Development of a spectropolarimeter for space use

Auteur : Vasilescu, Bogdan
Promoteur(s) : Loicq, Jerome
Faculté : Faculté des Sciences
Diplôme : Master en sciences spatiales, à finalité spécialisée
Année académique : 2018-2019
URI/URL : http://hdl.handle.net/2268.2/6186

AGO - Department of Astrophysics, Geophysics and Oceanography

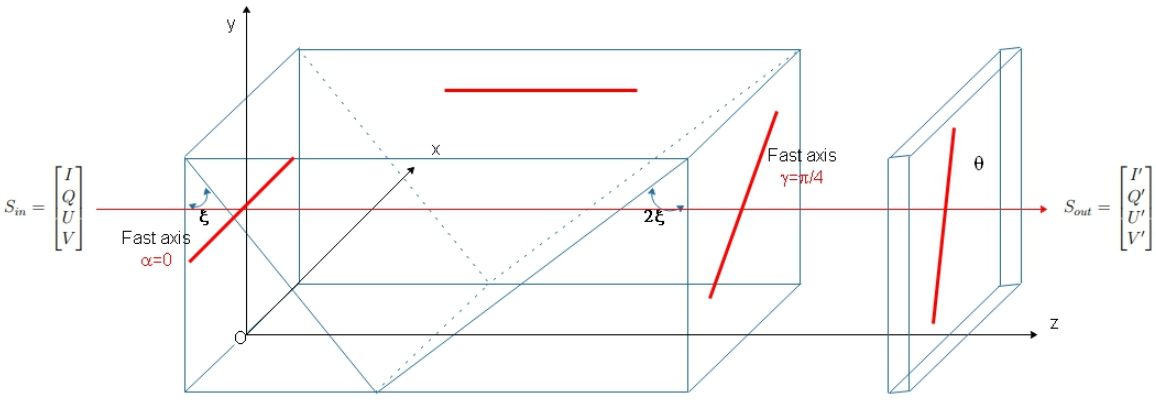
Master in Space Sciences

2017-2019

MASTER THESIS

Bogdan Vasilescu

Development of a spectropolarimeter for space use



Supervisor:

Prof. Jérôme Loicq

Contents

Acknowledgements

Firs of all, I would like to thank my coordinating teacher, Prof. Jérôme Loicq. His guidance and his clear vision about the subject and the purpose of our work helped me to keep the right approach during the entire research. His enthusiasm gave me the energy to overpass even the most difficult moments.

I would like also to express my gratitude to my optics teacher, Prof. Serge Habraken, and not only for the precious advices related to this thesis but also for the entire support during the last five years of study. Without the deep and passionate incursion in the world of optics through which he guided us nothing of this would be possible today.

An immense support come also from Dr. Yaël Naze. Her advices and hard work for correcting this thesis helped me to reach the necessary clarity and to fully meet the requirements of a scientific research.

A deep acknowledgement I have also for Christian Kintziger, Denis Defrére, Marin Ferrais and Luca Schifano. The long discussions with them helped me to clarify countless aspects related to this work.

Last but not least, I would like to thank to my family, for the entire support during all these years.

Introduction

The study of the polarization of light can be a very resourceful instrument in most of the fields where the light is the key player. Understanding and controlling this phenomenon can open the door to astonishing devices and instruments in various domains ranging from medicine until to military as well as to a better knowledge of the emitting bodies, whether they are around us whether they are extremely remote. And the astronomy is one of the domains that fully realized the importance of this type of study, taking advantage of it even from its discovering, in the mid of the 18th century. Since then, it never stopped to push further the development of polarimetry and of its applications, aside spectroscopy, this field of optics becoming one of the most appropriate instruments to sound the astronomical objects.

Despite of its capital importance in this type of researches, the polarimetry is not a very appealing domain: the hardness of the mathematical apparatus, the extreme sensibility of instrumentation as well as the hard work of interpretation of the results have been a huge hindrance in its development.

Nevertheless, behind all these difficulties lies a wonderful science, that made important steps ahead during the last decades, if it will be only to count here the Stokes formalism, the Jones calculus or the Mueller matrices: techniques which are able to completely describe the polarized light today. Still, the polarimetry didn't become fully a scientific instrument in the service of astronomy until the diffraction and the scattering theory were rigorously developed. It was only then that the significance of certain type of polarization retrieved in astronomical observation was understood and helped to the description of the sources. Only starting from that moment, the polarization was able to bring information about magnetic fields, about the presence of certain particles in the atmospheres of the observed celestial bodies, about their concentration or shape.

Over the time, polarimetry grew into a fully mature domain of optics, finding applications in countless sectors. Thus, for example, in atmospheric remote sensing, polarimetry allows the study of aerosols particles, enabling to retrieve capital properties like optical thickness, size distribution and shape, information which are both relevant for air traffic flow and for climate modeling. Sometimes comparative laboratory studies and modelling enable even the determination of the chemical structure of the molecules or particles responsible for the polarization. In the defense field, one may quote the use for target detection. In astronomy, where the degree of precision must be extremely high due to the tiny intensity of the signals that must be detected, the imaging polarimetry and the spectropolarimetry have been developed. The first is used to retrieve the circumstellar structures for example, while the latter can reveal the environment of stars and exoplanets.

Furthermore, the polarization can yield information about the presence of liquid water, of oxygen and, ultimately about the existence of some complex organic molecules (like amino acids, sugar, DNA and chlorophyll) which are able to affect on a certain level only the circular polarization.

The aim of this thesis is then to contribute, even if in a tiny measure, to the advance of this field

of research which is the polarimetry, by the study of a spectropolarimeter concept aimed to the space use.

Usually satellites and space telescopes which are equipped today with instruments able to reveal the polarization of light are using a technique that suppose the relative rotation of several components so that the electromagnetic radiation could be observed through different media, each time a certain property (or Stokes parameter) being pointed out. Because of its complexity and fragility this type of polarimeter is not very widespread among space missions. Just by looking to the structure of payloads from the spacecrafts, whether they are developed by NASA or the European Space Agency (ESA), we see that they are not very popular: only a few instruments for each Agency in the recent history. Overall there is an extremely small number compared to the other type of optical devices and with the importance of information that can be brought by the polarization of light. The solution is then to try to develop furthermore this instrument so that it becomes "space reliable". And the first steps were already made: a series of studies were conducted in the last years towards the development of a "static" polarimeter, a complex device based on the use of birefringent materials.

Our work will come also in this very direction: we will try to analyze one of the proposed devices from the recent literature, a spectropolarimeter build from a birefringent crystal of MgF_2. The device is supposed to provide simultaneously access to all four Stokes parameters, the intensity of the emerging light being also modulated spatially and regarding the wavelength. Concretely, the aim of this thesis is to provide the detailed and the complete deduction and description of the modulation function of the device and to see the eventual limitations imposed by this function. The sensibility of the concept will be tested by error analysis and numerical simulations in the presence of noise. In the end, the off-axis capabilities of the instrument will be also verified, trying to fully understand its functioning, its limitations and how can be improved.

We will see that this kind of analysis is not an easy task: it will require a lot of computations, a deep diving into the entire theory of polarization, and a very intensive numerical analysis. It will be also a great opportunity to better seize difficult concepts like those regarding the distribution of the refraction indices in the birefringent media, the ray tracing for this type of environment as well as regarding the usage and importance of polarization for astronomers.

Thus, in the first part of this work we will pass in review the general description of the phenomenon of polarization and of the processes responsible for this which are mostly used in astronomy. Briefly we will pass through the description of the Zeeman effect, of the Rayleigh and Mie scattering and through the Hanle effect. For a better image about the manifestation of the polarization and its importance we will see then a few results from the Solar System observation and from the interstellar medium.

In the third chapter we will encounter a few polarimeters and spectropolarimeters designed for the space use. We will see for example the POLLUX instrument intended to be part of the LUVOIR payload and SPEX, a project based on birefringent media, recently developed.

Further, the mathematical description of the polarization will be introduced. We will pass through the Stokes, Muller and Jones formalism with an emphasis on the computations and notions that are useful for our purpose. In the end of this mathematical incursion we will have the matrices able to describe the functioning of the spectropolarimeter.

Because the spectropolarimeter studied here uses the property of birefringence of MgF_2, a chapter is dedicated to the description of the propagation of light in this type of crystal. At this occasion we will see also the principle of Huygens constructions. All the properties presented here will be used further, in the last chapter, when a series of corrections will be applied to the model from the scientific literature.

With the required formalism established we will pass then to the analyze of the polarimeter. Because all the equations and formulas which surround the working principle of a polarimeter are often quite extensive and do not allow to see at once the mechanism behind the detection of polarization, the example of CLASP instrument will be detailed. Thus, we will see in the Chapter 5 why the rotation is needed to examine all four Stokes parameters. Finally, the design of our instrument will be introduced and described. In the end of the Chapter 6, the modulation function of the spectropolarimeter is derived.

The functioning of the instrument at different wavelengths forces us to take also into consideration aspects related to the variation of the refraction indices with the wavelength, or the chromaticity. Two models for this variation will be developed in the Chapter 7: one will be based on the Sellmeier method and the other on the interpolation of experimental measurements. This variation will be then integrated into the modulation function of our system and the consequences will be pointed out.

A first series of limitations for the studied design will become visible at this moment: limitations regarding the transmittance of the material and about the minimal dimensions of the optical device.

The Chapter 8 is completely dedicated to the error analysis. A linear adjustment by least square method will give us the theoretical error on each Stokes parameter. Secondly, through the simulation of a noisy signal we will detect new limitations for our design.

In the Chapter 9 we will approach the off-axis problem, trying to understand the behavior of this device when tilted rays are passing through. Mathematically, this is the most challenging part of the study. Because of this, in order the simplify the lecture, tedious mathematical developments will be here often replaced with numerical simulations and graphical analysis. In the end of this chapter we will be able to present the last series of constraints for the studied spectropolarimeter.

Chapter 1

Polarizing processes in nature

The polarization is a very widespread phenomenon in nature. But because our eyes are not sensible to it, we must point it out using polarizers.

The polarization is determined by the emission conditions (or re-emission, in the case of dispersion and reflection). In a more subtle way, this phenomenon occurs where is a symmetry problem: a scalar dissymmetry manifested into the concentration of scattering particles or their dimension is accompanied for example with the circular polarization, while vectorial dissymmetry like the presence of a magnetic field is accompanied most often by the linear polarization[42].

Theoretically, the linear polarization corresponds to the situation in which the electric field of an electromagnetic wave vibrates in a single plane regardless to the moment of observation.

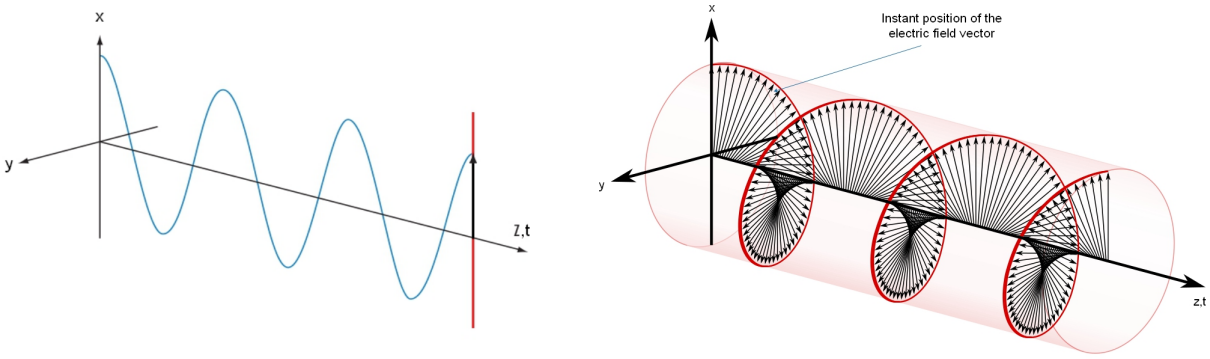


Figure 1.1: Linear (left) and circular (right) polarization.
Representations adapted after *Wikipedia.org*[8] (circular) and *edmundoptics.com*[2] (linear)

On the other hand, if the vector describing the electric field rotates in time around the direction of propagation we are dealing with the circular polarization. A full mathematical description of those situations will be developed in the Chapter 4.

At the deep level of the quantum physics the light is always circularly polarized. The photons are presenting all the time with an equal probability a left- or right-handed circular polarization, according to the orientation of their spin. If we would have the capability to perceive the light photon by photon it will exhibit always then a circular polarization. But this is far beyond the capability of the detectors that are dealing with packages of photons. Because of this the detected polarization is always merely an average property of that packages, average which is governed by the superposition of the individual states of the photons in the limits of the coherence conditions.

Aside the Rayleigh, Thomson and Mie diffusions that are responsible all the time for a certain type of polarization and revel to be extremely important for the study of scattering medium, one of the most important sources of polarization largely studied and used by astrophysicist is the Zeeman effect. Basically, it is through this phenomenon that we may have access to the projected absolute value of the magnetic field of a remote body. Nevertheless, we do not have to ignore the Hanle effect, which open also a door to the determination of the magnetic field in certain situations.

1.1 Zeeman effect

The Zeeman effect is one of the most important phenomena used by astronomers to determine the projected value of the magnetic field of a star, and which highlights the contribution of polarimetry. In just a few words, this effect, with a profound quantic origin, is translated by the splitting of the spectral lines of an atom in the presence of a magnetic field. Besides this, the radiation emitted during the transition process exhibits all the time a precise polarization.

In order to fully understand the Zeeman effect, we must remember a few basic concepts and notations from the quantum physic and electromagnetism. Thus, an electron spinning around a nucleus will generate a magnetic momentum $\vec{\mu}$ and as a consequence can be seen as a dipole. Hence, by placing the atom in a magnetic field $\vec{B}$, a torque will be exerted upon this dipole. The magnetic potential energy characterizing the interaction is given by:

$$U = -\vec{\mu} \cdot \vec{B}$$

When the Zeeman effect was firstly discovered the spin of electron was still a mystery, and because of this the magnetic momentum $\vec{\mu}$ was associated only to the angular momentum of the electron. According to the quantization of the orbital momentum ($\vec{L}$), we have:

$$\vec{\mu} = -\Big(\frac{e}{2m_e}\Big)\vec{L} \qquad L = \hbar\sqrt{l(l+1)}$$

where $\vec{L}$ is the angular momentum of the electron and l is the angular momentum number ($l = 0, 1, ..., n$, for n, the principal quantum number).

If we consider a magnetic field oriented along z axis, then the projected orbital angular momentum is:

$$L_z = m_l\hbar$$

in which m_l is the number associated to the projection of l:

$$m_l = -l, -l+1, ..., l-1, l$$

An intuitive representation of the projection of $\vec{L}$ is given in the Figure 1.2, where the particular case of l=2 was depicted.

The quantization of the angular momentum determines the quantization of the potential energy of the dipole exposed to the magnetic field:

$$U = \frac{e}{2m_e}L_zB = \frac{e\hbar}{2m_e}m_lB = \mu_Bm_lB$$

where μ_B is the Bohr magneton:

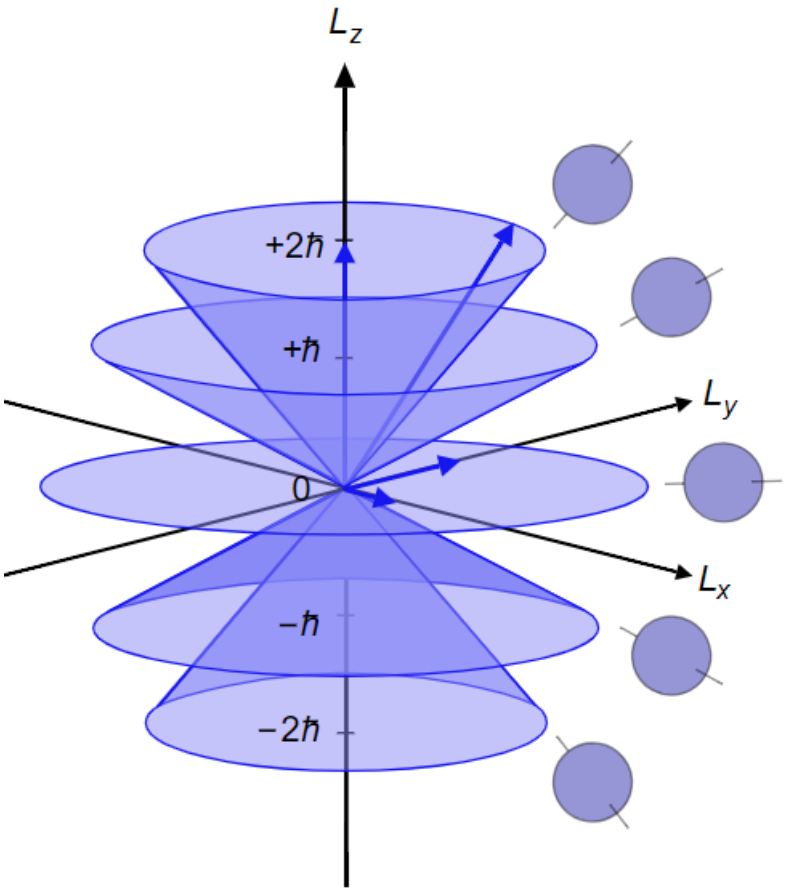


Figure 1.2: Angular momentum representation and its projection along z axis for l=2. Credits: *https://www.wikipedia.org*[7]

$$\mu_B = \frac{e\hbar}{2m_e} = 9.2740154 \cdot 10^{-24} J/T = 5.788382 \cdot 10^{-5} eV/T$$

Based on this assumptions, the difference in energy with regard to the zero field level is:

$$\Delta E = m_l \mu_B B$$

This is known as the Zeeman effect. It tells us that in the presence of a magnetic field the transition energy is splitted into m_l levels uniformly spaced. However, if the spin of the electron is also considered, then the total magnetic momentum should comprise also the spin momentum ($\vec{S}$):

$$\vec{\mu} = -\left(\frac{1}{2}\frac{e}{m_e}\vec{L} + \frac{e}{m_e}\vec{S}\right) = -\frac{1}{2}\frac{e}{m_e}(\vec{L} + 2\vec{S})$$

By considering $\vec{J} = \vec{L} + \vec{S}$ as the total angular momentum (Figure 1.3), then the last equation becomes:

$$\vec{\mu} = -\frac{1}{2}\frac{e}{m_e}(\vec{L} + \vec{J})$$

Once again, we have $\left|\vec{J}\right| = \hbar\sqrt{j(j+1)}$, for $j = l \pm s, l+s-1, ...|l-s|$, and for the projection:

$J_z = m_z\hbar$, $m_j = -j, -j+1, ..., +j$. In the case of the electron, where the spin number is $s = \frac{1}{2}$ we will have $j = l \pm \frac{1}{2}$.

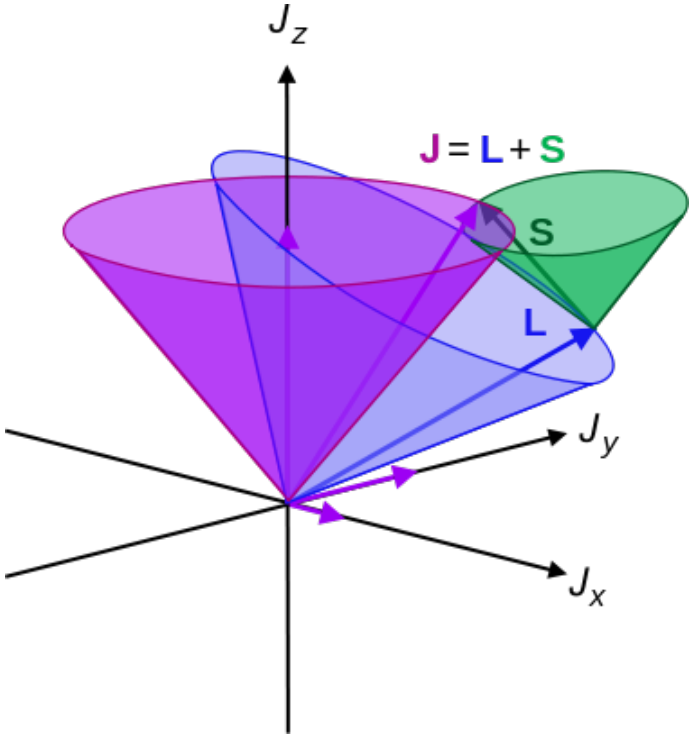


Figure 1.3: Vector cones for the total angular momentum $\vec{J}$, spin momentum $\vec{S}$ and orbital momentum $\vec{L}$. Credits: *chem.libretexts.org*[4]

As a consequence, the difference of energy with regard to the zero field state will take the form:

$$\Delta E = \mu_B g m_j B$$

where g is the Landé factor, and is given by:

$$g = \frac{3}{2} + \frac{s(s+1) - l(l+1)}{2j(j+1)} \tag{1.1.1}$$

The conclusion is then that by paying attention to the spin of the electron the transition energy will be splitted not into m_l levels, but m_j levels. The phenomenon is known as the "anomalous Zeeman effect", just because at the time of observation it was unaccountable. According to the selection rules of transitions, we may have only $\Delta m_j = 0, \pm 1$.

- $\Delta m_j = 0$ (π transition): in a transverse view (perpendicular to $\vec{B}$) the emitted radiation is linearly polarized in a direction parallel to $\vec{B}$; in a longitudinal view (parallel to $\vec{B}$) the transition is forbidden (see Figure 1.4).
- $\Delta m_j = \pm 1$ (σ transition): in a transverse view the emitted radiation is linearly polarized in a plane perpendicular to the field; in a longitudinal view the polarization is circular.

In order to explicit the Zeeman effect we may take a look to the transition ${}^1D_2 \rightarrow {}^1P_2$ (or, equivalently, from (L=2,S=0,J=2) to (L=1,S=0,J=1)) of Cadmium (Figure 1.5).

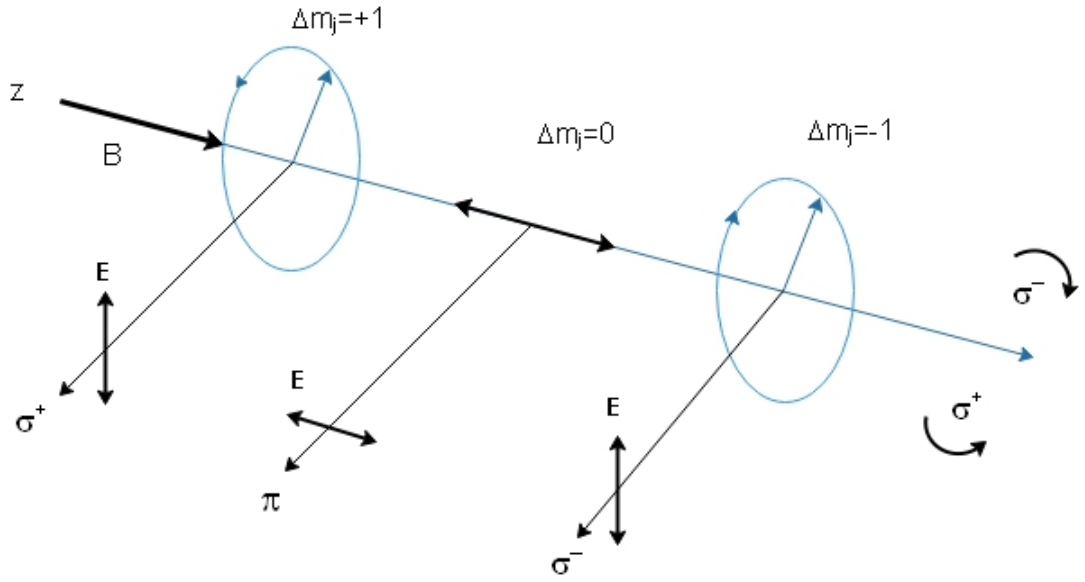


Figure 1.4: Intuitive reprezentation of the polarization in the Zeeman effect. The reference frame is arbitrary chosen and the type of polarization should be considered only with regard to the orientation of the magnetic field $\vec{B}$. Looking perpendicular to $\vec{B}$ we will see only linear polarization, while looking along $\vec{B}$ we will see only circular polarization.

The energy difference between the transition levels is then:

$$\Delta E = \mu_B(g'm_j^{'} - gm_j)B$$

and in this case we have also $g = g' = 1$. When $\Delta m_j = 0$ we obtain the π transition which is equivalent to the red ray of Cadmium without magnetic field ($\lambda = 643.8nm$). On the other hand, when $\Delta m_j = \pm 1$ we have a σ transition:

$$\Delta E = h\Delta\nu = \pm\mu_B B = \pm\frac{eh}{4\pi m_e}B$$

$$\Delta\nu = \pm\mu_B B = \pm\frac{e}{4\pi m_e}B = \pm 1.3995 \cdot 10^{10}B$$

Concerning the separation of the spectral lines, generally we have:

$$\Delta\lambda = 4.7 \cdot 10^{-13}B\lambda^2$$

where λ is expressed in Angströms and B in Gauss. Thus, if $\vec{B}$ is very weak, it will become impossible to observe directly the splitting of the spectral lines. But observing the spectrum through a polarizer can reveal the presence of circular or linear polarization, hereby of all the spectral lines.

Therefore, translated into Stokes parameters, the Zeeman effect can give us access to the value of the projection of the magnetic field of a star on the plane of the sky and along the line of sight (Figure 1.6). It can be shown [41] that for a field $\vec{B}$ making an angle γ with the line of sight, we will have:

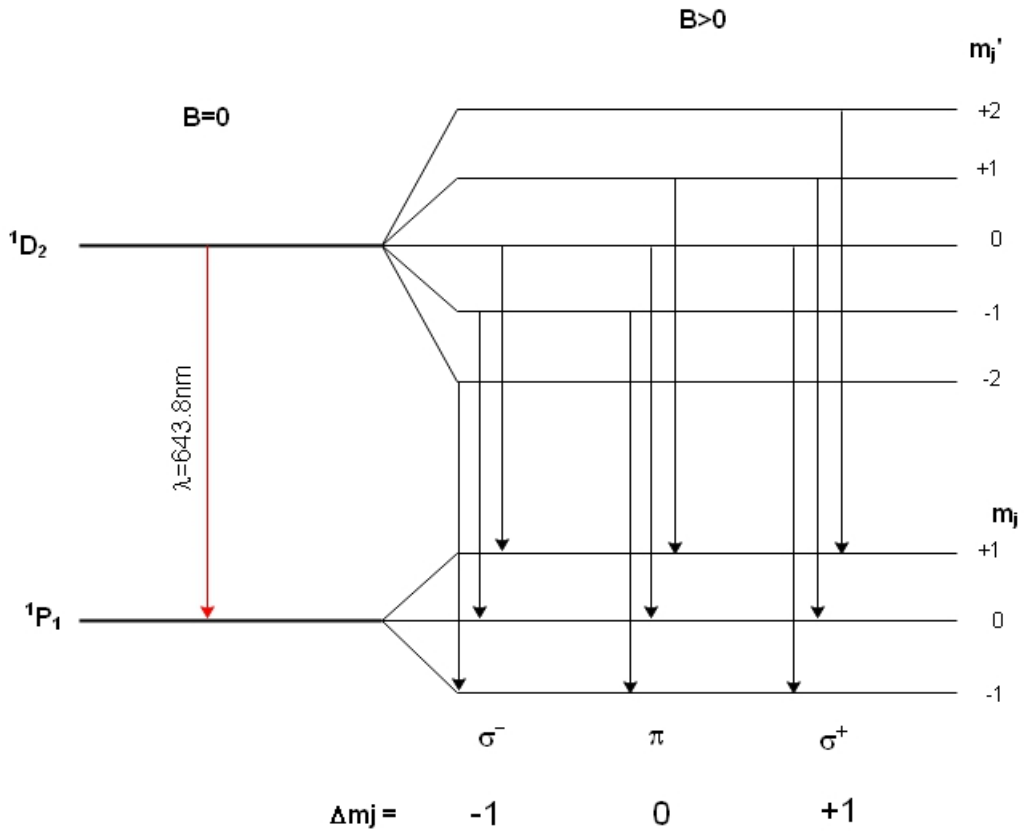


Figure 1.5: Transition ${}^1D_2 \rightarrow {}^1P_2$ for Cadmium with "normal" Zeeman effect (S=0)

$$\begin{cases} Q \propto B^2 \sin^2 \gamma \\ V \propto B \cos \gamma \end{cases}$$

As a consequence, measuring V will give us access to the projection of B along the line of sight ($B \cos \psi$) while with Q we will have access to the projection on the plane of the sky ($B \sin \psi$). It must stress here that in the Zeeman effect only Q and V parameters are concerned, following the geometry exposed in the Figure 1.6. The U parameter is not present in this geometry. Generally, the transverse component (linear) is fainter than the parallel component (circular), reason for which the linear component is hard to be measured by a selective approach. Using then a birefringent spectropolarimeter, which requires no selection process, but provides simultaneously access to all four Stokes parameters it will be of paramount importance in the study of the Zeeman effect and so in the determination of the magnetic fields of stars.

