction and Theory'*.
- [20] pco.knowledge base, *'warm/hot pixel'*.