It is worthy also to be noticed here that the Zeeman effect takes place only at relatively faint magnetic field. In a very strong one (more than $10^4 G$[1]), it is replaced by the Paschen-Back effect which suppose that the angular momentum $\vec{L}$ and the spin angular momentum $\vec{S}$ are decoupled and more coupled to the magnetic field around which they precess separately. For stars, the Paschen-Back effect is meet mostly for lithium, which requires only 3T to happen. At the same strength of the field Sodium does not undergo neither the Zeeman effect. For this it requires at least 18T. Generally,

[1]Alternatively in literature and here also the magnetic flux density (magnetic induction) is expressed in Gauss or Tesla: $1G = 10^{-4}T$

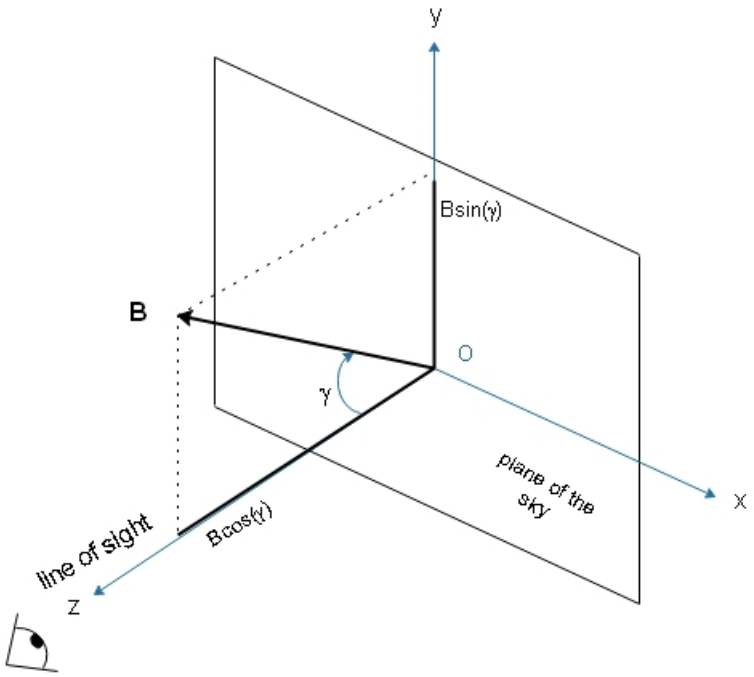


Figure 1.6: Hypothetical orientation of the $\vec{B}$ with regard to the line of sight (z axis). y and x axes can be rotated so that $\vec{B}$ to be comprised in the in the (yz) plane.

the magnetic field of stars exhibits strong variations being related to generating mechanism. For example, the magnetic field of the Sun is of about 1G, but it can reach even 4000G near the massive sunspots. In brawn dwarfs and ultracool dwarfs it was detected even a field of 17 kG. Meanwhile, white dwarfs and neutron stars may exhibit even 10 kG. Among the massive stars, the O-type stars can reach 1 kG, while among Ap and Bp stars the magnetic field varies between 300-300 kG.

1.2 Rayleigh diffusion

Another import source of polarization also of great interest in astronomy, is the Rayleigh scattering. When the light is absorbed by molecules of gas for example, simultaneously the molecule is starting to vibrate and to radiate an electromagnetic radiation of the same frequency and wavelength as the incoming one. Closer the incident beam is to the molecule resonance frequency, stronger the molecule will oscillate, in order to reach a maximum in the UV, where all the molecules have resonances[2]. The shape of the scatterers is not so important in this phenomenon, of more importance being the dimension of the molecules, because the amount of diffused light is related to the diameter of molecules via

$$Q \propto \frac{D}{\lambda} \tag{1.2.1}$$

where D is the molecule diameter.

In order to understand the relation with the polarization, we may use the the situation depicted in the Figure 1.7, where a source of unpolarized light lies on the Oy axis, illuminating the molecule

[2]A very detailed description of the phenomenon can be retrieved in Hecht[16][p.97]

from M.

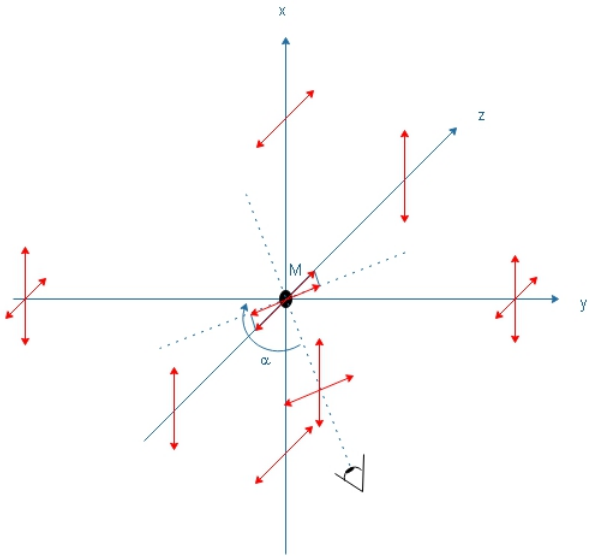


Figure 1.7: Rayleigh diffusion by a molecule: at the base of the phenomenon stays the emission of a vibrating dipole

If this undetermined state of polarization is represented, in a general manner, by two orthogonal vibrations, along Oz and Oy, then we see that the vibration along Oy will force the molecule to vibrate in the same direction and, consequently, to radiate on a plane perpendicular to its vibration. Thus, a vertical vibration will propagate in the plane (zy). Now, if we look to the case of the horizontal vibration, then by the same principle we will retrieve a vibration parallel to the incoming one, in the (xy) plane. Consequently, an observer placed on the x or z direction will receive only one type of vibration, thus a linearly polarized light. On y, just in front of the source, the light will be completely unpolarized, while on an intermediary position (as is depicted in the Figure 3.4) we will receive the projected vibration.

By using the relation (1.3.14), for the degree of polarization, then, in accord with Leroy [26][p.33] we will have a polarization rate that is given by:

$$P = \frac{C^2 - C^2\cos^2\alpha}{C^2 + C^2\cos^2\alpha} = \frac{\sin^2\alpha}{1+\cos^2\alpha} \tag{1.2.2}$$

where α is the angle between the line of sight and the illumination direction of the dipole, and C is the ratio between the amplitude at the observer and the amplitude at the molecule level.

For the Rayleigh scattering, this ratio is proportional to $\frac{1}{\lambda^2}$ which means that the diffused intensity will vary as $\frac{1}{\lambda^4}$. Therefore, smaller is the wavelength, bigger will be the received intensity.

The graphic from the Figure 1.8 shows clearly that at an angle of observation of $\frac{\pi}{2}$ the incoming light is totally polarized, while at 0° is unpolarized.

Another type of diffusion which occurs on free electrons (and thus mostly in hot atmospheres) is the Thomson diffusion. This follows the same mechanism as the Rayleigh diffusion, the degree of polarization being expressed through the same formula as before. The only difference is the fact that the constant C does not depend anymore on the wavelength.

The importance of the Rayleigh scattering in the astronomical polarimetry comes from the fact that measurements at different phase angles allows a relatively fast identification of the phenomenon

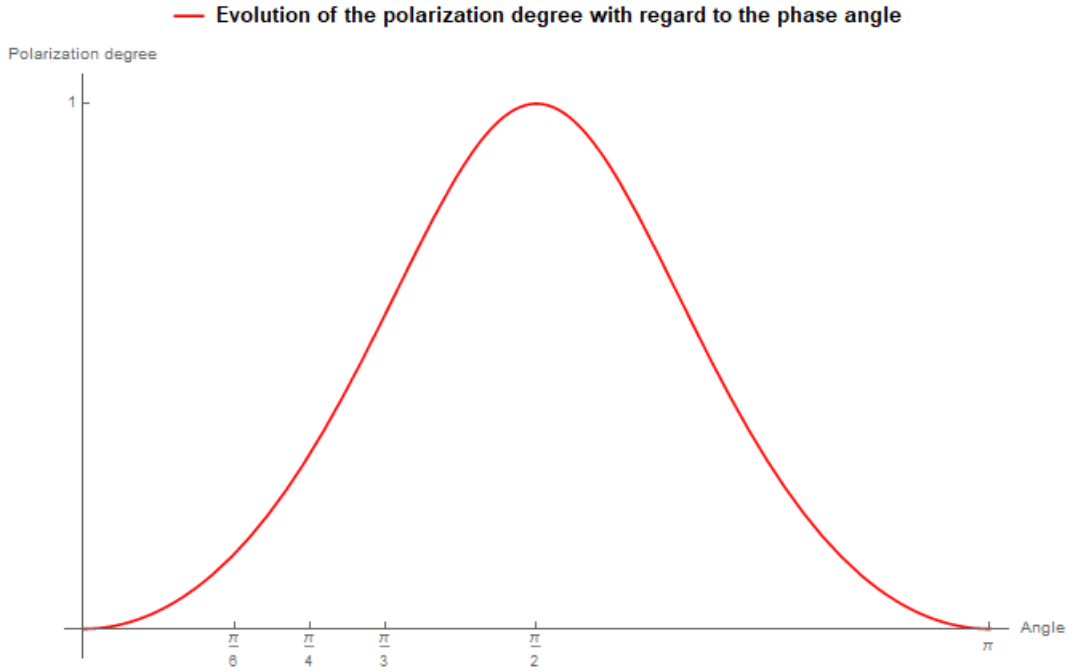


Figure 1.8: Polarization degree in Rayleigh diffusion

through the curve from the Figure 1.8. Observing then at various wavelengths we may have access to the diameter of molecules. Afterwards, comparative studies may revel even their type.

1.3 Hanle effect

Aside the Zeeman effect, the Hanle effect is also a very used phenomenon that can reveal the value of a magnetic field of a star. Its advantage comes from the fact that it works mostly when the magnetic field is relatively weak (compared with the required field for observing the Zeeman effect). At the base of the Hanle effect is the scattering phenomenon which is manifested this time in the presence of a magnetic field. Using the previous image of a vibrating dipole excited by an electromagnetic wave if we add now to the picture a magnetic field, then necessarily we must account also for the Larmor precession around this field. The dipole will have thus two movements: a vibration and a precession. Because the emission should be all the time perpendicular to the direction of movement, will change accordingly.

The best way to see this phenomenon is to consider the situations depicted in the Figure 1.9.

In this simple case the source emits along z axis a linearly polarized light (in the xz plane). In the absence of a magnetic field, we saw that the observer from y will perceive also a linear polarization along x, according to the Rayleigh mechanism. Let us consider now that $\vec{B} \neq 0$. If:

- $\vec{B}$ is along x: no change in the polarization along y (situation b)) is detected;
- $\vec{B}$ is along z: if B is very small or very big then is no effect on y; otherwise the degree of polarization will decrease until reaching 50% (situation c));
- $\vec{B}$ is along y: if B is weak then the linear polarization along y will be tilted. If B is strong then we obtain a complete depolarization.

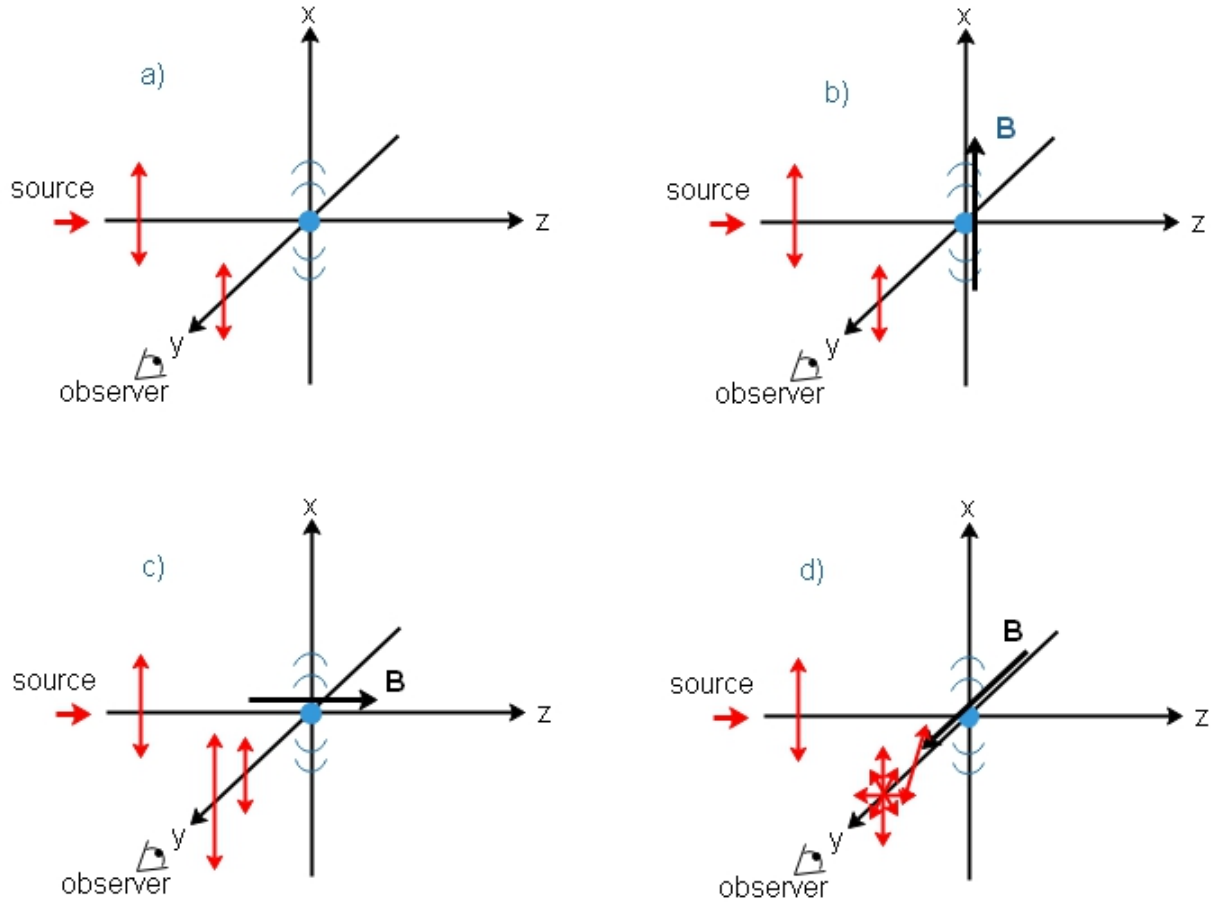


Figure 1.9: Hanle effect for an incoming linearly polarized light and different orientations of the magnetic field $\vec{B}$

The situation along y is determined by the competition between the dumping rate and the precession rate. The resultant vector will describe a rosette pattern like in the Figure 1.10. Here the field strength increases from left to right. The net effect of this variation will be then to continuously rotate the plane of vibration. If the magnetic field is strong enough, there will be no dominant direction of vibration and the light will be completely depolarized.

The polarization degree on z direction is related to the intensity of the magnetic field on z by a relation derived from the quantum mechanics:

$$P = \frac{P_0}{1 + \frac{e^2}{m_e^2 c^2} g^2 \tau^2 B^2}$$

where P_0 is the polarization degree in the absence of a magnetic field, g is the Landé factor and τ is the lifetime or the decay time of the particle (about $10^{-8}s$). In practice, the Hanle effect is used to diagnose magnetic fields of the level of a few G and until about $100G$. Thus, for the study of the Sun corona (B¡$100G$) is the best instrument, and observations have shown that the suitable wavelength range is the UV and EUV[34].

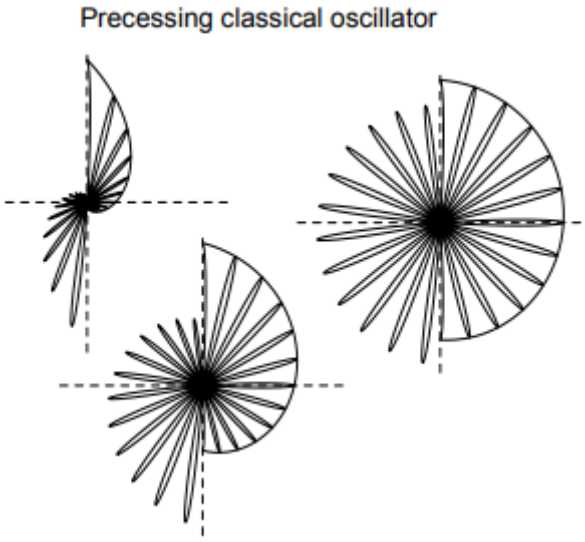


Figure 1.10: Rosette pattern of the trajectory of the momentum vector of the dipole in the presence of a magnetic field. Credits: Stenflo[41]

1.4 Mie diffusion

If the Rayleigh diffusion take place only on molecules, thus for dimensions smaller than λ, the Mie diffusion is for the domain of particles, with a "diameter" bigger than the wavelength of the incident light. The two theories are complementary one to another: below a dimension of about $\lambda/10$ they give the same result. After this level, Rayleigh approach is no more valid and Mie theory should be used.

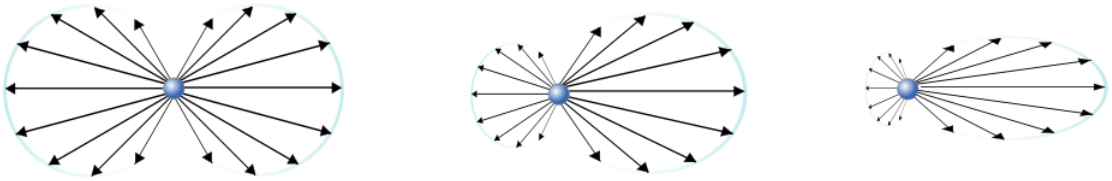


Figure 1.11: Scattering process: left - Rayleigh scattering, middle - intermediary scattering, right- Mie scattering. It is supposed here that the incident wave is coming from the left. WE notice that in the case of Mie scattering, most of the light is diffused towards the direction of movement. Credits: www.sesp.esep.pro

The Mie theory is based on the analytic description of the electromagnetic interaction of light with a particle, using the Maxwell equations. In its pure form is a very complicated tool, reason for which is always preferred the numerical approach which employs algorithms in order to retrieve the most important features. One of the principal inputs for the Mie calculus is the shape parameter (x):

$$x = \frac{2\pi r}{\lambda} \tag{1.4.1}$$

in which r is the radius of the particle and λ the wavelength of the incoming radiation. According to Mie, the degree of polarization will depend, just like in the case of Rayleigh, to the phase angle,

but also to the shape parameter, x. At small values of x, we will retrieve the curve presented in the Figure 1.8, corresponding to Rayleigh process. But after this value, the curve will change drastically from one x to another, presenting even negative values, which signifies that the polarization is radial.

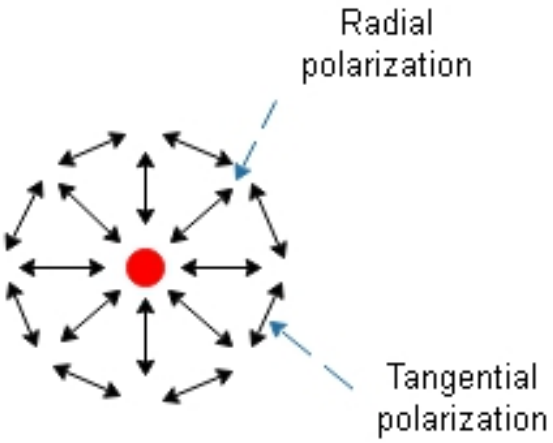


Figure 1.12: When observing extended sources, like the Sun or Jupiter, the polarization is often cataloged as radial or tangential. In the Mie calculus a negative polarization corresponds to the radial direction while a positive one to the tangential case. Relative to the scattering plane, if the vibration of the $\vec{E}$ field is in the scattering plane, we have negative polarization, while to the perpendicular case a positive polarization is associated.

Finding the variation of the polarization degree with the phase angle between the incident ray and the diffused one we may have access to the shape parameter and then, after a comparative and laboratory analyze, even to the type of particles.

It must be stressed here that almost all the time when the scattering theories are used in tandem with the degree of polarization, is the linear polarization that is invoked. And that happens not because the circular polarization would be absent in the observed targets, but because the scattering theories are giving results only for the linear polarization. For the circular polarization we do not have a similar dependency of the degree of polarization with the phase angle[43]. Nevertheless, that does not mean that the circular polarization would be poor in information or would be not involved in scattering processes. On the contrary, the detection of the circular polarization testify usually for the existence of a multiple scattering process, or, hypothetically, for the presence of organic molecules.

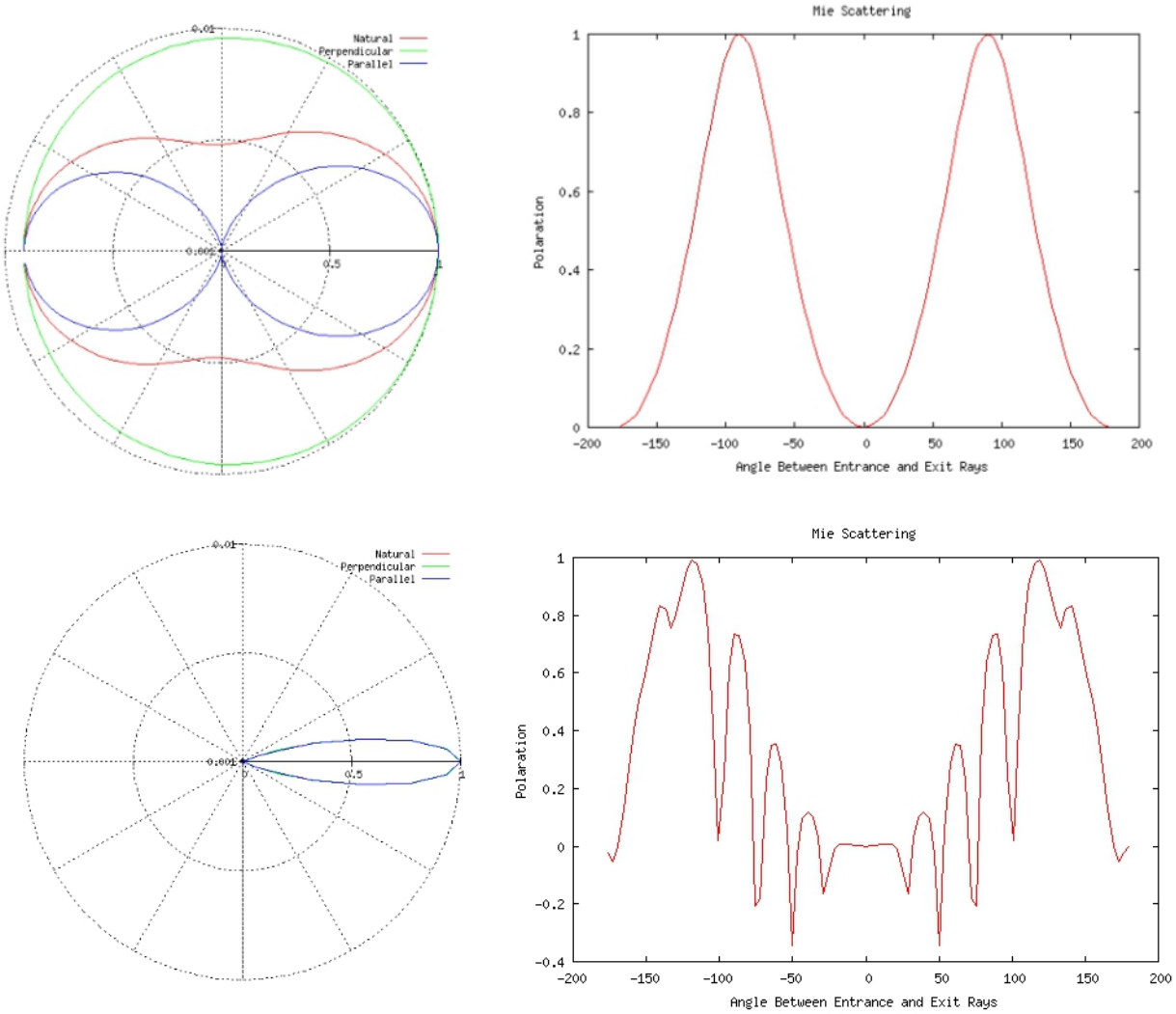


Figure 1.13: Mie scattering simulations for $\lambda = 0.5\mu m$ and $r = 0.02\mu m$ (up) and $r = 1\mu m$(down). On the left side we have the profile of the diffusion pattern (polar view) while on the right side we see the graphic for the degree of polarization as a function of the angle between the incoming and outgoing ray. In the upper part we have the classical situation of Rayleigh scattering: a very small particle which gives the same evolution of the polarization as in the Figure 1.7. In the lower part we have a big particle presenting even negative values of the polarization degree. Simulation conducted with the help of Mie Scattering Calculator (https://omlc.org)

Chapter 2

Polarization in astronomy

The study of the polarization is an important field in astronomy, this attribute of light being able to bring valuable information about the atmosphere, magnetic field, distribution of particles or reflectivity, dimension of particles or even shape and nature. Thus, by the study of the polarization we may have access to the real part of the refractive index, to the distribution of particles in clouds with regard to their size and to the structure of aerosols layers[29]. Also, the polarization study could help estimate the concentration of aerosols from the atmosphere as well as the pressure of the atmosphere. Regarding the type of the perceived polarization, a linear polarization (Q,U) is mainly interpreted as a first order scattering process in the upper layer of an atmosphere or is correlated to a magnetic field, while a circular polarization (V) suggest the presence of multiple scattering processes.

The first studies about the polarization effects in astronomy were conducted for the Moon by Arago, more than a century ago. A that time Arago observed that the shady areas of the Moon were stronger polarized that the brighter zones. Later, this phenomenon was interpreted as a consequence of a multiple-order scattering occurring in the brighter areas, which leads to a depolarization effect.

Further, another important discovery was made by Umov: at large phase angle, the degree of linear polarization is related to the albedo of the celestial body:

$$P \propto \frac{1}{a}$$

where P is the degree of linear polarization and a is the albedo of the observed body. It has to be stressed here that the phase angle does not refer to the phase of the electromagnetic wave, but at the angle between the incident light and the reflected light, as can be observed in the Figure 2.1.

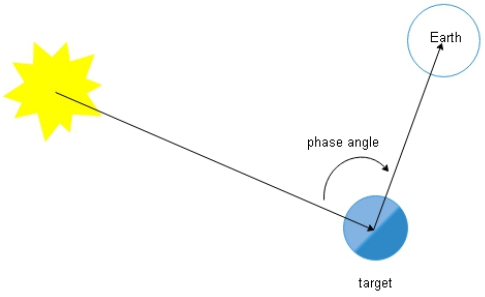


Figure 2.1: Phase angle for an astronomical object

Overall, this was called the Umov effect and it was observed mostly for the atmosphereless bodies of the solar system. In the case of the bodies with atmosphere, and under additional conditions for comets, this effect is observable only at small values of the phase angle.

Another important concept used in polarimetric studies is the inversion angle, which was firstly noticed by Lyot. According to this, for the atmosphereless bodies of the solar system there is a certain angle (α_i), called the inversion angle, at which the polarization angle of the incoming light is switched by 90°. As an example, for the Moon, this α_i is comprised between 15° and 27°. During the time, studies have revealed that there is a correlation between the inversion angle and the wavelength,

even if a precise expression of this dependency was not yet derived. Moreover, the rotation of the polarization angle is not a sharp process, but rather a gradual one, taking place during a certain $\Delta\alpha_i$ interval. Laboratory studies have proved that this interval is strongly related to the composition and the structure of the reflective surface, and consequently, polarimetric studies can be used to describe and to understand the surface of the atmosphereless objects.

For the comets, if there is no important gaseous emission, the concept of inversion angle can be applied, but when the emission starts to become important, this notion cannot be used anymore, and the theory of the polarization in the continuum should be used.

For the celestial bodies with thick atmosphere it is almost impossible to derive a common polarimetric feature, because all the time the polarization of light will depend on the concentration, the size and the shape of the particles responsible for the scattering process.

Another important concept used in polarimetric studies is the opposition effect (opposition surge), described by "the surge in brightness of a particulate medium observed near zero phase angle"[15]. At first, this phenomenon was explained through shadow hiding, which means that for a phase angle near zero the "quantity" of shadows or the shadowed area reaches a minimum in the case of non-Lambertian surface. Then, the phenomenon was also interpreted as a coherent backscattering process (or weak photon localization or time reversal symmetry). According to this theory, if the scattering particles are comparable in size with the wavelength of the incoming light and the distance between particles is bigger than λ, then incoming light and scattered light may interfere, giving birth to an increase of the brightness[9][15].

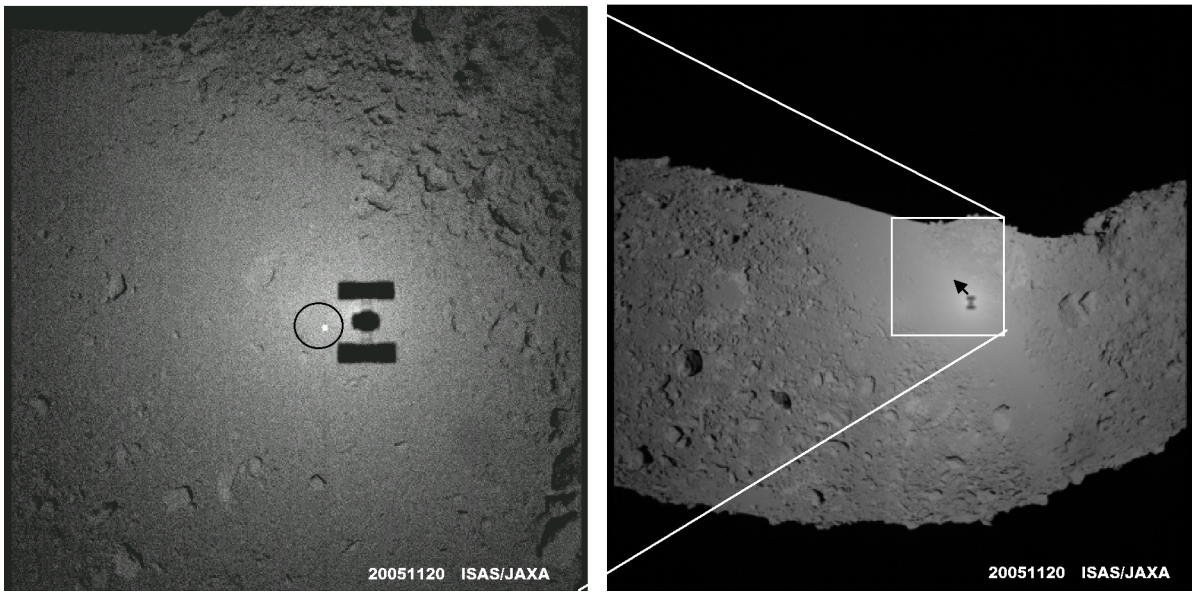


Figure 2.2: Opposition effect near the shadow of the space probe Hayabusa on the surface of the asteroid Itokawa. Credits: ISAS/JAXA

Thus, using the polarization opposition effect, important information about the sizes of particles can be retrieved. Concerning the type of polarization, usually the planets of our solar system exhibit a linear polarization resulting from single scattering processes. The circular polarization has a very poor presence, being noticed mostly for Jupiter and Saturn in the polar regions (phenomenon which is called polar scattering effect, or polar effect), where the degree of circular polarization reaches 0.1% compared with a regular ratio of only 0.001% or 0.01% for the other planets[20].

2.1 Polarization in the Solar System

Excepting the Sun for which the polarization can also be related to Hanle and even Zeeman effects, the rest of the bodies of the solar system are analyzed through the scattering processes: Rayleigh or Mie scattering. And the most common technique is based on the determination of the variation for the polarization degree with the phase angle. This type of curve says immediately if Rayleigh or Mie scattering are involved in the polarization process. Once the process is identified, one can go further towards the determination of the nature and properties of scattering particles.

2.1.1 The Sun

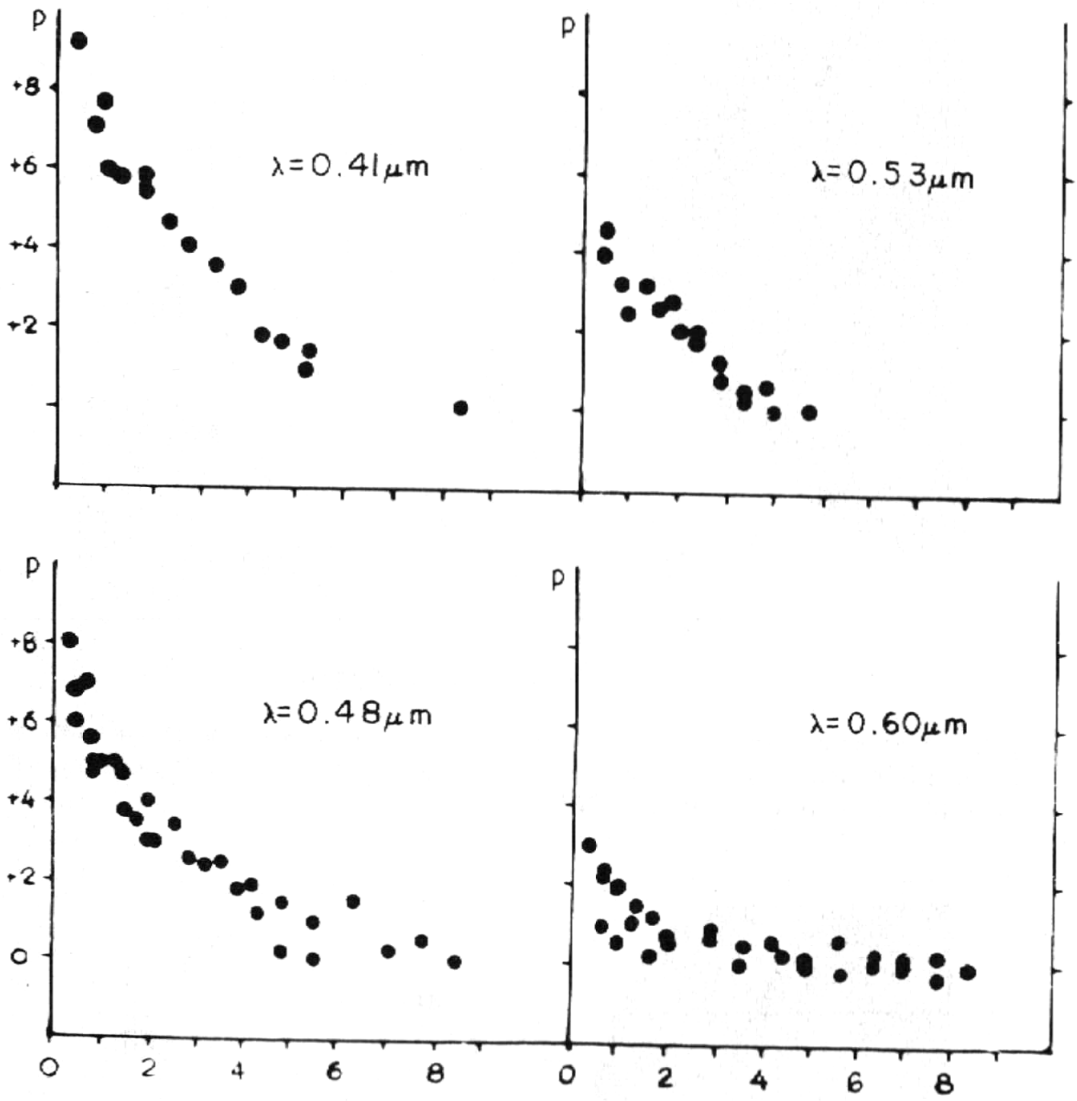


Figure 2.3: Polarization ($p \cdot 10^{-4}$) inside the limb (arcmins) at four wavelengths. Source: [20][p.713]

Stars can be usually considered as ”standard candles” for the nonpolarized light, more indicated to be used as calibration targets than an artificial source. And the main reasons for this is the distance and the symmetry[42][p.27] which nullify the resultant polarization over the disk. But not the same

is the story with the Sun: in this case we have access to a resolved image and in this way local features of the polarization and magnetic field can be measured.

Studies started even from the years 1940s have reveled a relatively weak polarization in the limb which varies with the distance inside the limb and with the wavelength. As an order of magnitude, in the visible spectrum, the maximum was found around $\lambda = 0.41\mu m$, where the degree of polarization reaches $9 \cdot 10^{-4}$.

The main source of this polarization is a combination of Rayleigh scattering on molecules and Thomson scattering on electrons occurring in the photospheric environment. Concerning the magnetic field diagnose, the most suitable instrument is the Hanle effect. The Zeeman effect remains poorly resolved in a low magnetic field and because of this is mostly used in the study of sunspots, where the magnetic flux reaches kG.

2.1.2 Venus

One of the first case studies of polarimetry in astronomy was conducted by Lyot in 1929, through the observation of Venus. Despite the rudimentary apparatus, the results of Lyot have never received important corrections. Its measurements can be observed in the Figure 2.4 where the degree of polarization was plotted as a function of a phase angle for different wavelengths. The entire spectrum is dominated by a linear polarization oriented in the plane described by the Sun, Venus and the Earth. The only exception is in the UV region, where a variable orientation was measured with a preference for the transverse direction.

Currently there is no total agreement about the causes for the "strange" behavior registered in the UV. Variations in altitude for the top layer of clouds, variation of the albedo and changes in the mean-free-path of photons have been advanced as possible explanations.

During the time, the only major improvements of the Lyot measurements were the extension of the phase angle range and the follow up of the variation for the degree of polarization at different wavelengths.

The results for degree of polarization, presented in the Figure 2.5, follows closely the multiple scattering calculations for particles with a radius near the value of the wavelength[20]. The measurements of the polarization allowed the development of a model for the atmosphere of Venus: the relatively close fitting between the polarization of the planet and the polarization obtained by scattering on droplets of about $1.2\mu m$ have suggested that a good candidate for these droplets is the sulfuric acid. Also, the rollover of the direction of polarization in UV could be explained by the presence of a Rayleigh component due to the existence of very fine layer of gas atop of the atmosphere.

We should underline here the tiny values of the polarization, ranging between 0% and 10%, and which suggest that any attempt to build a spectropolarimeter should refer to this level of sensibility.

2.1.3 Mars

The polarimetric study of Mars know a full advancement in the years 1970s. Almost all the studies used the assumption that the atmosphere of Mars is thin enough to be neglected in the process of scattering, at least during the good weather conditions, and above the UV spectral range. Aside the estimations for the atmospheric pressure who suggested a value of about 5 mb at the surface, another reason to neglect the contribution of the atmosphere was the fact that the evolution of the polarization with the phase angle was almost the same as for the Moon. The only differences appeared when weather conditions started to change.

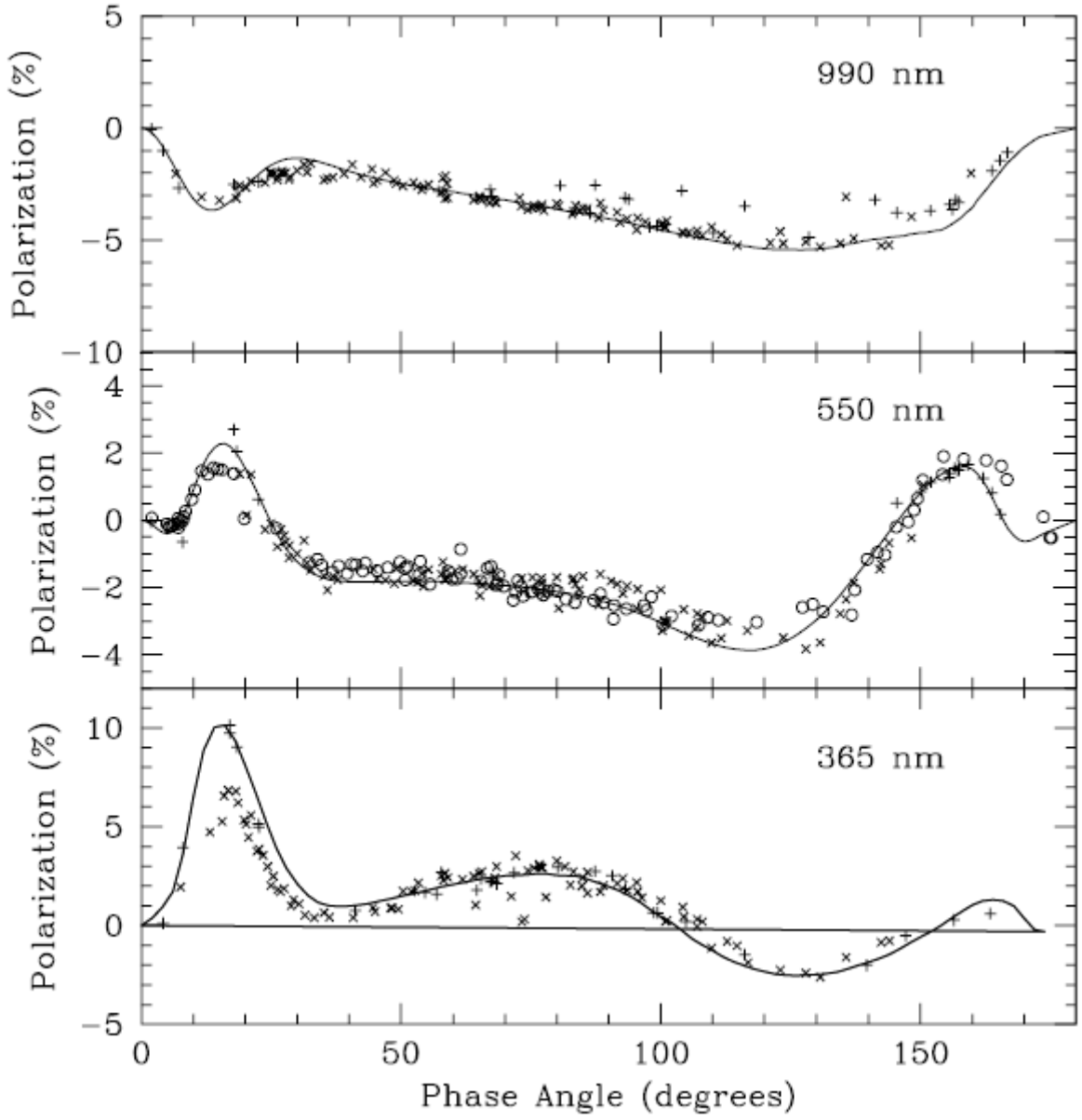


Figure 2.4: Polarization of Venus for three wavelengths: circles (Lyot, 1929), crosses (Coffeen and Gehrels, 1969), plus signe (Dolfus and Cofeen, 1970), continuous line (Bailey). Credits: Bailey [10]

Nevertheless, in time astronomers focused also on regional details: elaborate studies were conducted for white clouds, for blue clouds at the limbs, for yellow haze, etc.[20][p.567]. Mostly, the researches were based on Rayleigh and Mie scattering, the maximal value of the degree of the polarization being of about 2.5%. Among the most important results of the polarimetric studies of Mars one can count the evidence for the presence of water ice crystals as well as CO_2 ice crystals in the atmosphere, the development of an optical profile of the atmosphere[18].

2.1.4 Gas giants

The polarimetric properties of the outer planets are studied from very long time. A few similarities were observed but, overall, the conclusions about the scattering medium are still discussed. Before

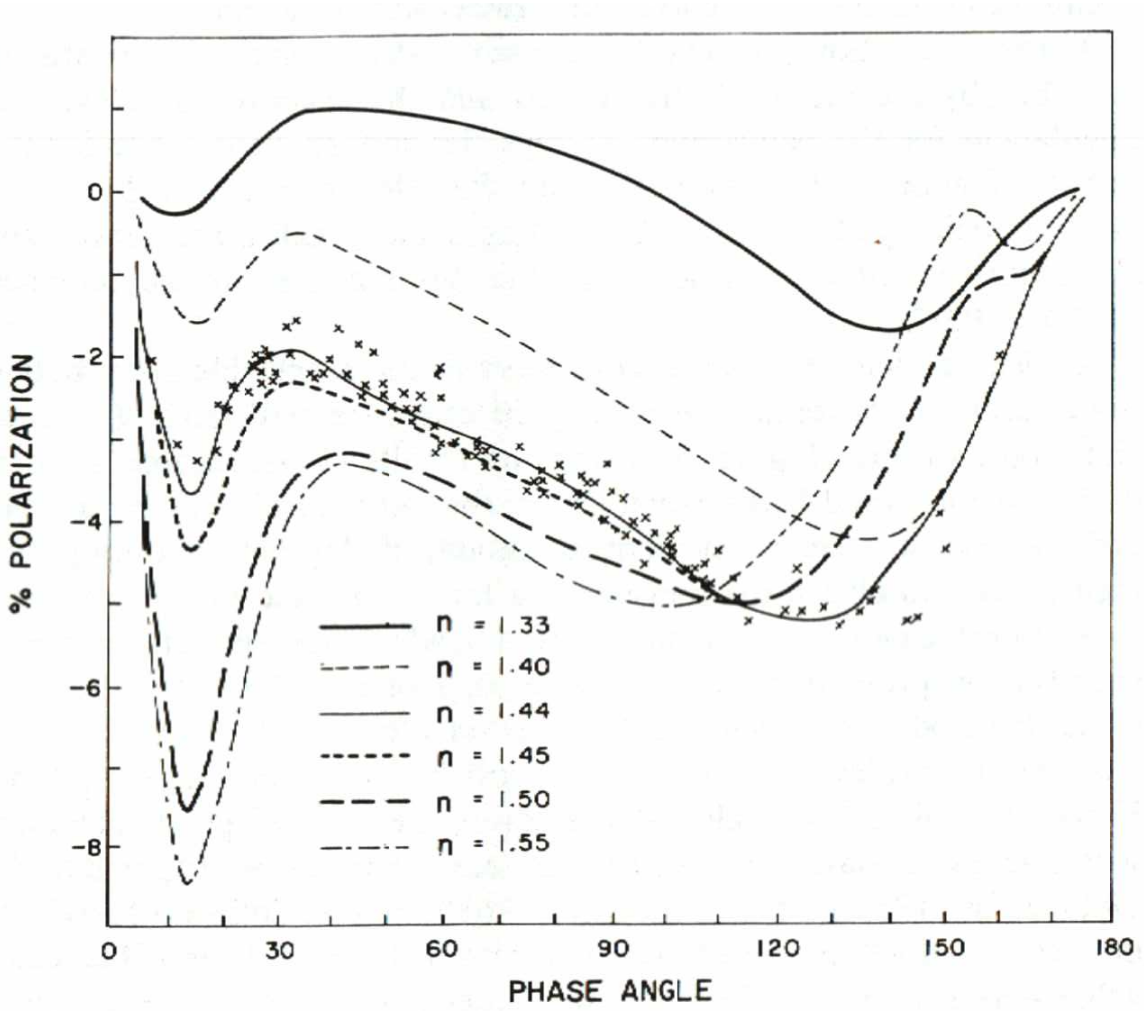


Figure 2.5: Variation of the polarization degree of Venus with regard to the phase angle: crosses (Coffeen and Gehrels, 1969, $\lambda = 0.99\mu m$). Assuming different values of the refraction index n, and a radius of particles between 0.7 and $1.2\mu m$, six comparison curves were obtained. Credits: Gehrels [20]

the space missions the data was lacking a relevant phase angle range. From the Earth Jupiter can be followed for example with phase angle ranging only between 0° and 12°. Thus, aside the resolution, the space missions improved also this aspect. The probes which have monitored the polarization features in the outer solar system have been: Pioneers 10 & 11, Voyager 2, Galileo, Cassini and Huygens[23].

According to gathered data, Jupiter and Saturn present a linear (negative) polarization which increases from the equator to the poles, to reach 7-8$\%$. Generally, the polarization is negative (i.e. tangential), but in the polar regions there is a strong variation with the wavelength, so that at $750nm$ it becomes positive. Also, circular polarization was spotted into the polar regions. According to Kolokolova [23][p.324], the main hypothesis explaining this feature are:

- multiple scattering processes in the outer atmosphere because of the presence of a large quantity of aerosols;
- scattering by aligned non-spherical aerosols;

- scattering by optically active molecules.

Despite all the detailed data about the polarization of the gas giants, the conclusions are controversial. The Mie approaches on spherical scatterers have failed to give a satisfying explanation while the modeling based on ellipsoids is still discussed [26][pp.72-75], [20][pp.568-571], [23][p.321-339].

2.1.5 Atmosphereless bodies

Concerning the atmosphereless bodies of the Solar System, how it was stated before, the researches have started long ago, with the observation of the Moon.

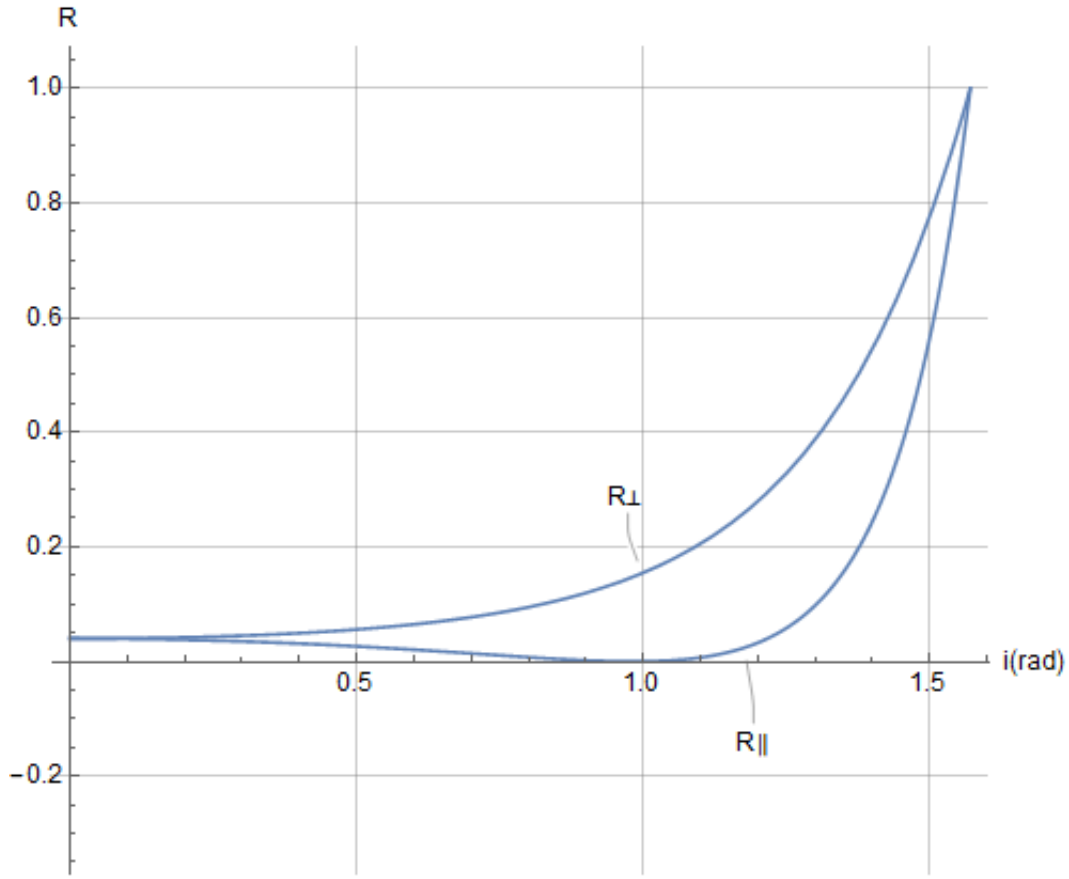


Figure 2.6: Reflectance for a dielectric surface according to Fresnel equations, for parallel and orthogonal vibrations

According to Fresnel equations, we see that the perpendicular component is dominant in reflection (Figure 2.6), and, therefore if a remote body will be only a reflective one, we should observe an orthogonal polarization. The polarization discovered for the Moon was, nevertheless parallel. Thus, based only on polarimetry, the first conclusion was that the surface of the Moon behaves more as a scattering surface than as a reflective one. Using then the Umov effect and a lot of laboratory studies for the polarization obtained with different types of ground samples, scientists deduced the texture of the lunar soil long before the Apollo missions.

A common feature that has to be remembered for the atmospherless bodies of the Solar System is the fact that for small phase angles (generally less than $20°$) the positive polarization (corresponding

to the vibration perpendicular to the plane of incidence) becomes negative. For the moment, there is no full theoretical description of this phenomenon, and it is seen as change of the diffraction pattern with the phase angle as well as a contribution of high order diffusion. The phenomenon was and still is largely studied in laboratory and helped to the identification of the soil composition for the Moon, for Mercury, for the satellites of Jupiter and for many asteroids. For the last ones, the polarimetry could help in the determination of diameter as well as of the composition. Thus, the ISOPHOT measurements of the subsurface infrared emission of asteroids (6) Hebe and (9) Metis [25] revealed the existence of a linear polarization at $25\mu m$, with a degree between 0.2% and 1.6%. Because for a spherical shape the resultant polarization of a non-resolute body should be zero, the manifestation of the polarization could be then a proof for the non-spherical shape. Further analysis can reveal even more details about their structure[24].

2.2 Interstellar medium

Even though is very hard to distinguish between the polarization induced by the interstellar medium and the polarization of the stars itself, astronomers have mapped the polarization of the sky since 1949. The main method employed to differentiate between environmental and stellar polarization is based on the variation of the degree of polarization with the wavelength. The underlying hypothesis is the fact that polarization processes taking place into the stars are governed by the Thomson scattering on electrons, which is independent of the wavelength. Thus, establishing a dependence of the polarization over the wavelength will be enough to prove the interstellar origin of the phenomenon. The empirical low governing this variation with the wavelength is given by:

$$\frac{p(\lambda)}{p(\lambda_{max})} = \exp\left(-1.15\ln^2\left(\frac{\lambda_{max}}{\lambda}\right)\right) \tag{2.2.1}$$

in which

- $p(\lambda)$: is the degree of linear polarization;
- $p(\lambda_{max})$: the maximum degree of polarization;
- 1.15: constant that was replaced in more recent studies by $k = 0.01 + 1.66\lambda_{max}$[26]

Concerning the type, the polarization is dominated by a linear polarization distributed mostly in the galactic plane. In a great measure, the phenomenon is due to the diffusion by the dust, thus being also considered as a proof for the distribution of matter in the interstellar medium. In fact, the polarization is generated by the alignment of the asymmetrical particles in the galactic magnetic field, distribution which acts as a dichroic environment. When the pattern of distribution changes along the line of sight as a result of variations in the magnetic field, the interstellar medium (ISM) can behave even as a birefringent one, capable to yield a circular polarization. Typically, the circular polarization rate is 100 times smaller than for the linear polarization.

A phenomenon worthy to be remembered here is the rotation of the polarization with the wavelength: for many stars it was observed that near λ_{max} a rotation of the plane of vibration take place, probably because of the birefringence of the interstellar medium [Coyne S.J in [20]].

Meanwhile, any variation of the polarization in time is generally interpreted as a variation of the source and is used as a method to distinguish between the interstellar polarization and the source polarization.

2.3 Polarization and the study of exoplanets

For the most part, stars like the Sun exhibits no polarization when the effect is considered across the entire disk, but the reflection of this light by a planet can show a high degree of polarization according to the phase angle and to the scattering processes taking place on the surface or in the clouds.

The polarization mechanism is based on the idea that the flux received from an exoplanet is composed by the reflected flux coming from the sun and the own thermal flux of the planet. Or the reflected flux is polarized due to the scattering.

This reflected flux can be expressed as:

$$F_s(\lambda, \alpha) = \frac{R_p^2}{d^2}\frac{1}{4}S(\lambda, \alpha)\pi F_0(\lambda)$$

where α is the phase angle, R_p is the planet radius, d the distance between the observer and the star, πF_0 the flux received by the planet from the star and $S(\lambda, \alpha)$, the planetary scattering matrix. Assuming that the incoming flux F_0 is not polarized, the corresponding Stokes vector will have only F_0 as element:

$$F_O(\lambda, \alpha) = \begin{bmatrix} F_0(\lambda, \alpha) \\ Q_0(\lambda, \alpha) = 0 \\ U_0(\lambda, \alpha) = 0 \\ V_0(\lambda, \alpha) = 0 \end{bmatrix}$$

Then it can be easely shown[1] that the received Stokes vector, at the level of the observer is:

$$F_s(\lambda, \alpha) = \begin{bmatrix} F_s(\lambda, \alpha) \\ Q_s(\lambda, \alpha) \\ U_s(\lambda, \alpha) = 0 \\ V_s(\lambda, \alpha) = 0 \end{bmatrix}$$

with:

$$\begin{cases} F_s(\lambda, \alpha) = \frac{R_p^2}{d^2}\frac{1}{4}a_1(\lambda, \alpha)\pi F_0(\lambda) \\ Q_s(\lambda, \alpha) = \frac{R_p^2}{d^2}\frac{1}{4}b_1(\lambda, \alpha)\pi F_0(\lambda) \end{cases}$$

in which $a_1(\lambda, \alpha)$ is the total flux scattered by the planet during its orbit, and depends on the wavelength and the phase, while $b_1(\lambda, \alpha)$ is a function reflecting the scattering properties of the planet and which has to be determined in order to understand the scattering process taking place. By measuring the polarization degree for an exoplanet we may have access to the ratio $b_1(\lambda, \alpha)/a_1(\lambda, \alpha)$.

Despite its simple mathematical formulation, the main problem of this technique remains the resolving power of telescopes and the difficulty to separate the faint and polarized flux coming from the planet from the flux of the star. For this a sensibility above 10^{-4} is required. Nevertheless, Sloane[23] cites five cases of measured polarization for exoplanets.

However, the polarimetry can be used combined with other techniques of observations, like radial velocity or transits. And in these cases, a differential polarimetry can be used as a tool for the confirmation of an exoplanet.

[1] The details of this calculus as well as a comprehensive overview of the topic can be found in Kolokolova [23][p.439-457]

More than that, recent studies suggest that the search of the circularly polarized light could be a good instrument for sounding the presence of life, given the fact that homochirals molecules, like most of the amino acids, have the ability to reflect circularly polarized light [35].

2.4 Polarization at high values of the magnetic field

When the magnetic field is stronger than $10^4 G$, then the Zeeman effect is no longer visible, and is replaced by Paschen-Back effect, in which the angular momentum and the spin will be decoupled. This phenomenon is visible for Ap and Bp stars, which are remarkable for their strong magnetic field, usually conceived like a rotating off-axis dipole and rich in elements like strontium, chromium or europium. Moreover, in stars like the white dwarfs, where the magnetic field is even bigger than $10^7 G$, a circular polarization of the entire spectrum can be observed, phenomenon which is called *magnetic circular dichroism*.

2.5 Red giants polarization

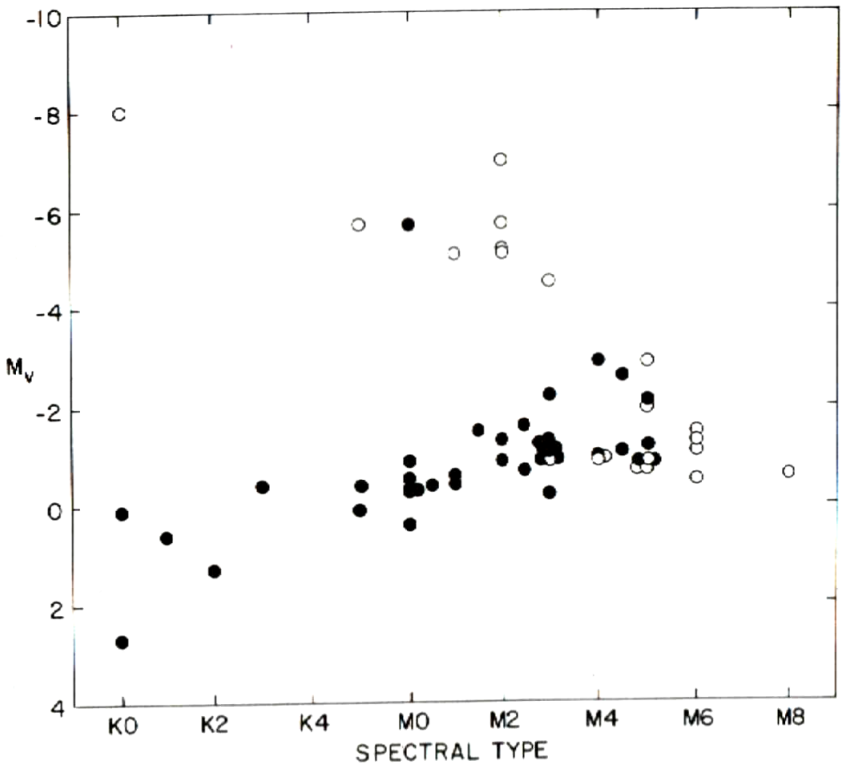


Figure 2.7: Polarized ($\circ$) and unpolarized ($\bullet$) stars in HR diagram. Source[20][p.827]

We saw before that generally stars do not exhibit a certain polarization. But there is an exception: the red giants (late-type stars), which present an intrinsic variable polarization (mostly linear), with a higher value towards UV. Many theories and observations have been conducted about this objects[2]

[2]A comprehensive summary of this work can be found in[20][Shawl, Stephen J: *Polarimetry of late-type stars*]

and until now several explanations about the origin of this polarization have been accepted. Thus, in certain cases it was admitted that the primary source is the scattering from the circumstellar shell (scattering on gas or on dust particles). In other cases, the hot spots model was advanced, which presumes that periodically convective cells surge on the surface emitting a highly polarized and oriented jet of radiation. The degree of linear polarization for this stars may reach even 20-25$\%$, while the circular polarization can be even 10^{-3}.

Chapter 3

Photo-polarimeters and spectropolarimeters in space

Most of the polarimeters and spectropolarimeters used in space are extremely complex instruments, occupying a relatively large volume and employing moving components. The need for turning parts is coming mostly from the fact that polarimeters are not able to provide simultaneously access to all the Stokes parameters. Different configurations are needed in order to cover all the necessary data and to infer the state of polarization. Is a difficult procedure, exposed to a high risk of failure in space conditions. Also, another reason to use moving components is related sometimes to the wavelength range of interest. The limited performance of polarimeters and of detectors with the wavelength is solved by changing the polarimeters and/or the detectors from one channel to another.

In this chapter we will see a few examples of polarimeters used or proposed for space missions along with their working procedures. It is merely an introduction in this problematic, with the only purpose to give a general view about this type of instruments. In detail, for most the them we lack the information, while for others is so huge that is impossible to be summarized here as a simple review.

3.1 ISOPHOT

One of the first polarimeters used in space by the European Space Agency was ISOPHOT, designed for the Infrared Space Observatory (ISO) Mission (1995 - 1998).

According to ESA, "ISO was one of the most sensitive infrared satellite ever launched"[1], aimed to study the interstellar medium between 2.4 and 240 microns. The ISOPHOT imaging photo polarimeter onboard of ISO spacecraft was conceived to cover this entire waveband.

The device used three single detectors which were disposed into three separated subsystems, that could be selected through the rotation of the wheels. According to the opto-mechanical description[21], the three sub-systems were:

- ISOPHOT-C: photopolarimeter with imaging capability working in 40-200 μm range;
- ISOPHOT-P: multi-band and multi-aperture photopolarimeter for the range 30-10 μm;
- ISOPHOT-S: dual grating spectrophotometer for the bands 2.5-5 μm, 6-12μm.

[1]For more details about this mission you may refer to https://www.esa.int/Our_Activities/Space_Science/ISO_overview

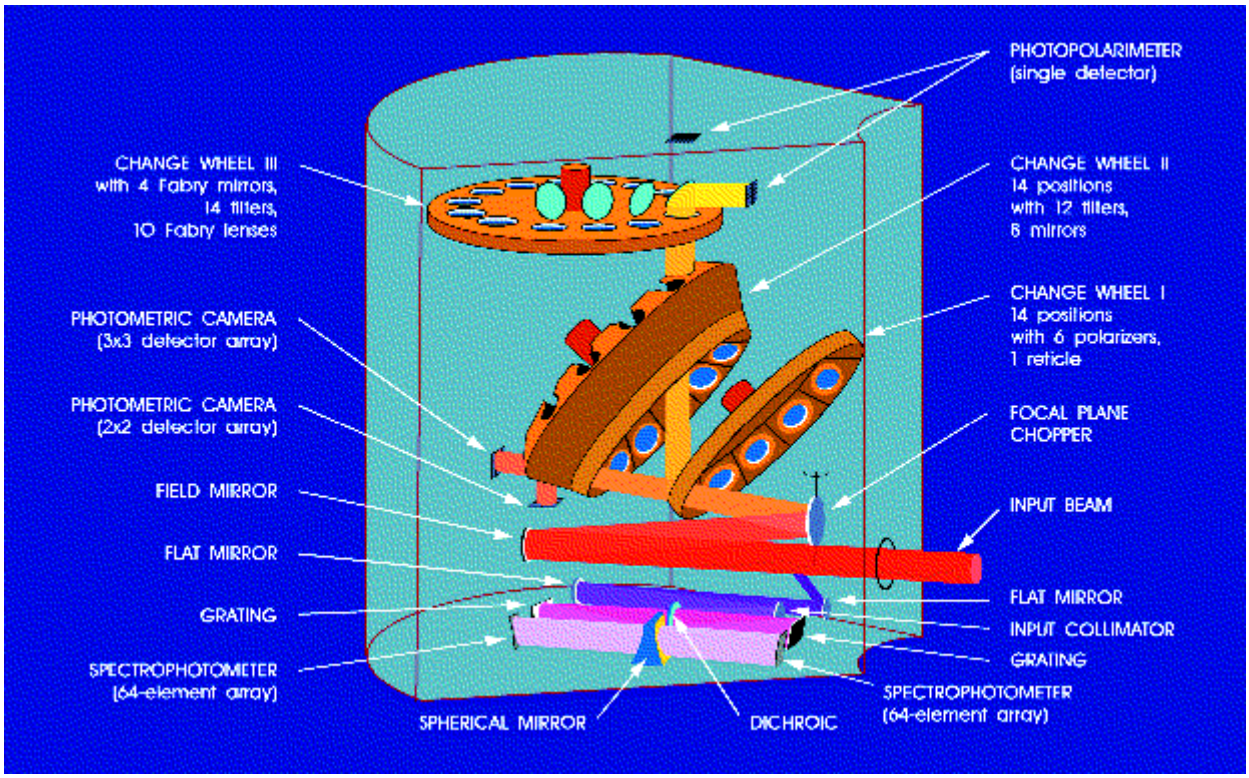


Figure 3.1: The ISOPHOT polarimeter of the Infrared Space Observatory (ISO) (1995-1998). Credits: ESA

The movement was driven by magnetic coils. Three types of detectors were used for this instrument: Si:Ga, Si:P BIB and Ge:Ga.

3.2 CLASP

The Chromospheric Lyman-Alpha Spectro-Polarimeter (CLASP) is an instrument developed by NASA Marshall Space Flight Center together with National Astronomical Observatory of Japan (NAOJ) to observe the polarization in the Lyman-α lines ($\lambda = 121.57nm$) and the Hanle effect in the line core. The device was optimized for sub-orbital missions[19] and operated only during the short interval of five minutes of flight above the atmosphere.

The purpose of CLASP was to measure only the linear polarization until a degree of 0.1% in the Lyman-alpha emission lines in order to determine the magnetic field of the Sun. The modulator of the instrument was a continuously rotating waveplate. In detail, the spectropolarimeter was formed by:

- Cassegrain telescope;
- Rotating half waveplate;
- Slit jaw imager for pointing verification;
- Spectrograph: two spectra are produced for two orthogonal polarization states;
- Analyzer (x 2);

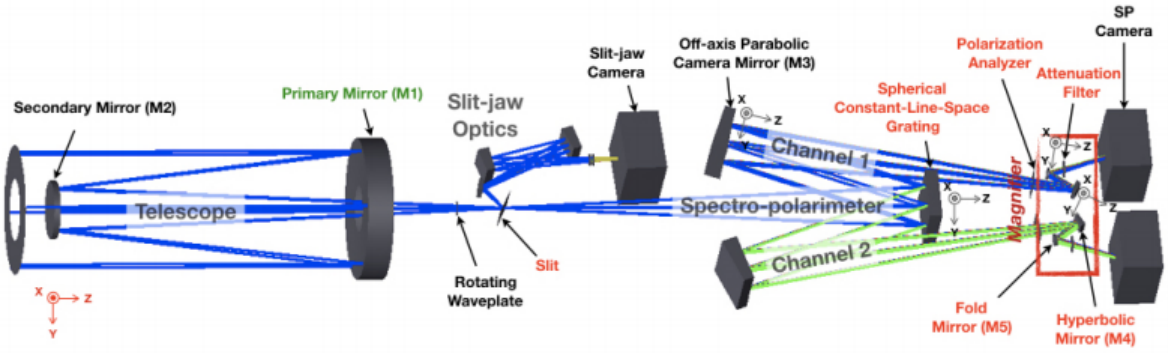


Figure 3.2: Optical design of CLASP Spectro-Polarimeter after Song[39]

- Camera

The movement of the waveplate was synchronized with the camera triggering so that the identification of spectra could be possible. A more detailed discussion about the operating mode of this instrument will be developed in the Chapter 6.

3.3 POLLUX

Even though the James Webb Space Telescope was not yet released, NASA is already planning the next mission that will replace this astonishing construction still under test with another even stronger in performances and dimensions: LUVOIR. Thus, at international level the discussions about this project have already started, and the biggest tasks have been already assigned to teams from the scientific community. The launch date of the project: mid 2030s...

For the moment, there are two possible architectures for LUVOIR (Large UV Optical Infrared Telescope): one with 15.1 m mirror (LUVOIR - A), and another at 9.2 m (LUVOIR - B).

As main structure, LUVOIR at 15.1 m diameter is designed to have four serviceable instruments, while the version at 9.2 m is planned to operate with 3-4 instruments. The primary mirror of the first version - and the only studied for the moment - will be a three-mirror anastigmatic system (TMA), while the secondary mirror will be a fine steering mirror (FSM) located at the exit pupil of the instrument.

The four instruments of LUVOIR-A are: ECLIPS, HDI, LUMOS and POLLUX.

ECLIPS (Extreme Coronagraph for Living Planetary Systems) will enable for the first time the direct imaging and spectroscopy of Earth-like exoplanets. It is conceived to work within 3 bands: UV (200 -400 nm), Vis (400 - 850 nm) and NIR (850 nm - 2.5 μm). And for each channel there will be two deformable mirrors for adaptive optics maneuvers. Also, the instrument will be provided with a series of coronagraph masks, with low-order/out-of-band wavefront sensors and separate imagers and spectrographs.

HDI (High Definition Imager) is the primary imager of the instrument and it will work within two channels: one covering the UV and the visible spectrum (200 - 950 nm), and another for NIR (800 nm - 2.2 μm). This instrument will be able to cover a field of view of 2 x 3 arcminute.

LUMOS (LUVOIR Ultraviolet Multi Object Spectrograph) will be a spectrograph working into 100 nm - 400 nm channel and providing a medium-low resolution multi-object imaging spec-

troscopy. [2]

POLLUX it will be a spectropolarimeter component which will be developed by a European team with support from CNES. The instrument is supposed to cover 3 UV channels (90 -124.5 nm, 118.5 - 195 nm, 195 - 390 nm), with a resolution bigger than 120,000. The device is conceived to give access to UV circular and linear polarization.

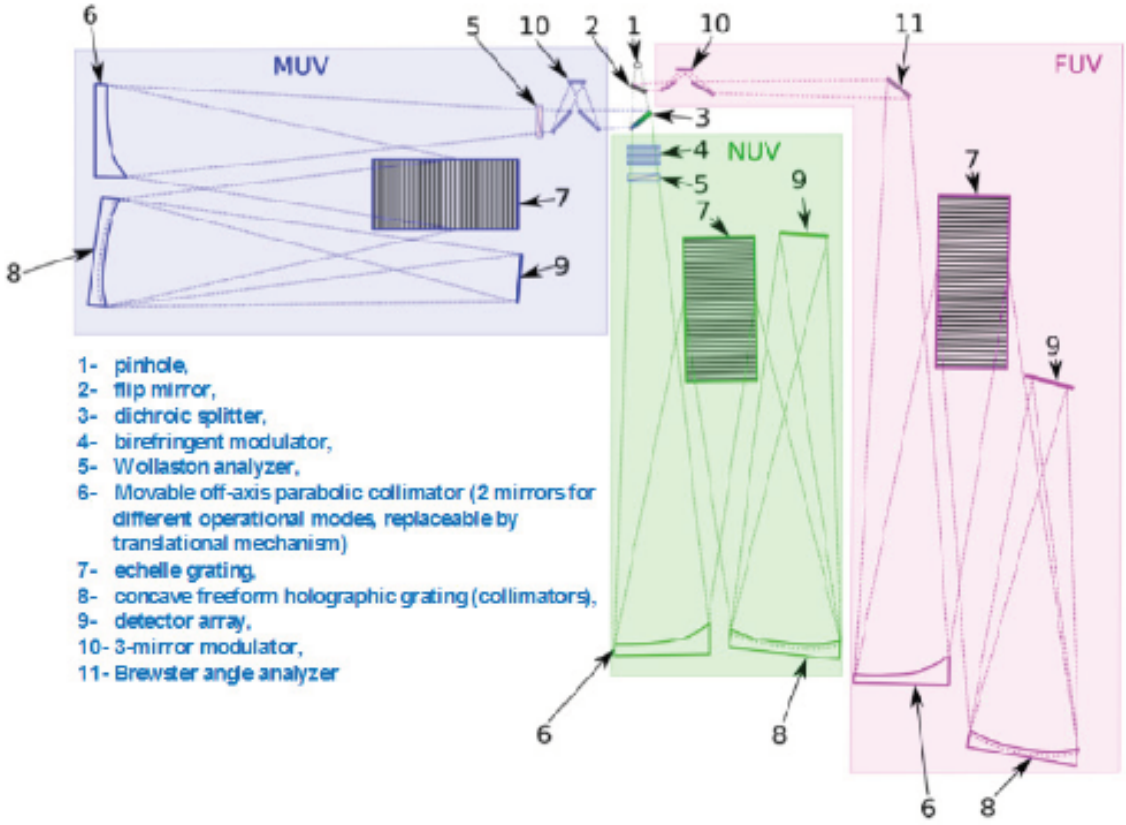


Figure 3.3: POLLUX baseline. Credits: Interim Report[31]

According to the last advancements into this topic[31] each channel will be provided with its own spectropolarimeter: for NUV wavelength range, a birefringent solution is preferred while for MUV and FUV the modulator will be based on the use of mirrors (Figure 3.4). A detailed view of the both solutions is presented in the Figure 3.5. The separation between MUV and NUV is realized using a dichroic splitter, while for the FUV a flip mirror will be used (because there are no dichroic materials available for this spectral range).

Because of this technique, only MUV and NUV channels may work simultaneously. Also, for these two channels, the polarimeters are conceived to be retractable so that a pure spectral survey mode to be accessible. The path difference will be compensated through the movement of the parabolic collimator (number 6 - in the Figure 3.4).

The FUV channel is supposed to use three SIC mirrors, while for MUV the best option is AL+LIF. The analyzer will be a Wollaston prism.

[2] A multi object spectrograph can provide simultaneously spectra for several hundred objects.

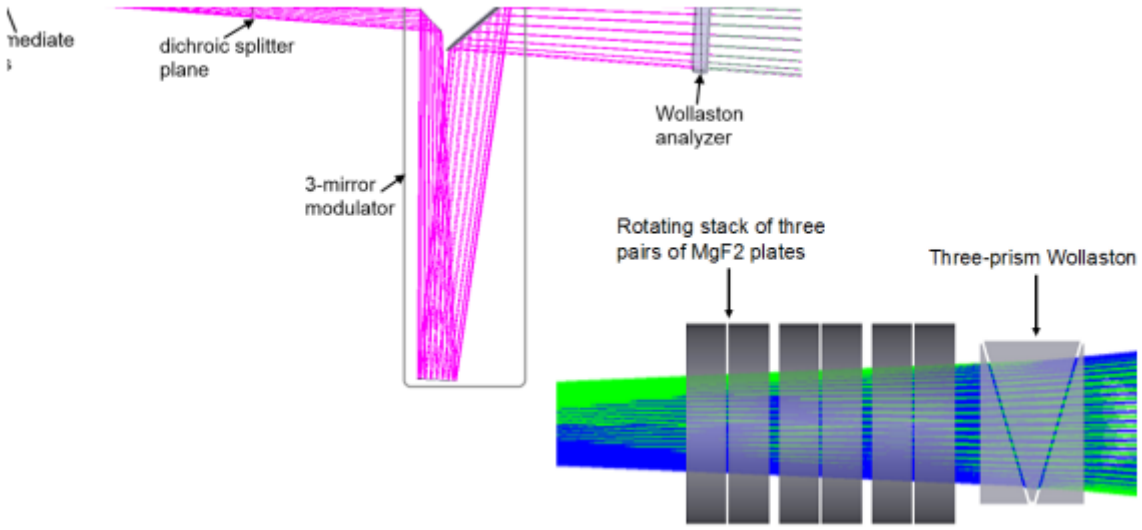


Figure 3.4: Optical design of MUV(left) and NUV(right) units of the LUVOIR project[30]

3.4 SPEX

SPEX (Spectropolarimeter for Planetary Exploration) is a prototype of stable and multi-angle spectropolarimeter aimed to airborne use developed by a Dutch team, able to cover 350-800nm spectral range within 3 channels.

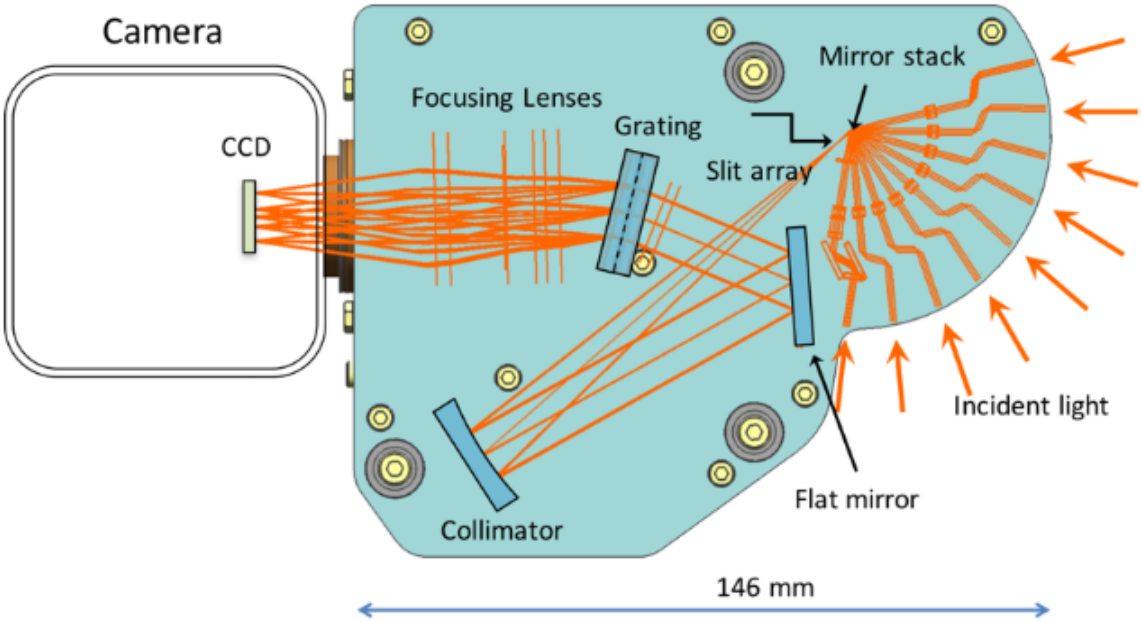


Figure 3.5: Ray tracing for SPEX prototype. Credits: www.sron.nl

According to its description[38], the instrument is able to detect simultaneously the spectral radiance and the degree of linear polarization with a precision of $2 \cdot 10^{-4}$. The design is quietly complex but ergonomic, the total volume being of about 1 liter. In order to cover a large field of

view extremely useful for identifying the scattering particles (-56° to $+56^{\circ}$) the instrument uses 9 entries separated by 14° each, acting overall as a push-broom scanner. The corresponding field of view of a single entry element (or "finger") is of 7° across-track and the images provided by the "fingers" are overlapping at the edges[3].

The polarimetric part is the first in the light pathway, being followed by the spectrometer.

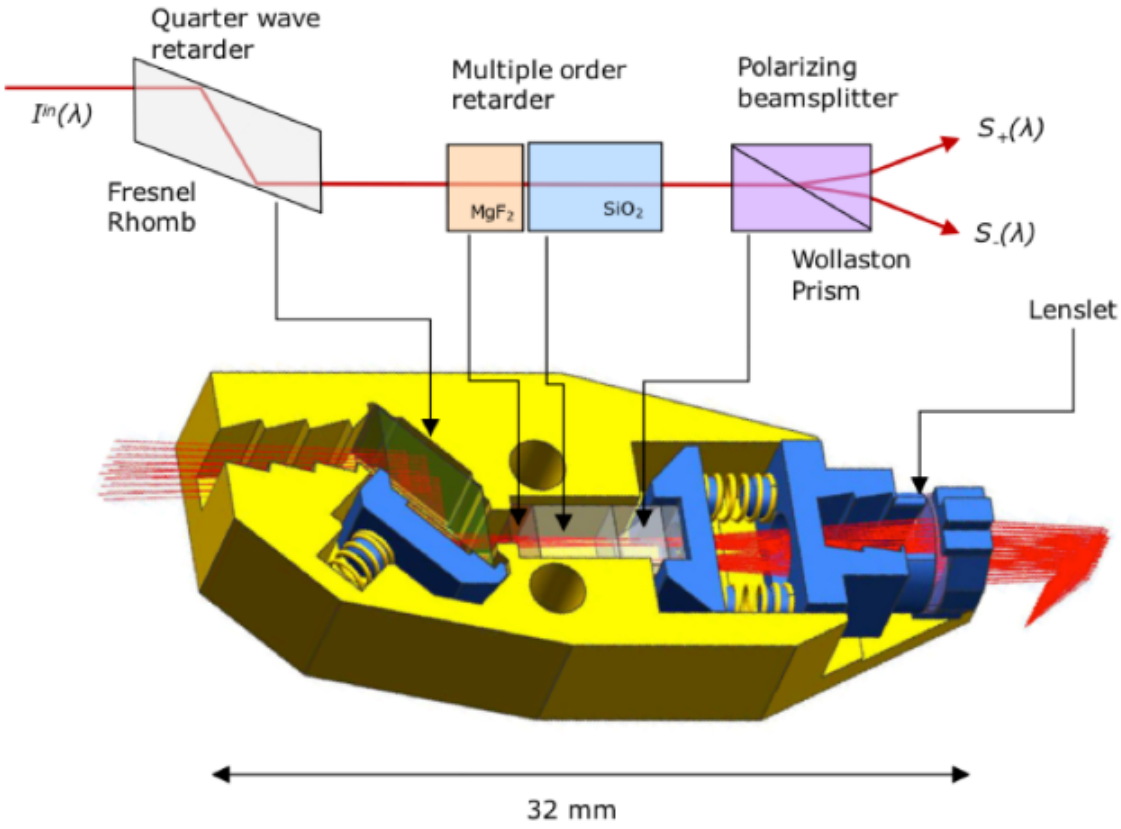


Figure 3.6: Schematics of the polarimeter part. Credits: www.sron.nl

How it can be seen in the Figure 9.5, the light passes at first by a Fresnel Rhomb that will induce a phase difference of $\pi/2$, such that any linearly incoming polarization will be converted into an elliptical one. Then the light passes through a multiple order retarder where will acquire a wavelength-dependent phase. Finally, the orthogonal states are splitted with the help of a Wollaston prism in order to ensure the determination of the incoming intensity.

The measurement of the outgoing intensity will provide access to the degree of linear polarization (DoLP) and to the angle of linear polarization (AoLP) according to:

$$I_{out} = \frac{1}{2} I_{in} \left[1 \pm DoLP \cdot \cos\left(\frac{2\pi\delta(\lambda)}{\lambda} + 2 \cdot AoLP \right) \right]$$

where $\delta(\lambda)$ is the phase induced by the multiple order retarder.

Thus, for each entry of the spectropolarimeter two beams will be produced and directed towards the detector, which means 18 beams for the entire instrument.

[3]Further technical details can be found on the web-site of the developer, *https://www.sron.nl/earth-instrument-development/spex*

Exactly as in our case, the limitations of this design are represented by sensitivity of the birefringent media to the off-axis rays, the chromaticity of the material and by the sensitivity to noise. In the same time, the very big number of elements from the ray paths can be a hindrance in the accurate determination of the throughput of the system.

3.5 Galileo photopolarimeter (PPR)

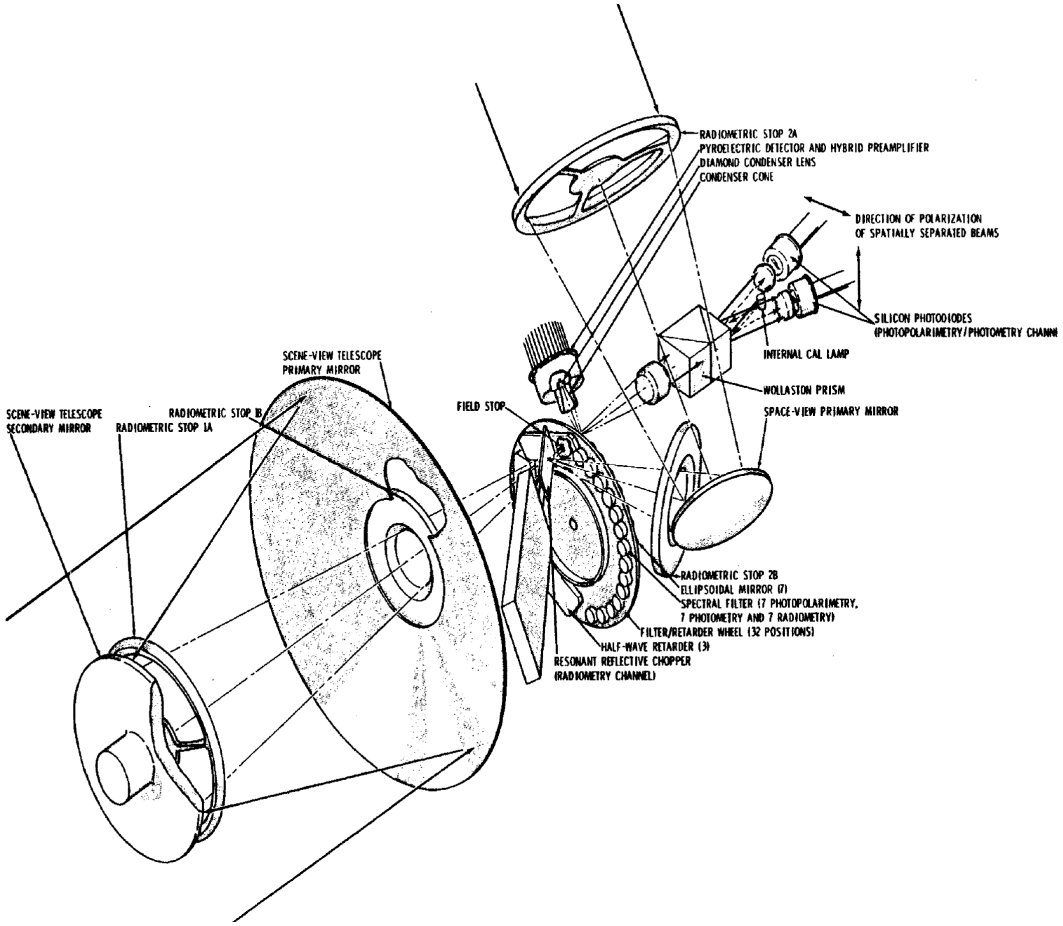


Figure 3.7: Optical design of Galileo photopolarimeter/Radiometer (PPR). Credits: Russell[36]

Galileo was the first mission of NASA aimed for the study of Jupiter and its moons. Launched in 1989, the spacecraft traveled for about 14 years, from which eight were spent on orbit around Jupiter. Among the ten instruments of the payload, the Photopolarimeter/Radiometer (PPR) was designed to measure the linear degree of polarization and the intensity of the light reflected by Jupiter as well as the flux of thermal radiation emitted by the planet. The wavelength range used for polarization measurements was divided in three channels, from 0.410 to 0.945μm, which could be selected with the help of the filter/retarder wheel. According to the instrumental description[36], the light collected by the telescope (see Figure 3.8) was focused on a circular field stop subtending 2.5 $mrad$. Then the rays were passing through a filter/retarder wheel. In the polarimeter mode the light was then directed to a half-wave plate and a spectral filter. By passing then through a Wollaston prism which worked as a polarizing beam-splitter the light was divided into two othogonally polarized outputs beams. With the help of the detector lenses this two beams were finally focused on silicon photodiodes.

The accuracy of measurements obtained with the polarimeter instrument was of $\pm 1\%$, while the sensibility was below 8% for the degree of linear polarization.

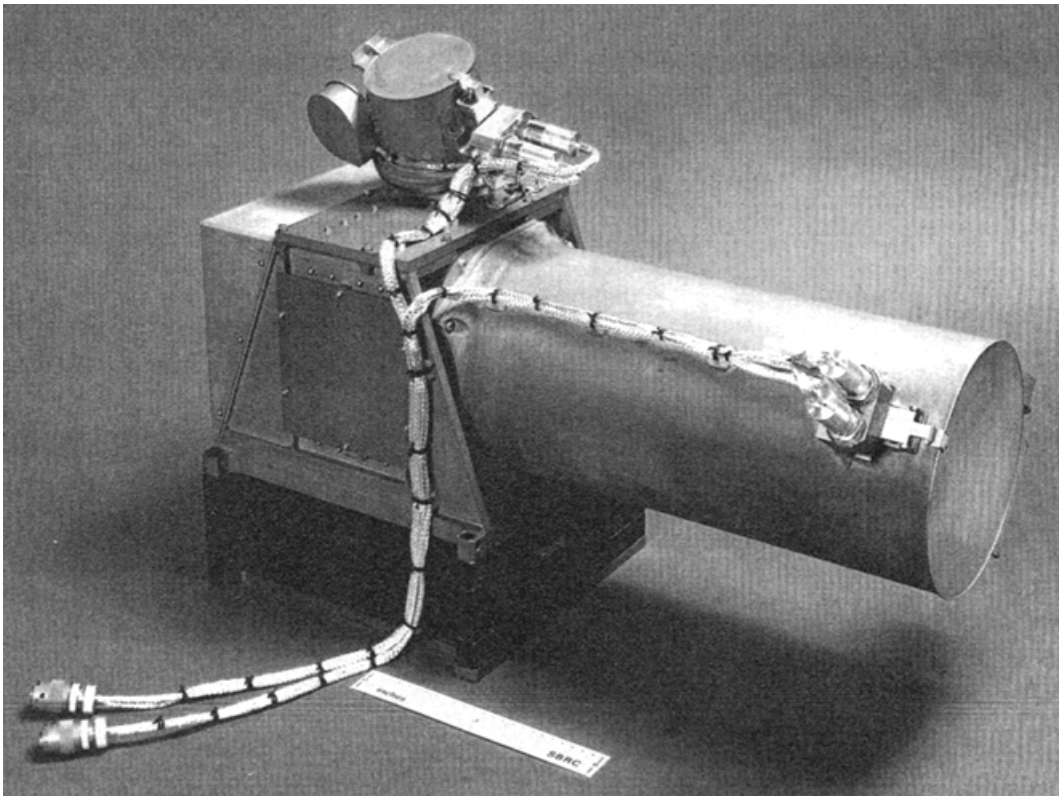


Figure 3.8: Galileo photopolarimeter/Radiometer (PPR) instrument. Credits: Russell[36]

Chapter 4

Mathematical description of the polarization

The analysis of the spectropolarimeter that makes the subject of our research is based on the Mueller calculus and Stokes formalism. Each element of the device will be described through the Mueller matrices and the corresponding output of the system will be calculated with the help of these. For a proper determination of the matrices we will pass in review the main notions and relations of the mathematical description of the polarization. At first, the elementary mathematical approach of the polarization will be revisited. Then the necessary Stokes formalism will be presented. Afterwards, because in this case we will work with Mueller matrices that are not very common in the specialized literature, we will have to derive ourselves. The best way to do this will be through the help of the Jones formalism. Therefore, a few notion and relations from the Jones approach will be used so that, in the end, the researched Mueller matrices to be established.

4.1 The concept of polarization

Besides its complex quantum behavior, light can be treated as well as an electromagnetic transverse wave, i.e. a wave for which the electric or the magnetic field vibrates in a direction perpendicular to the direction of propagation of the wave. Choosing a reference system oriented in such a way that the z axis coincides with the direction of propagation of the wave, $\vec{k}$, while the electric field $\vec{E}$ vibrates in the xy plane (Figure 4.1), we may observe that the oscillation of the electric field may embrace particular forms. The decomposition of the $\vec{E}$ field in the xy plane is given by:

$$\vec{E}(z,t) = \vec{E}_x(z,t) + \vec{E}_y(z,t) \tag{4.1.1}$$

where

$$\vec{E}_x(z,t) = \mathbf{e}_x \cdot E_{0x} \cos(kz - \omega t) \tag{4.1.2}$$

is the field component along x axis, and

$$\vec{E}_y(z,t) = \mathbf{e}_y \cdot E_{0y} \cos(kz - \omega t + \varepsilon) \tag{4.1.3}$$

is the field component along y axis. In this two last relations E_{0x} and E_{0y} are the field amplitudes on x and y, while $\mathbf{e}_x$ and $\mathbf{e}_y$ are the unit vectors describing the two directions. The phase difference between the vibrations is given by ε. Based now on these notations we can observe that a phase difference equal to 0 (or a multiple of π) will result in a so called linear polarization: in this case,

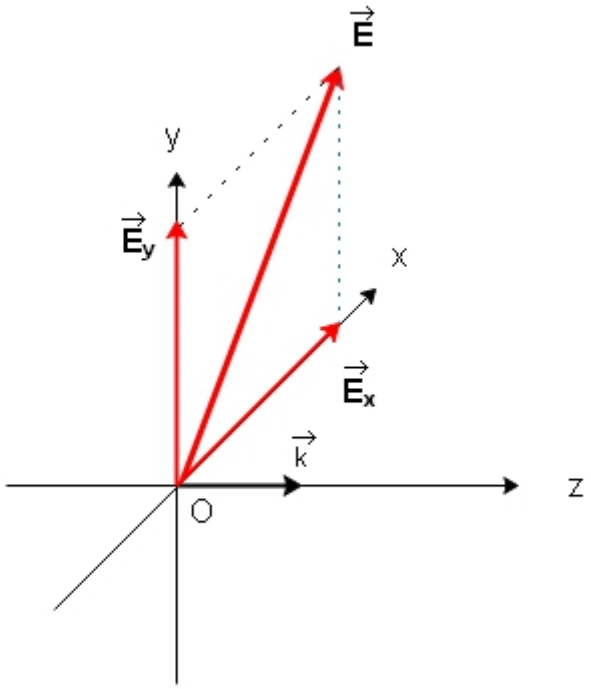


Figure 4.1: Instantaneous view of the $\vec{E}$ field and its decomposition in the plane orthogonal to the direction of propagation of the wave, $\vec{k}$

the projection of the electric field $\vec{E}$ in the (xy) plane will be, for any moment (t) and any value of z, on the same direction (Figure 4.2).

The inclination of the $\vec{E}$ field with regard to x axis will be determined in this case by the ratio $\frac{E_{0y}}{E_{0x}}$. Thereupon, two main subcategories can be described: the *horizontal polarization* and the *vertical polarization*. Thus, the horizontal polarization can be expressed via:

$$\vec{E}_y(z,t) = 0; \qquad \vec{E}_x(z,t) = \mathbf{e}_x \cdot E_{0x} \cos(kz - \omega t)$$

while, the vertical will be given by:

$$\vec{E}_x(z,t) = 0; \qquad \vec{E}_y(z,t) = \mathbf{e}_y \cdot E_{0y} \cos(kz - \omega t + \varepsilon)$$

By playing with the value of the phase difference ε, and of the ratio $\frac{E_{0x}}{E_{0y}}$ others types of polarization can be described. Thus, if the $\varepsilon = \pi$ then $\vec{E}_y$ becomes $-\vec{E}_y$, because $cos(A + B) = cos(A)cos(B) - sin(A)sin(B)$.

Also, if $\varepsilon = \pi/2$ and if $E_{0x} = E_{0y} = E_0$ then the vector $\vec{E}_y$ will become

$$\vec{E}_y(z,t) = -\mathbf{e}_y \cdot E_0 \sin(kz - \omega t + \varepsilon)$$

and totally, the $\vec{E}$ field will describe in time a circle in the (xy) plane, in a clockwise orientation. We say that the wave is *circularly right* polarized (RCP). In the same way, a $-\frac{\pi}{2}$ phase difference will induce a *circular left* polarization (LCP). The entire sitation is depicted in the Figure 4.4 and 4.5.

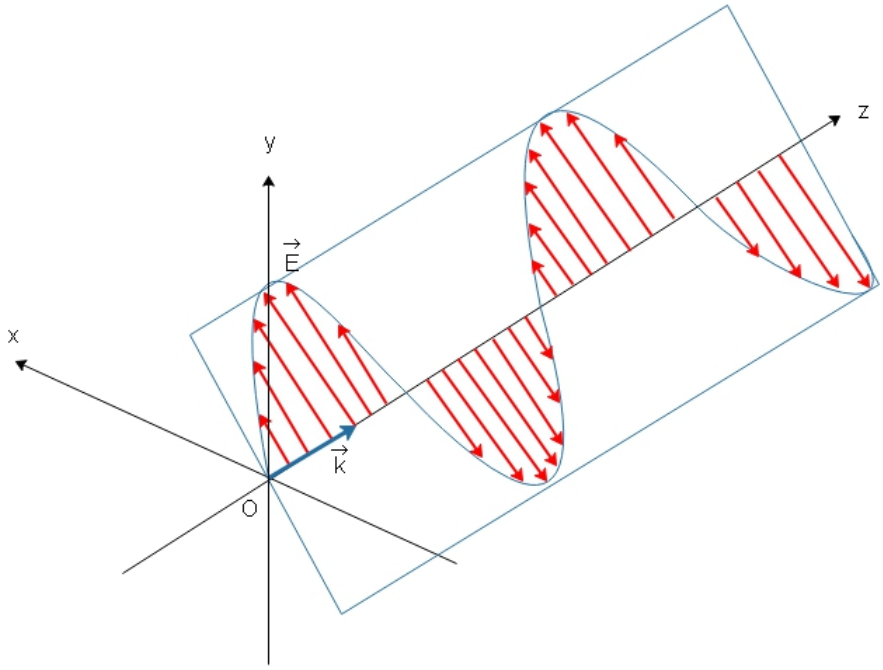


Figure 4.2: Schematic view of the linear polarization: at any moment the electric field vibrates in the same plane orthogonal to (xy) plane. The direction of propagation is given by the $\vec{k}$, oriented along z axis.

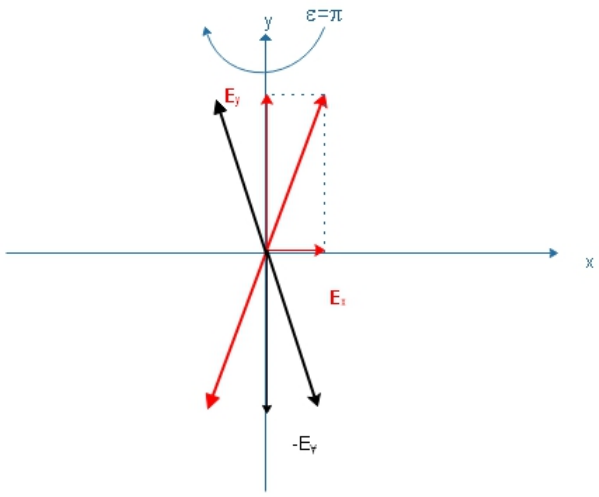


Figure 4.3: A phase difference of π will change the sign of Ey component

In the most general way, if $E_{0x} \neq E_{0y}$ and the phase difference has an arbitrary value, it can be shown that the projection of the electric field $\vec{E}$ in the xy plane describes an ellipse according to the relation:

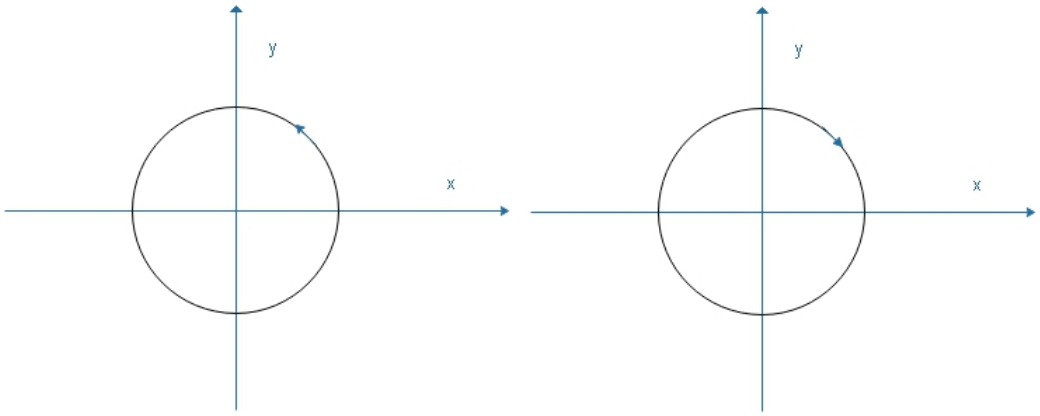


Figure 4.4: Circular left (LCP) and circular right (RCP) polarization. The wave is oriented toward us.

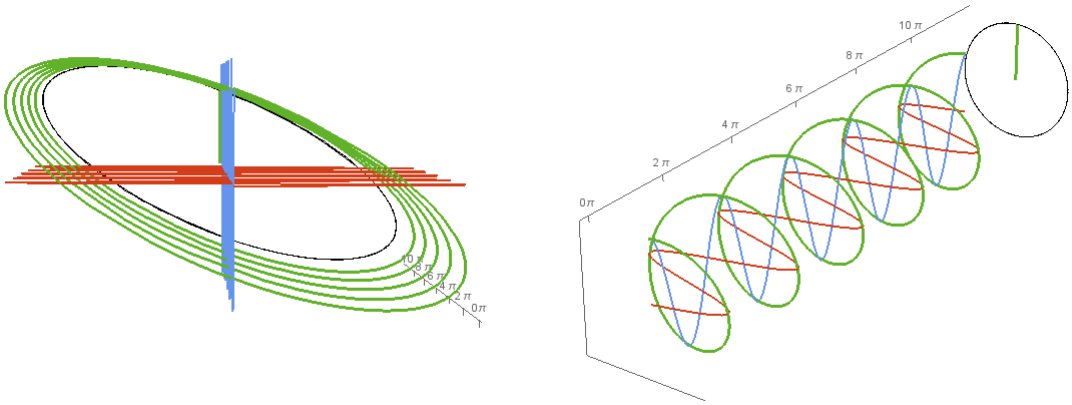


Figure 4.5: Elliptical and circular polarization simulations conducted with Mathematica software. With green we have the trajectory described by the point of $\vec{E}$, while in red and blue we have the two orthogonal components of the $\vec{E}$ field.

$$\left(\frac{E_y}{E_{0y}}\right)^2 + \left(\frac{E_x}{E_{0x}}\right)^2 - 2\left(\frac{E_x}{E_{0x}}\right)\left(\frac{E_y}{E_{0y}}\right)\cos\varepsilon = \sin^2\varepsilon \tag{4.1.4}$$

This ellipse is making an angle ψ with the x axis (the azimuth angle, or the orientation angle), given by the relation:

$$\tan(2\psi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 - E_{0y}^2}\cos\varepsilon, \qquad \psi \in [0;\pi] \tag{4.1.5}$$

Of interest is also the ellipticity angle, defined as the ratio between the semi-minor axis and the semi-major axis, or[1]:

[1]The full deduction of this identities is given the Appendix B: Azimuth and ellipticity calculation. Meanwhile, a more detailed presentation of the polarization parameters can be find in Collett (p.9)[12] and Kliger(p.108 - 116)[22].

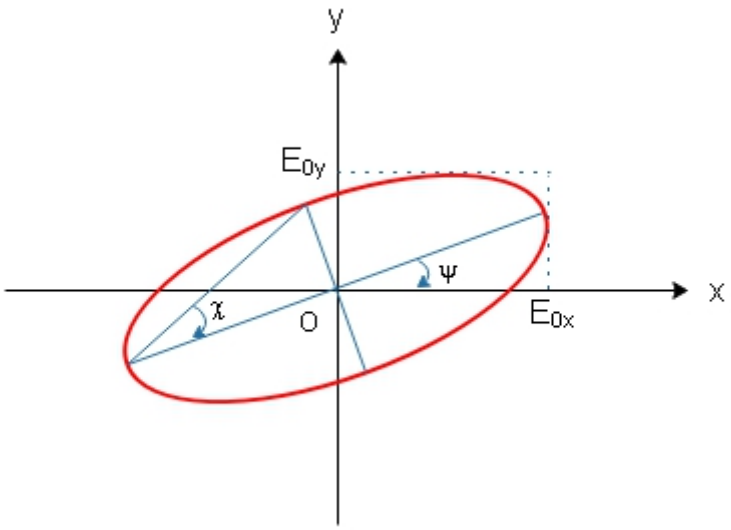


Figure 4.6: Ellipse described by the electric field in the xy plane. We may observe here the azimuth angle, ψ and the ellipticity angle, χ

$$\sin(2\chi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 + E_{0y}^2}\cos\varepsilon, \qquad \chi \in [-\frac{\pi}{4}; \frac{\pi}{4}] \tag{4.1.6}$$

The auxiliary angle of the ellipse is also defined through:

$$\tan(2\alpha) = \frac{E_{0y}}{E_{0x}}, \qquad \alpha \in [0; \frac{\pi}{2}] \tag{4.1.7}$$

Thus, in a general way we may say that the electrical field is characterized by an elliptical polarization, the sens of polarization being dictated by the value of the phase difference. In turn, the auxiliary angle is imposing the ellipticity of the polarization figure, which may vary between a straight line and a perfect circle, as can be seen in the Figure 4.7.

Because of this, it is a common practice to say that the light has generally an elliptic state of polarization, the linear and circular polarization being only particular cases of this.

4.2 Unpolarized light

Normally, the light coming from the sun, from the stars and from most of the artificial sources is not polarized, which means that the electric field of the radiation does not have a preferential direction of vibration. Instead we have a countless superposition of vibrations, summarized over a certain interval of time, which is the integration time of the detector (that can be a photodetector, a CCD or the eye).

If this integration time could be shrunk to the limit of the persistence of the particular vibrations ($\approx 10^{-14}s$, dictated by the temporal coherence) then we would have the possibility to observe all the time a polarized light. But how this is not possible with the current technique, we must be content with the statistical result of the integration over a longer period.

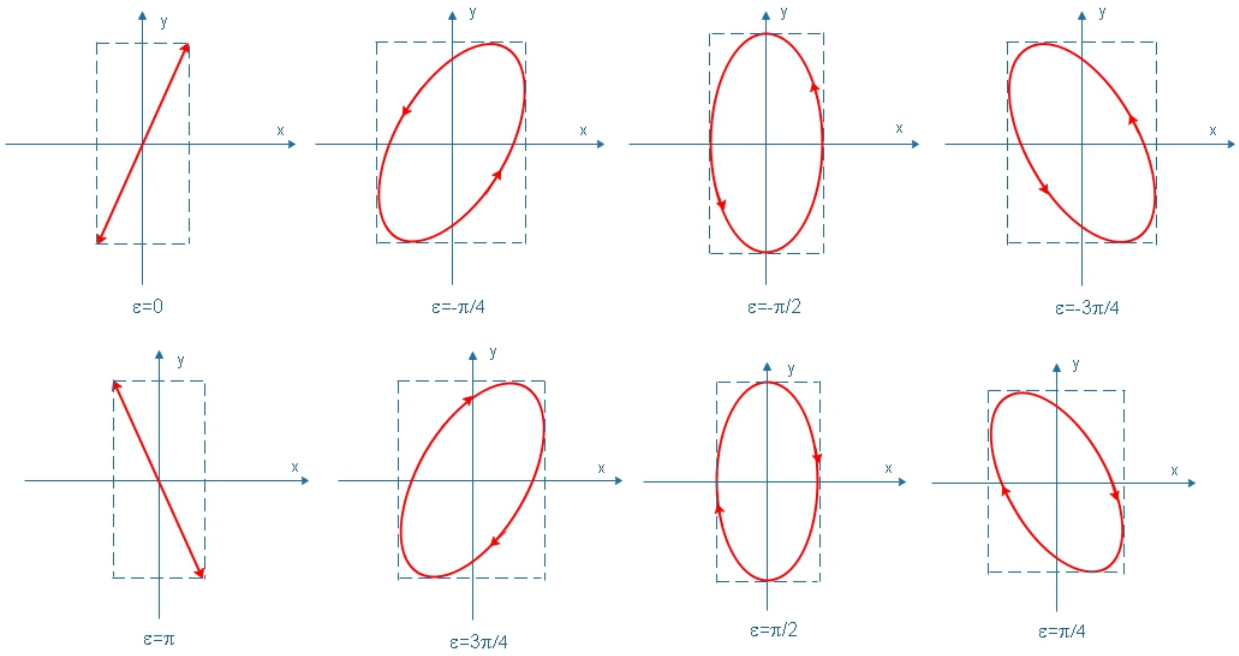


Figure 4.7: Evolution of the elliptical polarization as a function of ε for $E_{0y} = 2E_{0x}$

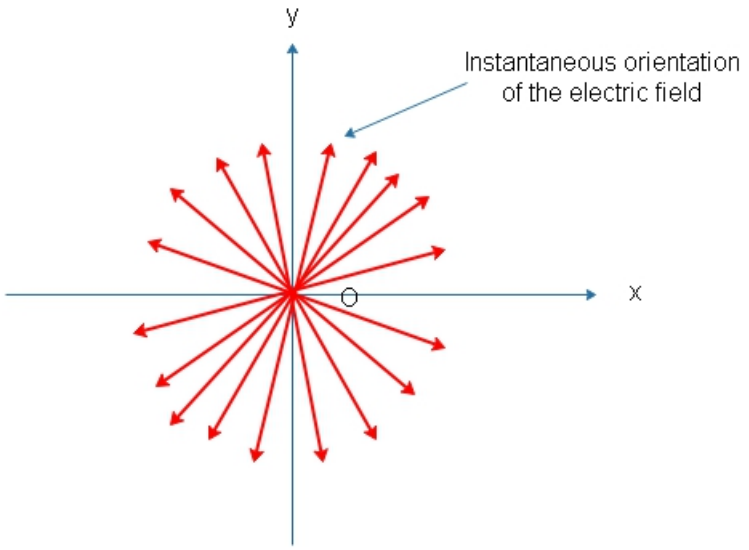


Figure 4.8: Unpolarized light: during the integration time of the detector the orientation of the $\vec{E}$ change continuously and randomly

And of course, the result of this statistical distribution can embrace any of the forms of polarization described above.

Apparently, the unpolarized light is of no interest for someone studying the polarimetry. The reality is rather contrary: it is by the presumption of a nonpolarized emission of certain stars that the polarization effect of interstellar medium is detected for example, or the polarization in the corona is understood[33].

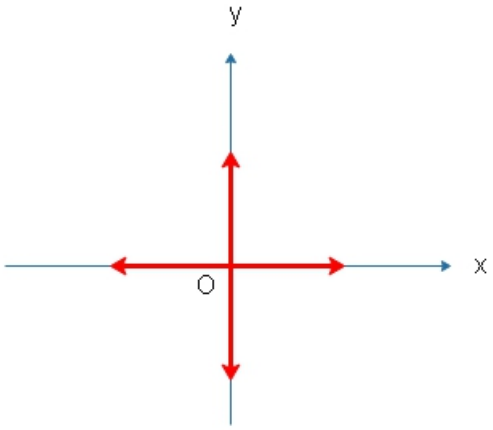


Figure 4.9: General representation of the unpolarized light: the projections along x and y, in the plane perpendicular to the direction of movement, are considered simultaneously.

4.3 Stokes formalism

The Stokes formalism is one of the most elegant way to deal with the polarization phenomenon, that consider all the transformations that light can endure into a medium, in order to give a representation of the emergent wave in terms of observables, and more precisely, of intensity.

This formalism is based on the Stokes parameters, a finding that allows us to describe all states of polarization, even the unpolarized light.

Considering that

$$\vec{E}_x(t) = \mathbf{e}_x \cdot E_{0x} \cos\left((kz - \omega t) + \varepsilon_x(t)\right) \tag{4.3.1}$$

$$\vec{E}_y(t) = \mathbf{e}_y \cdot E_{0y} \cos\left((kz - \omega t) + \varepsilon_y(t)\right) \tag{4.3.2}$$

then we may write the four Stokes parameters as

$$\begin{array}{lccc} I = & \langle I \rangle & = \langle E_{0x}^2 \rangle + \langle E_{0y}^2 \rangle \\ Q = & \langle I_0 \rangle - \langle I_{90} \rangle & = \langle E_{0x}^2 \rangle - \langle E_{0y}^2 \rangle \\ U = & \langle I_{45} \rangle - \langle I_{-45} \rangle & = \langle 2E_{0x}E_{0y} \cos\varepsilon \rangle \\ V = & \langle I_{RCP} \rangle - \langle I_{LCP} \rangle & = \langle 2E_{0x}E_{0y} \sin\varepsilon \rangle \end{array} \tag{4.3.3}$$

in which $\varepsilon = \varepsilon_y - \varepsilon_x$, and the brackets $\langle\rangle$ signify the average value over time. The calculus uses the averaged value in order to reflect the limited capacity of detectors and to encompass the unpolarized light. This last one corresponds to the case (Q,U,V)=(0,0,0). For the totally polarized light we have

$$Q^2 + U^2 + V^2 = I^2 \tag{4.3.4}$$

while to a partially polarized light:

$$Q^2 + U^2 + V^2 < I^2 \tag{4.3.5}$$

Kliger [22](p.77) explains this relation by the fact that the partially polarized light can be imagined as the superposition of two beams: of one which is totally polarized and another that is unpolarized. Therefore we can write:

$$\begin{cases} \text{total pol.:} & Q^2+U^2+V^2=I_1^2 \\ \text{non pol.:} & I_2 \end{cases}$$

Obviously we will have then:

$$Q^2+U^2+V^2<(I_1+I_2)^2 \tag{4.3.6}$$

because the received intensity is always positive.

From the definition of the Stokes parameters and from the previous description of the polarization ellipse, a formulation of this parameters in terms of ellipticity (χ) and azimuth (ψ) can be inferred, the average being implied. Thus, for an ellipticity governed by

$$\sin(2\chi)=\frac{2E_{0x}E_{0y}}{E_{0x}^2+E_{0y}^2}\sin\varepsilon$$

and an azimuth angle

$$\tan(2\psi)=\frac{2E_{0x}E_{0y}}{E_{0x}^2-E_{0y}^2}\cos\varepsilon$$

it can be shown that the Stokes parameters can be expressed as:

$$\begin{aligned} I &= E_{0x}^2+E_{0y}^2 \\ Q &= (E_{0x}^2+E_{0y}^2)\cos(2\chi)\cos(2\psi) \\ U &= (E_{0x}^2+E_{0y}^2)\cos(2\chi)\sin(2\psi) \\ V &= (E_{0x}^2+E_{0y}^2)\sin(2\chi) \end{aligned} \tag{4.3.7}$$

or, by normalizing with the intensity $I=E_{0x}^2+E_{0y}^2$:

$$\begin{aligned} i &= 1 \\ q &= \cos 2\chi\cdot\cos 2\psi \\ u &= \cos 2\chi\cdot\sin 2\psi \\ v &= \sin 2\chi \end{aligned} \tag{4.3.8}$$

where i,q,u,v are the normalized values of Stokes parameters (i=I/I, q=Q/I, etc.).

Those four parameters define the so called Stokes vector

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \text{ also notated } \left|\begin{matrix} I \\ Q \\ U \\ V \end{matrix}\right| \text{ or } \{I \quad Q \quad U \quad V\}$$

The degree of polarization is defined as:

$$p_t=\frac{\sqrt{Q^2+U^2+V^2}}{I}\in[0,1] \tag{4.3.9}$$

If the light is nonpolarized, then $Q=U=V=0$, and $p_t=0$. Also, if $I^2=Q^2+U^2+V^2$, then $p_t=1$ and we say that the light is totally polarized. For intermediary values, the light is considered as partially polarized.

Of great interest is also the degree of linear polarization, defined as:

$$p_l = \frac{\sqrt{Q^2 + U^2}}{I} \tag{4.3.10}$$

and the degree of circular polarization:

$$p_c = \frac{V}{I} \tag{4.3.11}$$

An alternative formulation for the degree of polarization which considers the radial and the tangential polarization as it was described in the first chapter is

$$P = \frac{I_\perp - I_\parallel}{I_\perp + I_\parallel}$$

for the degree of linear polarization, in which $I_\perp$ is the tangential intensity, whereas $I_\parallel$ stays for the intensity in the radial orientation.

4.3.1 Few particular examples

In this notation, I stays for the intensity of the beam, Q and U for the linear polarization and V for the circular polarization.

Thus, for example we have:

Table 4.1: Stokes vectors for different types of polarization

Stokes vector	Unpolarized	Parallel (Ox)	Perpendicular (Oy)
$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$
$+45°$	$-45°$	Right circular (RCP)	Left circular (LCP)
$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

An intuitive representation of this polarization states can also be retrieved in the figure below:

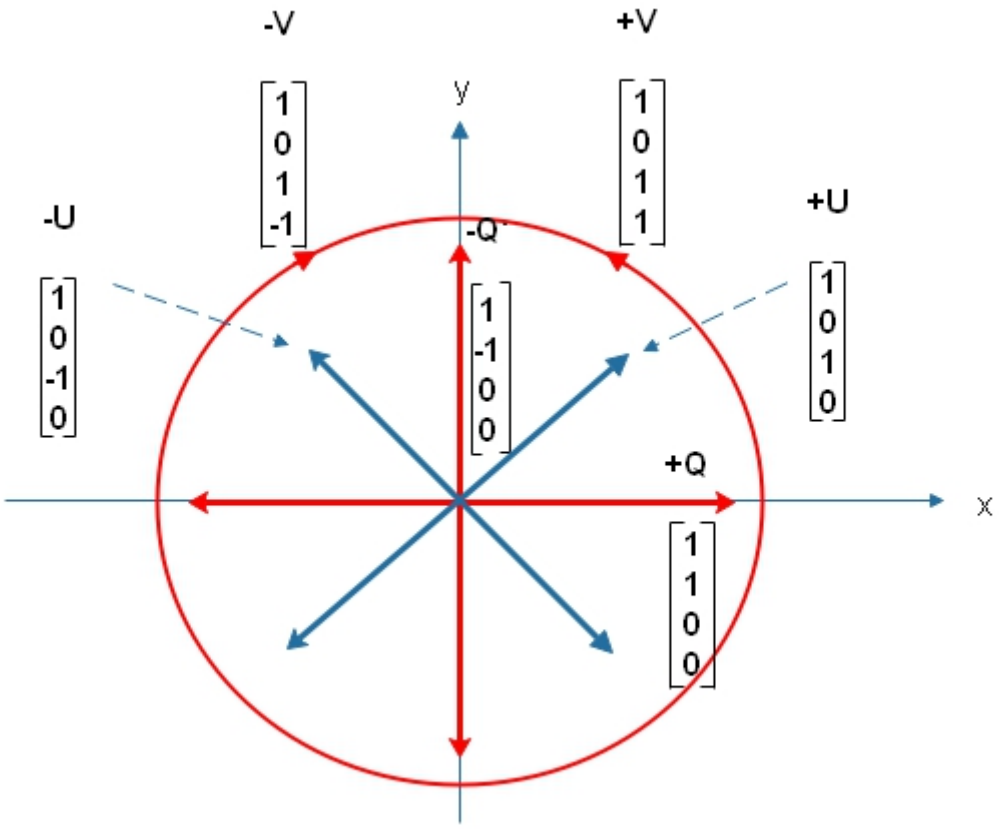


Figure 4.10: General representation of the Stokes vectors and corresponding polarization states

4.4 Mueller formalism

For a given polarized beam which passes through different media, the best way to determine the polarization state of the outgoing beam is to use the Muller matrices (M_i):

$$\vec{S_{out}} = M \cdot \vec{S_{in}} \tag{4.4.1}$$

where

$$M = M_n \cdot M_{n-1} \cdot .... \cdot M_1$$

is the resultant matrix of the system through which the light is passing. Each matrix M_i describes the passage of light through the element i of the system.

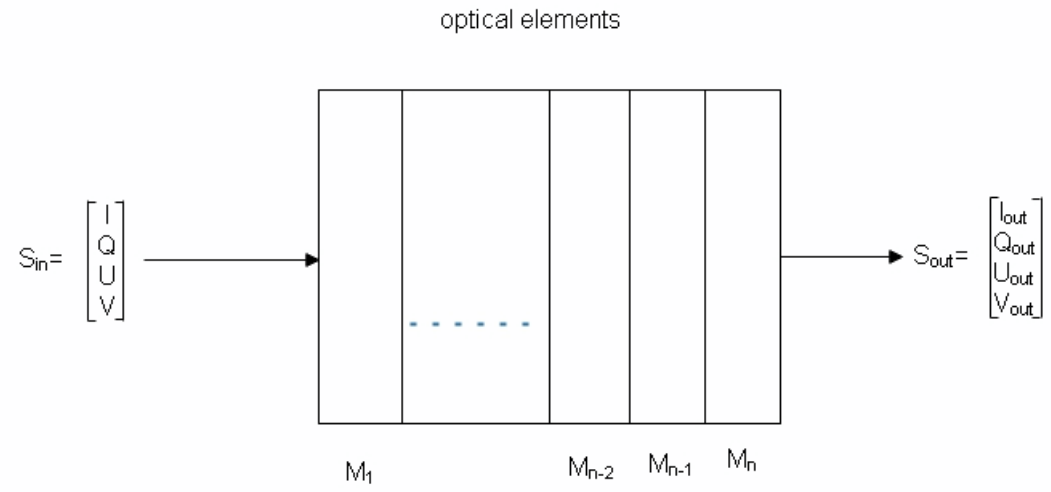


Figure 4.11: Each element of the optics, $M_1, M_2, ..., M_n$ will affect the intensity of the incoming beam, the corresponding Stokes vector undergoing multiple transformations during the passage

The underlying assumption of this representation is the fact the original Stokes vector undergoes a linear transformation through the medium:

$$S_{in} \quad \longrightarrow \quad S_{out}$$

or, mathematically:

$$\begin{aligned} I_{out} = &\quad m_{11}I + m_{12}Q + m_{13}U + m_{14}V \\ Q_{out} = &\quad m_{21}I + m_{22}Q + m_{23}U + m_{24}V \\ U_{out} = &\quad m_{31}I + m_{32}Q + m_{33}U + m_{34}V \\ V_{out} = &\quad m_{41}I + m_{42}Q + m_{43}U + m_{44}V \end{aligned}$$

the Mueller matrix of the system being then expressed as:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix}$$

Even though the Mueller calculus is a very elegant way to deal with any type of polarization able to give access to the outgoing Stokes vector, the calculation of the elements of the Mueller matrix can be sometimes a very hard task. Generally, the specialized literature provides all the time the most common matrices for the largely used optical components. Nevertheless, for more rare optical structures, like the uniaxial anisotropic medium that we will use for our instrument, the computation of these matrices must be done. Further we will try then to determine the corresponding Mueller matrix of such a component.

There exist several methods for the determination of Mueller matrices. For the sake of simplicity and clarity, we chose here a method based on the Jones calculus. This one is mostly used for describing phenomena in which the phase or the amplitude of the light plays a significant role (while the Mueller calculus resumes to the behavior of intensities).

Theoretically the elements of the Mueller matrices are related to the Jones calculus via:

$$m_{ij} = \frac{1}{2} Tr[M\sigma_j M^{\dagger}\sigma_i] \tag{4.4.2}$$

where Tr is the trace operator, M is the corresponding Jones matrix of the system, $\dagger$ is the conjugate transpose operator and $\sigma_{i,j}$ are the Pauli matrices.

As a consequence, first step into the attempt to find the Mueller matrix of the birefringent medium is to calculate the corresponding Jones matrix. Further on we will briefly pass through this computation.

According to the Jones representation[2], any beam of polarized light can be described via a vector **V**:

$$\mathbf{V} = \begin{bmatrix} E_{0x}e^{-i\phi/2} \\ E_{0y}e^{+i\phi/2} \end{bmatrix}$$

where ϕ is phase difference. Thus, for example, the linear polarization can be described by the vector:

$$\mathbf{V} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

while for the circular polarization we have

$$\mathbf{L} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \mathbf{R} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Generally, for any type of polarization, according to the Jones approach we have

$$\mathbf{V} \cdot \mathbf{V}^{\dagger} = \mathbf{V}^{\dagger} \cdot \mathbf{V} = I$$

and the normalization condition,

$$V_{11} \cdot V_{11}^{*} + V_{21} \cdot V_{21}^{*} = 1 \tag{4.4.3}$$

where $\dagger$ is the conjugate transpose operator, $*$ stays for the conjugate and I is unity matrix.

Moreover, for any anisotropic medium, which interests us most of all, Jones associates a square matrix of type

[2]This brief description of Jones calculus and of its relation with the Mueller method is based mostly on Huard[17]

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

which affects the phase of the incident light. According to this theory, the matrix M has two eigenvalues λ_1, λ_2 associated to two eigenvectors $\mathbf{V_1}, \mathbf{V_2}$. This two vectors are orthogonal and can be generally expressed via [17][p.89]:

$$\mathbf{V_1} = \begin{bmatrix} u \\ v \end{bmatrix} \qquad \mathbf{V_2} = \begin{bmatrix} -v^* \\ u^* \end{bmatrix}$$

Also they verify the normalization condition:

$$uu^* + vv^* = 1$$

Thereupon we may write:

$$M\mathbf{V_1} = \lambda_1 \mathbf{V_1}$$
$$M\mathbf{V_2} = \lambda_2 \mathbf{V_2}$$

And this system can be developed in:

$$\begin{cases} Au + Bv = \lambda_1 u \\ Cu + Dv = \lambda_1 v \\ -Av^* + Bu^* = -\lambda_2 v^* \\ -Cv^* + Du^* = \lambda_2 u^* \end{cases}$$

Solving this system for (A, B, C, D) by keeping in mind that we have a normalization condition, we find:

$$M = \frac{1}{uu^* + vv^*} \cdot \begin{bmatrix} \lambda_1 uu^* + \lambda_2 vv* & (\lambda_1 - \lambda_2)uv^* \\ (\lambda_1 - \lambda_2)vu^* & \lambda_2 uu^* + \lambda_1 vv* \end{bmatrix}$$

$$\Leftrightarrow M = \begin{bmatrix} \lambda_1 uu^* + \lambda_2 vv* & (\lambda_1 - \lambda_2)uv^* \\ (\lambda_1 - \lambda_2)vu^* & \lambda_2 uu^* + \lambda_1 vv* \end{bmatrix}$$

Considering now the case of an incident wave linearly polarized passing through a retarding waveplate, the $\mathbf{V_1}, \mathbf{V_2}$ can be expressed:

$$\mathbf{V_1} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad \mathbf{V_2} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

Also, according to Jones method, the eigenvalues λ_1 and λ_2 can be written as:

$$\lambda_1 = e^{in_1 k_0 e} = e^{i\phi_1} \qquad \lambda_2 = e^{in_2 k_0 e} = e^{i\phi_2}$$

in which n_i is the corresponding refraction index, k_0 is the wave number and e is the thickness of the plate. Substituting now with this identities into the M formula we will find:

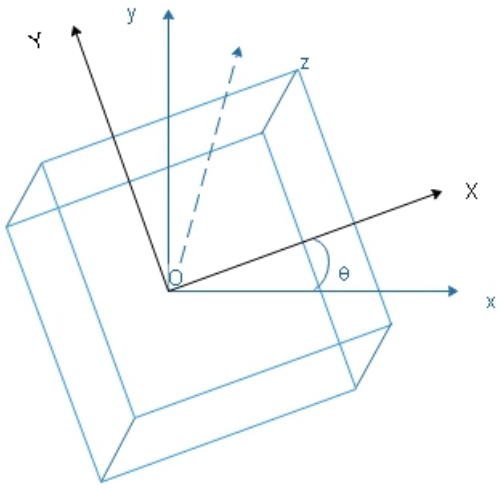


Figure 4.12: Retarder plate example: XOY is the reference frame of the plate and xOy is reference of the laboratory

$$M = \begin{bmatrix} e^{i\phi_1}\cos^2\theta + e^{i\phi_2}\sin^2\theta & (e^{i\phi_1} - e^{i\phi_2})\cos\theta\sin\theta \\ (e^{i\phi_1} - e^{i\phi_2})\cos\theta\sin\theta & e^{i\phi_2}\cos^2\theta + e^{i\phi_1}\sin^2\theta \end{bmatrix}$$

At this point of computation, a change of variable is useful:

$$\begin{cases} \phi = \phi_1 - \phi_2 \\ \psi = \frac{\phi_1+\phi_2}{2} \end{cases} \quad \Leftrightarrow \begin{cases} \phi_1 = \psi + \frac{\phi}{2} \\ \phi_2 = \psi - \frac{\phi}{2} \end{cases}$$

Thus, replacing with this new notations for the phase into the last form of the M matrix we will find

$$M(\theta,\phi) = e^{i\psi}\begin{bmatrix} \cos\frac{\phi}{2} + i\cos 2\theta\sin\frac{\phi}{2} & i\sin 2\theta\sin\frac{\phi}{2} \\ i\sin 2\theta\sin\frac{\phi}{2} & \cos\frac{\phi}{2} - i\cos 2\theta\sin\frac{\phi}{2} \end{bmatrix}$$

or, by ignoring the phase term that will play no role in the determination of the Mueller matrix:

$$M(\theta,\phi) = \begin{bmatrix} \cos\frac{\phi}{2} + i\cos 2\theta\sin\frac{\phi}{2} & i\sin 2\theta\sin\frac{\phi}{2} \\ i\sin 2\theta\sin\frac{\phi}{2} & \cos\frac{\phi}{2} - i\cos 2\theta\sin\frac{\phi}{2} \end{bmatrix} \tag{4.4.4}$$

Further, based on the fact that

$$m_{ij} = \frac{1}{2}Tr[M\sigma_j M^\dagger \sigma_i]$$

and because σ_i is the Pauli base:

$$\sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_4 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{4.4.5}$$

we can easely calculate the Mueller matrix of any anisotropic media.

For example, if we consider now the case of a quarter wave plate (QWP), for which $\phi = \frac{\pi}{2}$ and for simplification we replace θ by 0 (the orientation of the optical axis with regard to x), then the corresponding Mueller matrix will be:

$$M(\tfrac{\pi}{2}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

For a half-wave plate (HWP), for which $\phi = \pi$, the matrix is[3]:

$$M(\pi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Generally, for a plate able to induce a total phase difference ϕ, usually called a wave plate (WP), the Mueller matrix will be

$$M(\phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{bmatrix}$$

Continuing our analysis, if we consider now the case of a simple linear polarizer, oriented to an angle θ with respect to x axis, then, because only one direction will be allowed to pass through the system, the eigenvalues for $\mathbf{V_1}$ and $\mathbf{V_2}$ will be 1 and 0, and so the Jones matrix will be:

$$M_J(\theta) = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \tag{4.4.6}$$

And then, using the same procedure as before we may compute the corresponding Muller matrix for any type of linear polarizer:

$$M(\theta) = \begin{bmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \cos 2\theta \sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta \sin 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.4.7}$$

In the case of a rotator, the vectors $\mathbf{V_1}$ and $\mathbf{V_2}$ are given by:

$$\mathbf{V_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \mathbf{V_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

The corresponding Jones matrix of this type of element is:

[3] More examples can be find in the end of this thesis, in the Appendix F

$$M(\phi) = e^{i\psi} \begin{bmatrix} \cos\frac{\phi}{2} & -\sin\frac{\phi}{2} \\ \sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{bmatrix} \tag{4.4.8}$$

Replacing $\phi/2$ by θ, which is the rotation angle, we may compute the Muller matrix of a rotator:

$$M(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Now, if we came back to the problem of a birefringent medium which simultaneously induces a phase difference and rotates the plane of vibration according to the orientation of the optical axis, then, the associated Mueller matrix is expressed as:

$$M_{WP}(\theta, \phi) = M(-\theta) \cdot M_{WP}(0, \phi) \cdot M(\theta) \tag{4.4.9}$$

where $M(\theta)$ is the matrix of a rotator, previously determined. Developing this product we may finally find:

$$M(\theta, \phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \cos\phi \sin^2 2\theta & (\cos\phi - 1)\sin 2\theta \cos 2\theta & \sin\phi \sin 2\theta \\ 0 & (\cos\phi - 1)\sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos\phi \cos^2 2\theta & \sin\phi \cos 2\theta \\ 0 & -\sin\phi \sin 2\theta & -\sin\phi \cos 2\theta & \cos\phi \end{bmatrix} \tag{4.4.10}$$

This final result will be largely used in the next chapters, when optical systems based on rotated waveplates will be described. At this point it is worthy to mention that the literature in the polarization field could be confusing when it comes to present the general matrices for a wave plate, a rotator or a general rotator. Depending on how we consider the position of the system with regard to the reference frame of the laboratory, the rotation matrix can be expressed like

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta & 0 \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

As a consequence, the above matrix of a rotated wave plate, will become using the second matrix of rotation:

$$M(\theta, \phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \cos\phi \sin^2 2\theta & (1 - \cos\phi)\sin 2\theta \cos 2\theta & \sin\phi \sin 2\theta \\ 0 & (1 - \cos\phi)\sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos\phi \cos^2 2\theta & -\sin\phi \cos 2\theta \\ 0 & -\sin\phi \sin 2\theta & \sin\phi \cos 2\theta & \cos\phi \end{bmatrix}$$

and, obviously, the matrix of a rotator(2.39):

$$M(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha & 0 \\ 0 & -\sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

the main reason of all this differences being the projection of electrical field on the chosen reference frame in the Jones method. In order to be consistent with all the results presented above, further on we will use the identities established by the base of the decomposition illustrated in the Figure 4.12 and which correspond to the formalism used by Huard [17].

Chapter 5

Anisotropic media and Huygens constructions

5.1 Birefringence

The birefringence property exhibited by some crystals was discovered by the Danish scientist Rasmus Bartholin (or Bartholinus, in the Latinized form), in 1670. At that moment, the scholar observed that incoming natural light emerged from a piece of calcite crystal in two different directions simultaneously as though by a subtle mechanism the ray was split inside the medium.

Even though it was a relatively easy to reproduce phenomenon, it did not receive a complete explanation but a few hundred years later, after the formulation of the wave theory of light, by Fresnel and Arago.

Still, the propagation of light insight a birefringent medium remains a difficult problem requiring for certain cases a vast mathematical apparatus in order to be described. And this mostly because in such a medium the Poynting vector is not anymore oriented along $\vec{k}$, the electric field also is not perpendicular to $\vec{k}$ and all because the refraction index is not the same in all directions. It is the symmetry of the crystal, the arrangement of atoms inside that dictates all this behaviour: a cubic structure will ensure for example a homogeneous propagation of light, reason for which they are considered as isotropic, while trigonal, tetragonal or hexagonal structures will form something which is called uniaxial crystals. On the same way we may encounter biaxial crystals, such as mica. In a intuitive representation, an isotropic medium has the same index of refraction in all directions, while a uniaxial crystal for example will have a certain index along the so called optical axis and

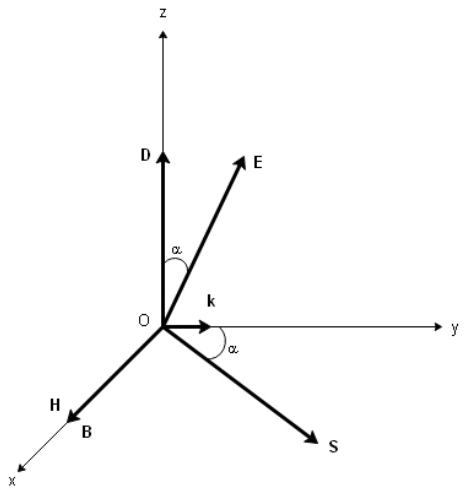


Figure 5.1: Representation of the vectors describing the propagation of a wave in an anisotropic medium: $\vec{k}$ is the direction of propagation of the incoming wave, $\vec{E}$ is the electric field, $\vec{B}$ the magnetic field, $\vec{D}$ the induced electric field, $\vec{H}$ the induced magnetic field and $\vec{S}$ is the Poynting vector.

another in the plane perpendicular to this direction.

The mathematical description of the propagation of light in anisotropic media is quite long and elaborate, representing by itself an independent field of optics [37]. Because of this, we will be forced to restrain our presentation only to the results which are useful to our analysis.

The most important idea, that governs the description of the propagation of light in an anisotropic medium (and not only), is the fact that the electric field of the incident wave $\vec{E}$ will generate an electromagnetic induction $\vec{D}$ that will be adjusted by the permittivity of the medium $[\epsilon]$. The existence of this electric induction will make that the POYNTING vector describing the direction of propagation of the energy flux to be no longer parallel to $\vec{k}$, who characterized the direction of propagation of the wave.

By considering that α, β and γ are the components of the unity vector $\vec{u}$, ($\vec{k} = k \cdot \vec{u}$) in the system $Oxyz$ in which the tensor $[\epsilon]$ is diagonal, while n_x,n_y and n_z are the main refractive indices, and n is the refractive index along the propagation direction of the wave, then the relations:

$$\vec{D} = \frac{k^2}{\omega^2 \mu_0}[\vec{E} - (\vec{u} \cdot \vec{E})\vec{u}]$$

and

$$\vec{D} = \epsilon_0 \cdot [\epsilon] \cdot \vec{E}$$

can be translated also into the condition:

$$\frac{n_x^2 \alpha^2}{n^2 - n_x^2} + \frac{n_y^2 \beta^2}{n^2 - n_y^2} + \frac{n_z^2 \gamma^2}{n^2 - n_z^2} = 0$$

which is the Fresnel equation. It is by the mean of this equation that we may have access to the refraction index n.

Solving this equation we find four values for n: $\pm n'$ and $\pm n$", the sign being dictated by the orientation of $\vec{u}$. If we consider only the positive values we may associate to them a phase speed:

$$\begin{cases} v' = \frac{c}{n'} \\ v" = \frac{c}{n"} \end{cases}$$

It is this splitting of the incident wave into two waves travelling at different speeds inside the medium that explains the phenomenon of birefringence.

Commonly, the mathematical development of the Fresnel equation make use also of something which is called the indicatrix surface. This is defined as the locus of the points of coordinates

$$X = n\alpha \quad Y = n\beta \quad Z = n\gamma$$

Using this notations, the Fresnel equation may also be expressed as:

$$[n_x^2 X^2 + n_y^2 Y^2 + n_z^2 Z^2][X^2 + Y^2 + Z^2]$$
$$-\Big[n_x^2 X^2(n_y^2 + n_z^2) + n_y^2 Y^2(n_x^2 + n_z^2) + n_z^2 Z^2(n_x^2 + n_y^2)\Big] + n_x^2 n_y^2 n_z^2 = 0$$

In the particular case of an uniaxial crystal, as is the case MgF_2 that is proposed for use here, the main indices n_x, n_y are equal: $n_x = n_y = n_o$ (the ordinary index). Along z axis we have $n_z = n_e$ (the extraordinary index).

Under this assumption, the previous equation can be decomposed in the product of a spherical surface and an ellipsoid of revolution:

$$\begin{cases} X^2 + Y^2 + Z^2 = n_o^2 \\ \frac{X^2}{n_e^2} + \frac{Y^2}{n_e^2} + \frac{Z^2}{n_o^2} = 1 \end{cases} \tag{5.1.1}$$

Considering a section in the plane $Y = 0$ we retrieve the circle of indices, corresponding to the case $X^2 + Z^2 = n_o^2$ and the ellipse $\frac{X^2}{n_e^2} + \frac{Z^2}{n_o^2} = 1$, as it can be seen in the Figure 5.2.

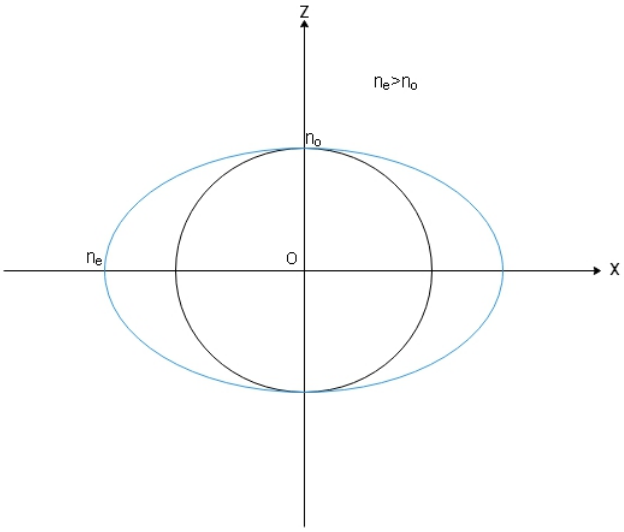


Figure 5.2: Circle and ellipse of indices in the plane Y=0 for the case $n_e > n_o$

Using also the relations for X, Y, Z introduced before we have:

$$\begin{cases} n^2(\alpha^2 + \beta^2 + \gamma^2) = n_o^2 \\ n^2\left(\frac{\alpha^2}{n_e^2} + \frac{\beta^2}{n_e^2} + \frac{\gamma^2}{n_o^2}\right) = 1 \end{cases}$$

Solving this system for $(\alpha^2 + \beta^2 + \gamma^2 = 1)$ we find that

$$\begin{cases} n = n_o^2 \\ n = \frac{n_e n_o}{\sqrt{n_o^2(\alpha^2+\beta^2)+n_e^2\gamma^2}} \end{cases} \tag{5.1.2}$$

Considering again the case Y=0, and an angle θ between $\vec{u}$ and Z axis (which is also the angle between the incident wave and the optical axis of the medium), we can write:

$$\begin{cases} n = n_o^2 \\ n^2(\theta)\left(\frac{\sin^2(\theta)}{n_e^2} + \frac{\cos^2(\theta)}{n_o^2}\right) = 1 \end{cases} \tag{5.1.3}$$

This last system tells us that for any incident beam perpendicular to the optical axis, with a general state of polarization, the component of $\vec{E}$ vibrating into a plane perpendicular to the optical axis will see a medium of index $n(\theta = \pi/2) = n_o$ in the direction of propagation. This vibration will propagate inside the crystal as an ordinary wave. On the other hand, the component of $\vec{E}$ vibrating parallel to the optical axis will see an index n_e. This is called the extraordinary ray. Consequently, a

phase difference will appear between the two directions of vibration. Generally, for an incidence at angle θ between the optical axis and the direction of the incoming wave, the extraordinary ray will see an index of refraction $n(\theta)$ given by:

$$\frac{1}{n^2(\theta)} = \frac{\sin^2(\theta)}{n_e^2} + \frac{\cos^2(\theta)}{n_o^2} \tag{5.1.4}$$

This relation is valid only if the optical axis is contained in the plane of incidence ($\beta = 0$). Otherwise we have to use the relation (5.1.2) to derive the value of n.

Based on this short introduction into the anisotropic problematic we may properly define now the birefringence of medium:

$$\Delta n = |n_o - n_e| \tag{5.1.5}$$

If $n_o > n_e$, then the crystal is negative, while if $n_o < n_e$ is positive.

5.1.1 Propagation of light: Huygens construction

Another aspect related to the birefringence and that will be largely used in the last part of this research is the trajectory of the rays of light in a uniaxial crystal. We will present here only the case of a positive crystal, that corresponds to MgF_2. In a first scenario we will suppose that the light is passing from an isotropic environment of refractive index n, into the crystal whose optical axis is oriented perpendicular to z axis (Figure 5.3).

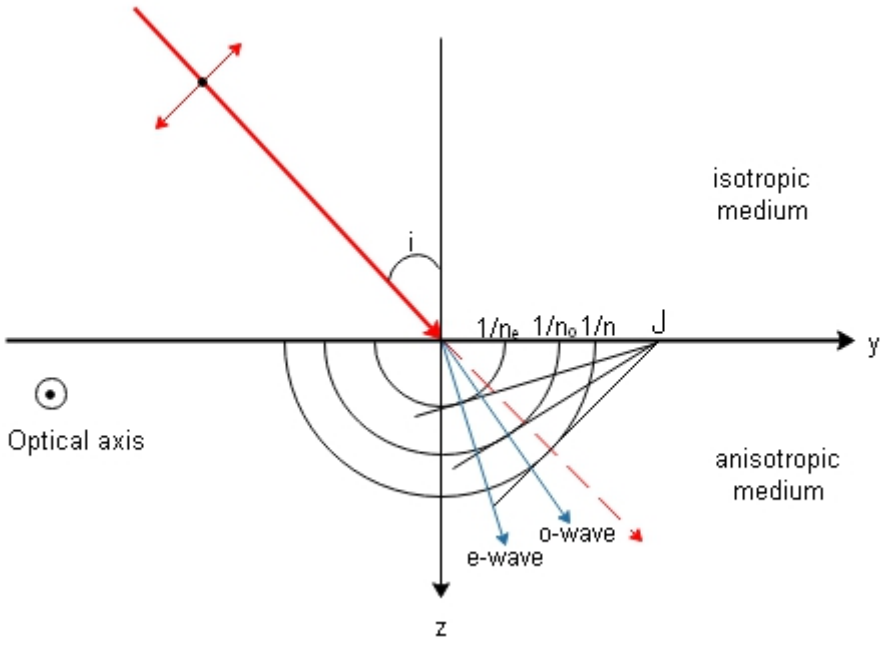


Figure 5.3: Huygens construction for the first scenario: optical axis perpendicular to z

If we consider that the incident wave is not polarized, then it can be represented by two orthogonal vibrations: one parallel to the optical axis and another perpendicular. Using the relations (5.1.3) we infer that the ordinary wave, vibrating perpendicular to the optical axis, will see in the crystal a

refractive index n_o in any direction, while the extraordinary will encounter a medium of index n_e in any direction, because the angle θ is 90° here.

Because both waves are traveling into the crystal like in an isotropic medium, the angles made with the z axis can be found using the Snell-Descartes low:

$$\begin{cases} n \sin i = n_o \sin r_o \\ n \sin i = n_e \sin r_e \end{cases}$$

Concerning the geometry of the problem, the simplest way to represent it is to use the Huygens construction. In the Figure 5.3 it is represented the construction based on the radial speeds. On the same way, a construction using only the surfaces of indices can be built.

Thus, in this approach, the first step is to represent a circle of radius $1/n$ centred in the incidence point and situated in the plane of incidence. Then we draw an extension of the incident wave (dashed red). From the point of intersection of this extension with the circle of radius $1/n$ we draw a tangent to the circle that will intersect the diopter surface in the point J. On the same way we will construct the circles of radius $1/n_o$ and $1/n_e$. Then, from J we draw the tangents to the new circles: the tangent points will dictate the direction of the ordinary and extraordinary waves.

Let's suppose now that the optical axis is oriented along z axis, as in the Figure 5.4.

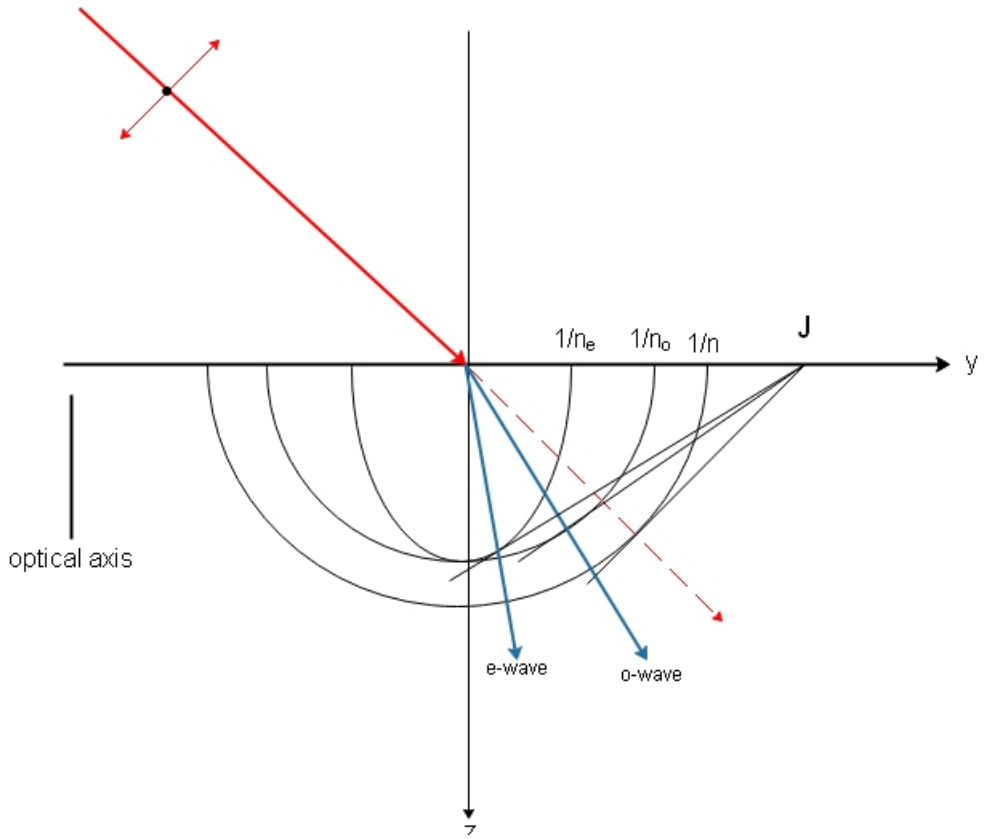


Figure 5.4: Huygens construction for the second scenario: optical axis along z

In this case, the ordinary wave, corresponding to the vibration perpendicular to the plane of incidence, will see an isotropic medium in all directions inside the medium. Because of this, just as before, we can find its orientation by the mean of the Snell-Descartes low. On the other hand,

the vibration that is parallel to the plane of incidence will be no longer parallel to the optical axis everywhere inside the medium, like before. Because of this, the extraordinary wave will see an anisotropic medium of index $n(\theta)$, corresponding to the relation (5.1.3). The Snell-Descartes low will help us to find the angle re :

$$n \sin i = n(\theta) \sin r_e \tag{5.1.6}$$

In a third scenario we will suppose that the optical axis is making an angle α with the z axis, as it can be seen in the Figure 5.5.

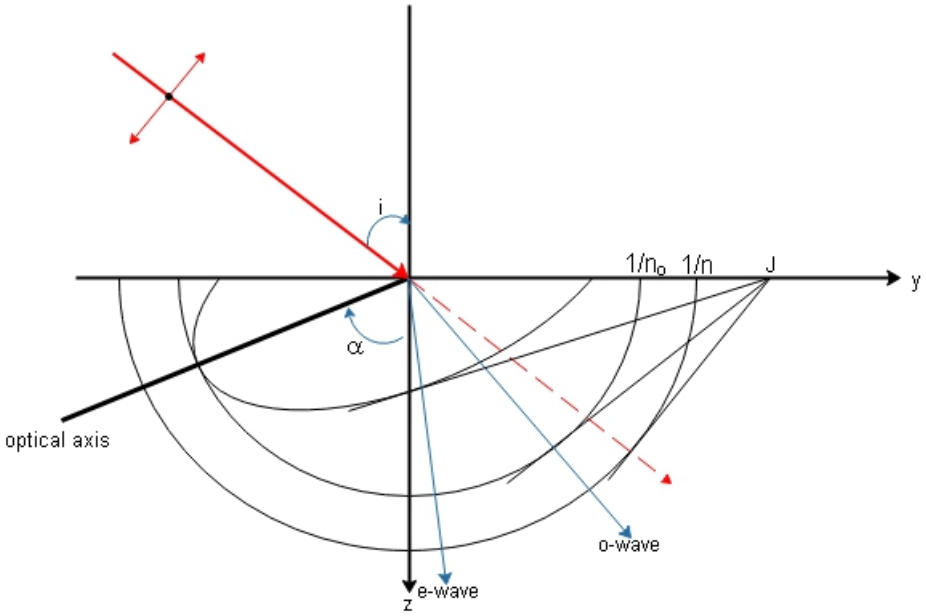


Figure 5.5: Huygens construction for the third scenario

Again, the vibration which is orthogonal to the plane of incidence will be all the time orthogonal to the optical axis, thus representing the ordinary wave. This ray will follow the Snell-Descartes low and will be refracted with an index n_o. On the other hand, the vibration parallel to the plane of incidence will form the extraordinary ray and will see a variable index inside the medium. If we consider r_e the refraction angle of the extraordinary wave, then in this particular situation we have:

$$\theta = \pi - (\alpha + i)$$

and

$$n \sin i = n(\theta) \sin r_e$$

Making use again of the relation (4.1.2) we may find r_e.

An important observation is the fact that if the isotropic medium is characterized by a refraction index $n > \max(n_e, n_o)$ then, for certain values of the incidence angle i, the point J can be situated inside the surface of radius $1/n_o$ or of the ellipse. In this situation the total internal reflection will occur, and the corresponding wave will not pass into the anisotropic medium.

A situation of this type is depicted in the Figure 5.6 where we can see that for the given angle i the ordinary wave does not exist into the anisotropic medium. Only the extraordinary wave can

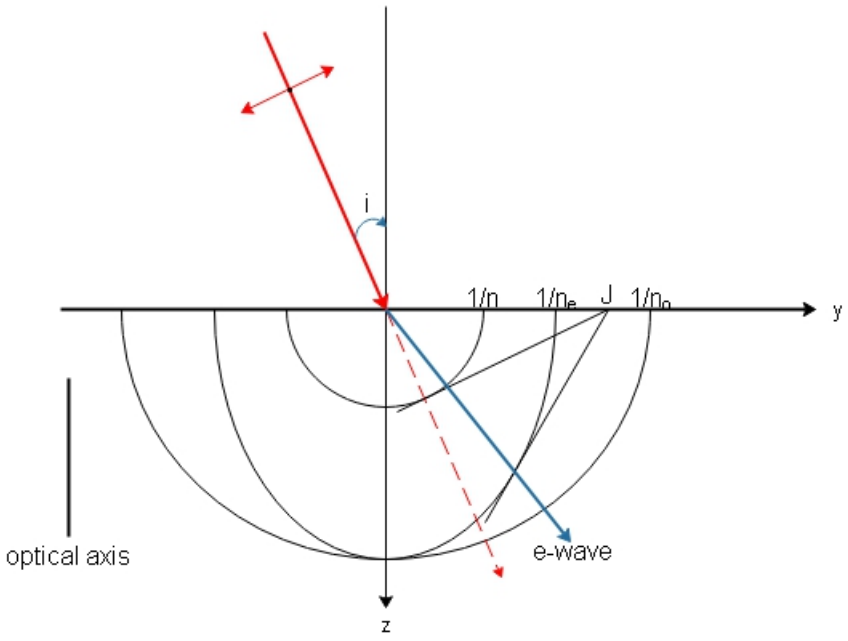


Figure 5.6: Huygens construction for a total internal reflection

propagate in this case. Consequently, the position of the point J is of paramount importance for this kind of situations. This position is given by:

$$y_J = \frac{1}{n \sin i} \tag{5.1.7}$$

If $y_J > \max(1/n_o, 1/n_e)$ then both rays will propagate through the anisotropic medium. Instead, if $y_J > 1/n_e$ and $y_J < 1/n_o$ then only the extraordinary wave will exist, while if $y_J < \min(1/n_o, 1/n_e)$, then both waves will vanish.

Because the device that will be introduced here suppose also the passage from an anisotropic medium to another, this situation must also be considered here from the point of view of ray tracing.

Thus, let us consider the case depicted in the Figure 5.7 of a ray passing from a medium of index n into an anisotropic environment with the optical axis along x (orthogonal to the plane of the page). Then the ray is passing into a second anisotropic medium with optical axis along z. The surface of separation between the two media is tilted to an angle ξ about y axis as it can be seen in the figure.

If the incoming ray is not polarized, then in the first medium we will have a separation between ordinary and extraordinary rays. However, both will see this medium as an isotropic one and because of this will follow the same path, at constant refraction indices. Because of the difference existing between n_e and n_o, a certain phase difference between the two orthogonal vibrations will be acquired during the passage through this medium. In order to understand what is happening at the entrance in the second anisotropic environment, let's follow at first the ordinary ray (o-ray) from the first medium. Regarding the optical axis of the second medium, this vibration is still orthogonal. Thus, the ray will pass in the second medium without deviation and without change in speed. In fact, the ordinary wave remains ordinary and sees no difference between one medium and another. With the extraordinary ray from the first bloc the situation is slightly different. This vibration, which is parallel to the optical axis in the first medium becomes now perpendicular to the optical axis. Thus, the e-ray is converted into an o-ray. This conversion is accompanied by a small change in orientation

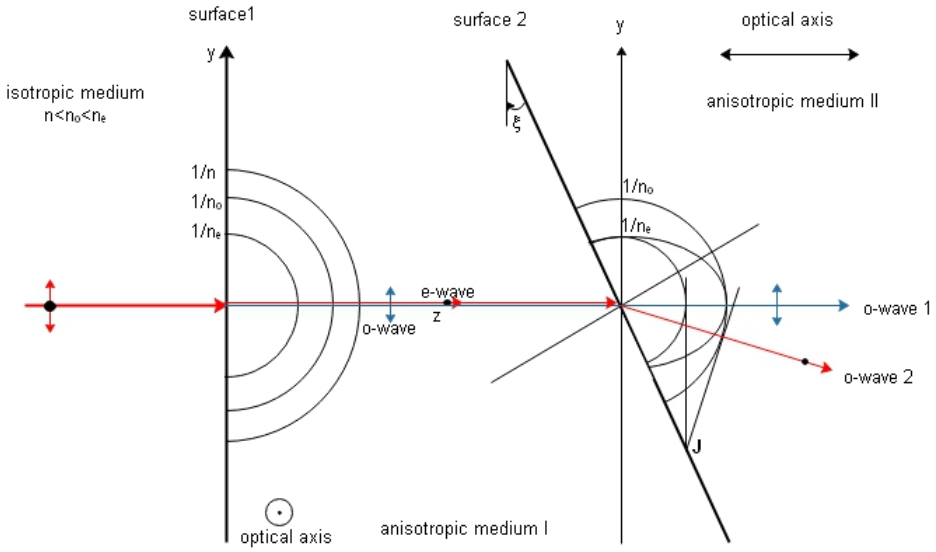


Figure 5.7: Huygens construction for two anisotropic media

which can be calculated via Snell-Descartes low:

$$\sin r_o = \frac{n_e}{n_o} \sin i = \frac{n_e}{n_o} \sin \xi$$

Moreover, we have to pay attention to the angle ξ. The relation (5.1.7) tells us that y_J should be bigger then $1/n_o$, otherwise the extraordinary ray from the first medium will be totally reflected.

Chapter 6

Static polarimeter design

The interest of using birefringent media in the construction of polarimeters is coming mostly from the fact this kind of material can separate different types of polarizations by sending them into different directions. In this way at least two measurements, of two of the Stokes parameters are accessible simultaneously.

Meanwhile, for certain configurations, rotating parts are no more needed. Because of this, the entire design can be drastically simplified. Also, for the space usage, it can prove to be much more stable, easier to implement and more ergonomic.

Currently, the technique uses mostly a combination of birefringent elements, aimed for the selection of different types of polarization, and turning analyzers or filters. An absolute stable design, without moving parts was never implemented in the space field. The main reason for this is the fact this technique is still in the research phase.

In this chapter we will see at first from where is coming the need of rotating components in a classical approach. For this we will consider the very easy example of the CLASP project, based on a rotating waveplate and an analyzer.

Then we will introduce the key model of our study: the static spectropolarimeter based on three birefringent blocs, glued together in a single piece. Therefore, we will see how by the simple modulation of the phase along the instrument we can keep the track of all four Stokes parameters.

6.1 The basic model: Why do we need rotation?

In the classical approach of a polarimeter based on birefringent media, the outgoing intensity can be expressed as a linear combination of the Stokes parameters of the incoming beam.

Simply stated, if the incident ray has a state of polarization that can be expressed through the Stokes vector S:

$$S = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$$

then the outgoing intensity which can be measured with help of a detector, can be expressed as:

$$I_{out} = \alpha I + \beta Q + \gamma U + \tau V$$

The coefficients $\alpha, \beta, \gamma, \tau$ depend on the properties of the system (Δn, thickness) and on the wavelength of the incoming ray.

As long as $\alpha, \beta, \gamma, \tau$ are constant in any point of the exist face of the polarimeter, there is no chance to find the value of I, Q, U and V from a single measurement of I_out. We have four unknowns; thus, we need at least four equations. And these equations can be obtained by changing the orientation of the analyzer or of the modulator.

An easy and good example for the "rotating" method is the CLASP project. The Chromospheric Lyman-Alpha Spectro-Polarimeter (CLASP) was a joint project of NASA and the National Astronomical Observatory of Japan, aimed for a sub-orbital mission of observation of the chromosphere in UV, EUV and soft X-ray field, from 2015.

Among other instruments, CLASP was also provided with a spectropolarimeter, as can be seen in the Figure 6.1. The main goal of this was to determine I, Q and U parameters of Hydrogen Lyman-Alpha at 121.60 nm. According to the technical specifications of the design, the polarization was measured with the help of rotating half-waveplate (HWP) and an analyzer.

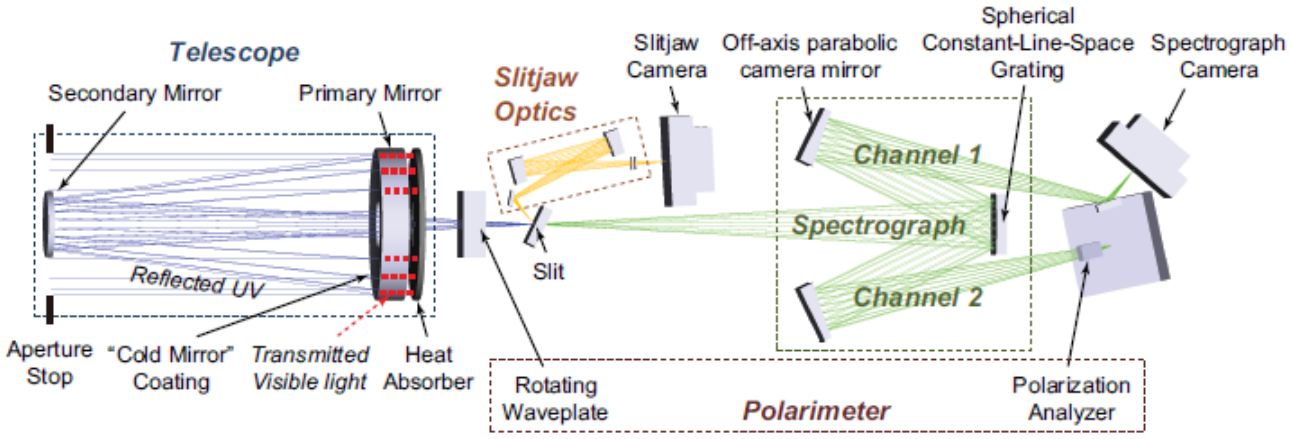


Figure 6.1: The CLASP polarimeter was based on a rotating half-wave plate (HWP) and an analyzer. Source: Clasp project[19]

As it was explained before, the reason of the rotation was to induce a variation of the $\alpha, \beta, \gamma, \tau$ coefficients. Thus, for example, if we fix the orientation of the analyzer at 0°, then as a function $\alpha(t)$, the angle of the rotating HWP, we may write:

$$\begin{cases} \alpha = 0 \Rightarrow & I_{out}(0) = \frac{1}{2}(I - Q) \\ \alpha = \frac{\pi}{4} \Rightarrow & I_{out}(\frac{\pi}{4}) = \frac{1}{2}(I + Q) \\ \alpha = \frac{\pi}{3} \Rightarrow & I_{out}(\frac{\pi}{3}) = (\frac{1}{2}I + \frac{1}{4}Q - 0.433U) \end{cases}$$

Based on this equations we may find the values of I, Q and U parameters, and then the degree of linear polarization. More than that, by increasing the number of readings the precision of the measurement can be also ameliorated.

It has to be noticed here that such a device is not suitable for the determination of the V parameter. In fact, no matter the incidence angle $\alpha(t)$, the corresponding coefficient τ will be equally zero. In fact, on the basis of the relation (4.4.10), the Mueller matrix of the rotating HWP is given by:

$$M(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\alpha) - \sin^2(2\alpha) & -2\cos(2\alpha)\sin(2\alpha) & 0 \\ 0 & -2\cos(2\alpha)\sin(2\alpha) & -\cos^2(2\alpha) + \sin^2(2\alpha) & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

That means the rotating HWP will alter the degree of linear polarization and it will only change the handedness of the circular polarization. For example, if $\alpha = 0$, then a beam linearly polarized at angle i with regard to the fast axis will be rotated with $2i$, while the elliptical and circular polarization will endure only a change in the handedness.

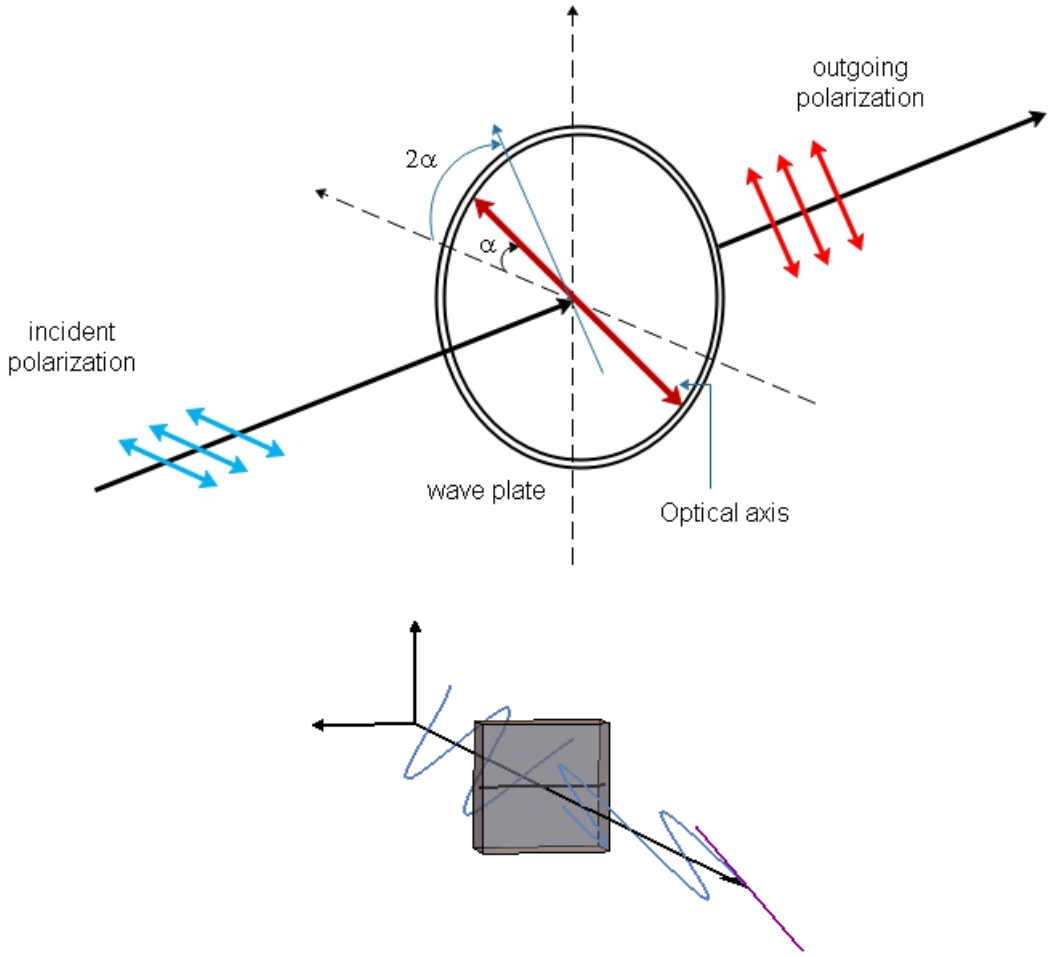


Figure 6.2: A HWP does not change the state of polarization but only flip the orientation of the vibration with regard to the fast axis. At the bottom: the same phenomenon simulated with Mathematica software

Speaking in therms of Stokes parameters it can be easily shown that the turning HWP changes the incoming Stokes vector S into:

$$S' = \begin{bmatrix} I \\ aQ + bU \\ bQ - aU \\ -V \end{bmatrix}$$

where a, b coefficients depend on α. Passing then through an analyzer, the circular component V will be totally lost.

For this configuration, the Stokes vector after passing through an analyzer oriented at angle θ with regard to x axis will embrace the general form:

$$S_{out} = \begin{bmatrix} \alpha I + \beta Q + \gamma U \\ \alpha' I + \beta' Q + \gamma' U \\ \alpha'' I + \beta'' Q + \gamma'' U \\ 0 \end{bmatrix}$$

where the coefficients before I, Q and U depends on α and θ. We see then that any trace of the V parameter has been lost. In fact, if the modulator is not able to convert the circular polarization into a linear one, the linear analyzer will not be able to give any information about this. At an even more general case, for any wavelength, where the wave plate does not act anymore as an HWP, it can also be proved that the measured intensity keeps no trace of the V parameter.

In the attempt then to build a polarimeter able to ensure the access to all four Stokes parameters we must conceive a modulator able to convert the circular polarization into linear polarization without losing the information about the rest of the parameters.

6.2 Static spectropolarimeter concept

One way to build such a device, able to give access to all four Stokes parameters of the incoming light was suggested by Sparks[40] in 2012.

The design, presented in the figure below (6.3), is not only able to ”conserve” the Stokes parameters of the incident light but also to facilitate their measurement without moving components.

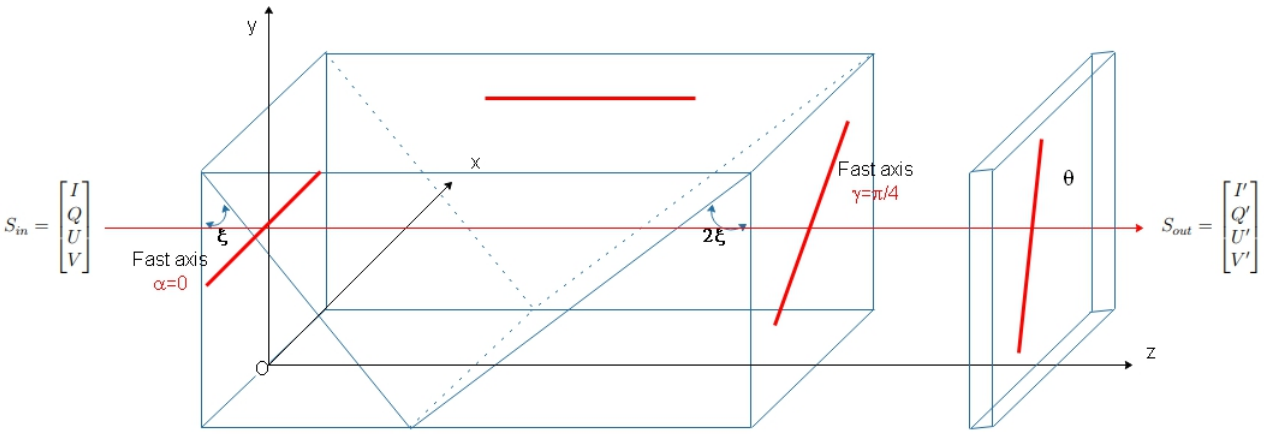


Figure 6.3: General representation of the design proposed by Sparks

The concept consider a birefringent bloc of MgF_2 divided into three parts: at the entry we have a wedge with a fast axis oriented at 0° with regard to x axis and an apex ξ; in the middle there is a prism of the same material with the fast axis along z while at the exit we have again a wedge with the fast axis oriented at 45° and apex 2ξ. The birefringent bloc is followed by an analyzer with an

angle θ about x. The impact of all the values as well as of the material will be analyzed in the next chapters. For the moment we will concentrate to the general behavior of such a device.

In the previous example of the CLASP polarimeter we saw that for a given position of the analyzer the incoming polarization was modulated about α, the angle of the wave plate. Multiple values of α were needed to determine the I, Q and U parameters.

In exchange, in the present model the rotation is not required anymore. In fact, if we consider for example an incoming monochromatic beam of wavelength λ, entering the device at 90° the phase difference induced between the horizontal and the vertical vibrations of this beam (with regard to the orientation of the optical axis) will vary on vertical (along y) as a consequence of the variation of the optical path length.

For the normal incidence, the phase difference (between the ordinary and extraordinary components) acquired in the first wedge by an incoming beam characterized by a wavelength λ and Stokes vector $S = [I, Q, U, V]$ will be:

$$\Delta\phi_1 = \frac{2\pi}{\lambda}\Delta n(\lambda)(h - y)\tan(\xi) \tag{6.2.1}$$

where $\Delta n(\lambda) = \left|n_e(\lambda) - n_o(\lambda)\right|$ is the birefringence of the material (which can vary with the wavelength), h is the height of the wedge and ξ the apex angle. In the middle part of the wedge, because of the orientation of the optical axis, in the case of a normal incidence no phase will be induced. The detailed physical reasons for this have been explained in the Chapter 5. However, for this to be true we have to neglect the impact of the angle ξ. In the last chapter of this thesis we will see how this angle affects the functioning of the device. Finally, the physicists that have proposed and discussed this type of spectropolarimeter[40][32] have also assumed that in the last wedge the phase difference is given by:

$$\Delta\phi_2 = \frac{2\pi}{\lambda}\Delta n(\lambda)(h - y)\tan(2\xi) \approx 2\Delta\phi_1 \tag{6.2.2}$$

As a consequence, the phase difference characterizing in the end a given beam will be $\Delta\phi(y, \lambda, \xi)$, a function varying along vertical, with the wavelength and with apex angle. If it is wide enough, this continuous variation of the phase difference along y ensure for example the fact that a circular incoming polarization to be converted into a linear polarization thus the V parameter to be "conserved". In fact, because of the angle ξ, we may imagine that at certain values of y the device acts as a quarter wave plate, then, at another y as a half wave plate and so on. Instead of a modulation obtained by rotation we obtain a modulation along y.

But for a proper understanding of the device, the corresponding Mueller matrix should be explored.

Based on (4.4.10), for $\theta = 0$, the Mueller matrix of the first wedge for the present configuration is:

$$M_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\Delta\phi_1) & \sin(\Delta\phi_1) \\ 0 & 0 & -\sin(\Delta\phi_1) & \cos(\Delta\phi_1) \end{bmatrix} \tag{6.2.3}$$

In the middle of the bloc the phase difference is zero and, as a consequence, the corresponding matrix is the identity matrix. For the last wedge, the fast axis is making an angle $\pi/4$ with x axis, thus in (4.4.10) we have $\theta = \frac{\pi}{4}$, and:

$$M_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\Delta\phi_2) & 0 & \sin(\Delta\phi_2) \\ 0 & 0 & 1 & 0 \\ 0 & -\sin(\Delta\phi_2) & 0 & \cos(\Delta\phi_2) \end{bmatrix} \tag{6.2.4}$$

The total Mueller matrix of this block (without analyzer) can be then easely computed:

$$M_{block} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\Delta\phi_2) & -\sin(\Delta\phi_1)\sin(\Delta\phi_2) & \cos(\Delta\phi_1)\sin(\Delta\phi_2) \\ 0 & 0 & \cos(\Delta\phi_1) & \sin(\Delta\phi_1) \\ 0 & -\sin(\Delta\phi_2) & -\cos(\Delta\phi_2)\sin(\Delta\phi_1) & \cos(\Delta\phi_1)\cos(\Delta\phi_2) \end{bmatrix} \tag{6.2.5}$$

If we look to the transformation of a Stokes vector through such a block we will find that:

$$S' = \begin{bmatrix} I \\ Q\cos(\Delta\phi_2) + V\cos(\Delta\phi_1)\sin(\Delta\phi_2) - U\sin(\Delta\phi_1)\sin(\Delta\phi_2) \\ U\cos(\Delta\phi_1) + V\sin(\Delta\phi_1) \\ V\cos(\Delta\phi_1)\cos(\Delta\phi_2) - U\cos(\Delta\phi_2)\sin(\Delta\phi_1) - Q\sin(\Delta\phi_2) \end{bmatrix}$$

This result certifies that the device is able to convert all incoming form of polarization into linear polarization because Q, U, V parameters are determining the value of Q' and U' parameters from the outgoing vector $S' = [I', Q', U', V']$.

If we consider now, in a general manner, that the analyzer is oriented to an angle θ with regard to x direction we may compute the entire Mueller matrix of the compound birefringent block + analyzer:

$$M = M_{Analyzer} \cdot M_{block} \tag{6.2.6}$$

The entire expression of the matrix M can be found in the Appendix D. By multiplying this matrix with the Stokes vector of the incoming beam we will retrieve the Stokes vector of the outgoing beam:

$$S_{out} = M \cdot S$$

The entire expression of this vector is far too long to be reproduced here but the first therm, which represents the detected intensity after the passage through the entire device is given by:

$$\begin{aligned} I_{out} = \frac{1}{2}(I + Q\cos(\Delta\phi_2)\cos(2\theta) + U(\cos(\Delta\phi_1)\sin(2\theta) - \sin(\Delta\phi_1)\sin(\Delta\phi_2)\cos(2\theta)) \\ +V(\cos(\Delta\phi_1)\cos(2\theta)\sin(\Delta\phi_2) + \sin(\Delta\phi_1)\sin(2\theta))) \end{aligned} \tag{6.2.7}$$

Thus, in the received intensity we can see the mark of the entire incoming Stokes vector. All four Stokes parameters are present here.

From this point it will be useful to introduce separate notations for the modulation terms:

$$\begin{cases} m(y,\lambda,\xi,\theta) = \cos\left(\Delta\phi_2(y,\lambda,\xi,\theta)\right)\cos\left(2\theta\right) \\ n(y,\lambda,\xi,\theta) = \cos\left(\Delta\phi_1(y,\lambda,\xi,\theta)\right)\sin\left(2\theta\right) - \sin\left(\Delta\phi_1(y,\lambda,\xi,\theta)\right)\sin\left(\Delta\phi_2(y,\lambda,\xi,\theta)\right)\cos\left(2\theta\right) \\ p(y,\lambda,\xi,\theta) = \cos\left(\Delta\phi_1(y,\lambda,\xi,\theta)\right)\cos\left(2\theta\right)\sin\left(\Delta\phi_2(y,\lambda,\xi,\theta)\right) + \sin\left(\Delta\phi_1(y,\lambda,\xi,\theta)\right)\sin\left(2\theta\right) \end{cases} \tag{6.2.8}$$

or simply

$$\begin{cases} m(y,\lambda,\xi,\theta) = \cos\left(\Delta\phi_2\right)\cos\left(2\theta\right) \\ n(y,\lambda,\xi,\theta) = \cos\left(\Delta\phi_1\right)\sin\left(2\theta\right) - \sin\left(\Delta\phi_1\right)\sin\left(\Delta\phi_2\right)\cos\left(2\theta\right) \\ p(y,\lambda,\xi,\theta) = \cos\left(\Delta\phi_1\right)\cos\left(2\theta\right)\sin\left(\Delta\phi_2\right) + \sin\left(\Delta\phi_1\right)\sin\left(2\theta\right) \end{cases} \tag{6.2.9}$$

being implied that $\Delta\phi_1$ and $\Delta\phi_2$ are functions depending on y, λ, ξ, θ. With this notation, the detected intensity is:

$$I_{out} = \frac{1}{2}(I + Q \cdot m(y,\lambda,\xi,\theta) + U \cdot n(y,\lambda,\xi,\theta) + V \cdot p(y,\lambda,\xi,\theta)) \tag{6.2.10}$$

Because ξ and θ play a role only in the optimization of the device and for any configuration they have to be fixed, then for the simplicity of the notations we will ignore for the moment any variation of the m, n and p with regard to this angles. Thus, in a even more simplified notation we may write:

$$I_{out} = \frac{1}{2}(I + Q \cdot m(y,\lambda) + U \cdot n(y,\lambda) + V \cdot p(y,\lambda)) \tag{6.2.11}$$

By choosing arbitrary values for $\xi, \theta, \Delta n, h$ in agreement with the existing literature[32] we may attempt to plot m, n and p functions. Thus, for $\xi = 1.5^\circ$, $\theta = 65^\circ$, $\Delta n = 0.014$, the height of the wedges, $h = 1.3cm$ and a wavelength $\lambda = 0.9\mu m$ we will obtain the graphics below (Figure 6.4):

At this level of analysis the most important observations are that for this particular case (at least), m, n and p are different functions with only a limited number of points in common and which are not constant, neither with regard to y, neither with regard to λ.

The very important question that may arise at this moment, and which govern the entire project by its importance, is related to the unequivocal character of the modulation function I_{out}. And this question is: for any incoming vector S can be found another incoming vector S', $S' \neq S$, such that $I'_{out} = I_{out}$ for the same values of λ and y? We considered here that I_{out} is the detected intensity for an incoming polarization described by S, while I'_{out} corresponds to S'. In other words, the question is if we can have the same profile of the intensity obtained from different polarizations. We will consider that for both functions I_{out} and I'_{out} we look at the same values y and λ simultaneously because only in this scenario the system may lead to confusions.

In order to properly answer to this question, let us consider that the incoming S vector is:

$$S = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$$

and the corresponding detected intensity after the passage through the optical device is:

$$I_{out} = \frac{1}{2}(I + Q \cdot m(y,\lambda) + U \cdot n(y,\lambda) + V \cdot p(y,\lambda)) \tag{6.2.12}$$

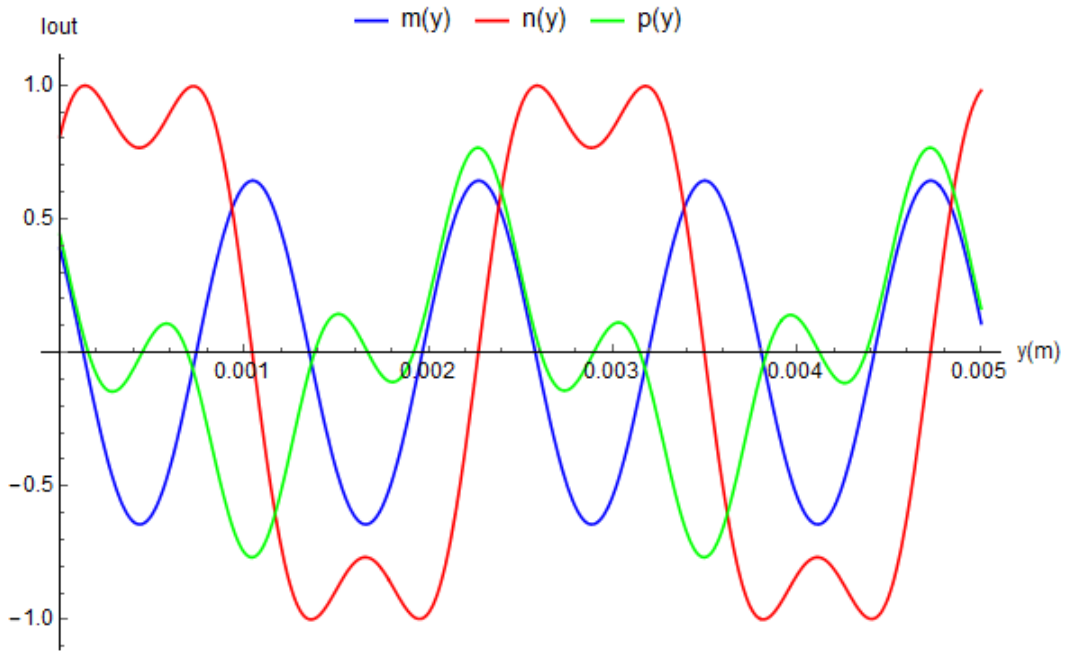


Figure 6.4: m(y),n(y) and p(y) for $\lambda = 0.9\mu m$ and y varying between 0 and 5mm

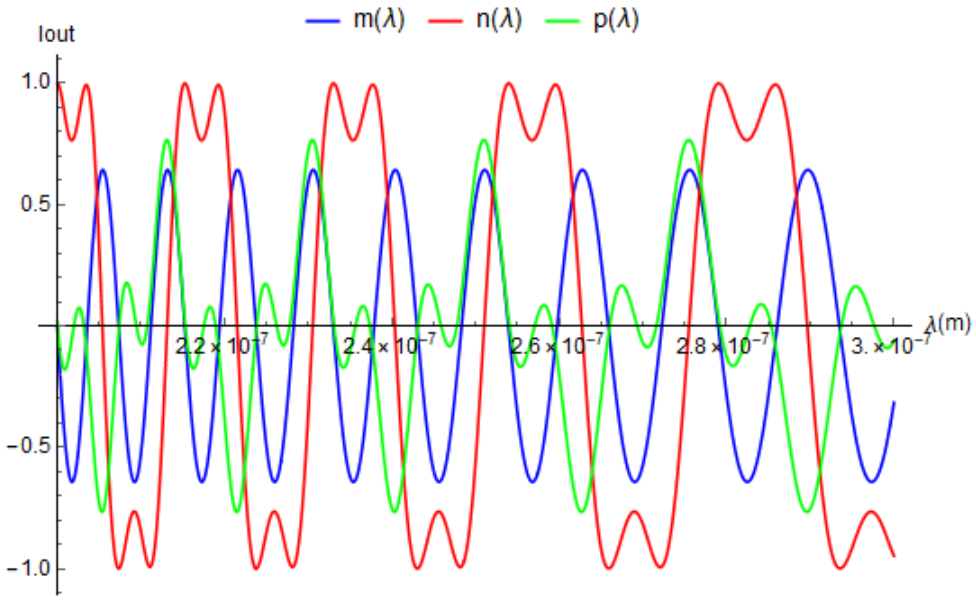


Figure 6.5: $m(\lambda), n(\lambda)$ and $p(\lambda)$ for $y = 0.5mm$ and λ varying between 0.2 and $0.3\mu m$

The second vector S' is:

$$S' = \begin{bmatrix} I' \\ Q' \\ U' \\ V' \end{bmatrix}$$

and the received intensity:

$$I'_{out} = \frac{1}{2}(I' + Q' \cdot m(y,\lambda) + U' \cdot n(y,\lambda) + V' \cdot p(y,\lambda)) \tag{6.2.13}$$

The hypothesis are that:

$$\begin{cases} S \neq S' \\ I_{out} = I'_{out} \end{cases}$$

But if $I_{out} = I'_{out}$ then $I_{out} - I'_{out} = 0$ which means:

$$\begin{cases} S \neq S' \\ \Delta I_{out} = I_{out} - I'_{out} = \frac{1}{2}(\Delta I + \Delta Q m(y,\lambda) + \Delta U n(y,\lambda) + \Delta V p(y,\lambda)) = 0 \end{cases}$$

Or this last identity is the intensity received by the detector for an incoming Stokes vector ΔS which is converted into ΔS_{out}:

$$\Delta S = \begin{bmatrix} \Delta I \\ \Delta Q \\ \Delta U \\ \Delta V \end{bmatrix} \rightarrow \Delta S_{out} = \begin{bmatrix} \Delta I_{out} = 0 \\ \Delta Q_{out} \\ \Delta U_{out} \\ \Delta V_{out} \end{bmatrix}$$

and tells us that there exists a polarization state that is completely destroyed by the device, at a such a level that no intensity can pass through. Or this could happen only if m, n, p functions are multiple one of another or simple constants. But we saw that this is not the case. As a consequence,

$$\Delta S = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

thus, $S = S'$, and we may conclude that for the same values of the wavelength and of y the system is univocal. Moreover, the conclusion is true only as long as y and λ are covering the minimal sampling distance. Otherwise no relation could be established between the incoming Stokes vector and the modulated intensity.

In a 2D representation, the variation of all the concerned modulating functions (m,n,p), again with regard to y and λ and for a normal incidence will have the allure from the Figures 6.6, 6.7, where an arbitrary spectral range between 4 and $6\mu m$ was chosen.

It is worthy to be mentioned that considered alone, each of the concerned functions will describe a particular incoming polarization, m corresponding to the linear case $S_{in} = [I, Q, 0, 0]$, n to a 45° polarization, $S_{in} = [I, 0, U, 0]$, while p will stay for the circular polarization, $S_{in} = [I, 0, 0, V]$.

Moreover, the graphics 6.6 and 6.7 shows clearly that this type of optical device is able to generate simultaneously a spectral and polarization modulation of the outgoing intensity. For each type of incoming polarization we will have a certain profile for the I_{out} function, profile that will vary almost linearly with λ.

In order to better seize the effect of the spectropolarimeter upon of an incoming ray of light with a certain polarization we will try to visualize a few particular cases.

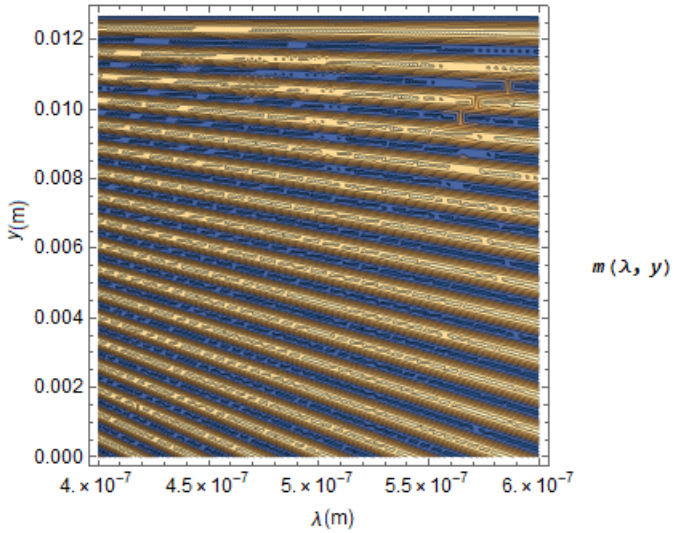


Figure 6.6: $m(y, \lambda)$ behavior in a contour plot on a y scale of $1.2cm$ and for λ between $400nm$ - $600nm$

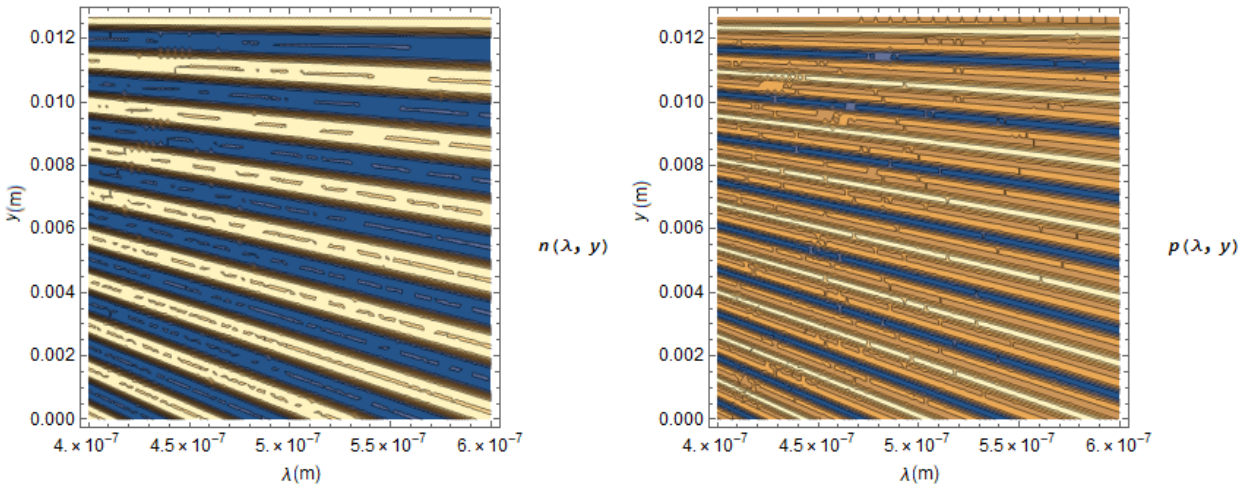


Figure 6.7: $n(y, \lambda)$ and $p(y, \lambda)$ behavior in a contour plot on a y scale of $1.2cm$ and for λ between $400nm$ - $600nm$

For example, let us consider at first that the incoming orthogonal ray is totally polarized, the corresponding Stokes vector being $S = [1, 0.55, 0.703, 0.45]$. The situation is represented in the graphics 6.8 and 6.9, where we can see a contour representation of the modulated intensity as well as a variation with respect to λ and y.

A second example could be a linear polarization (also in normal incidence), with a Stokes vector $S = [1, 0.3, 0.26, 0]$. This configuration corresponds to a degree of linear polarization of 0.4. The m, n, p functions are not affected by this new scenario because they depend only on y and λ. Instead the received intensity is linearly modulated by the Stokes parameters (Figure 6.10, 6.11).

Comparing the I_{out} profiles for the two situations (Figure 6.12) we may notice that, as could be expected, the decrease of the degree of polarization is translated into a general decrease of the received intensity (all the graphics here keeps the arbitrary units of the intensity: V or W/m^2) and in the change of the pattern.

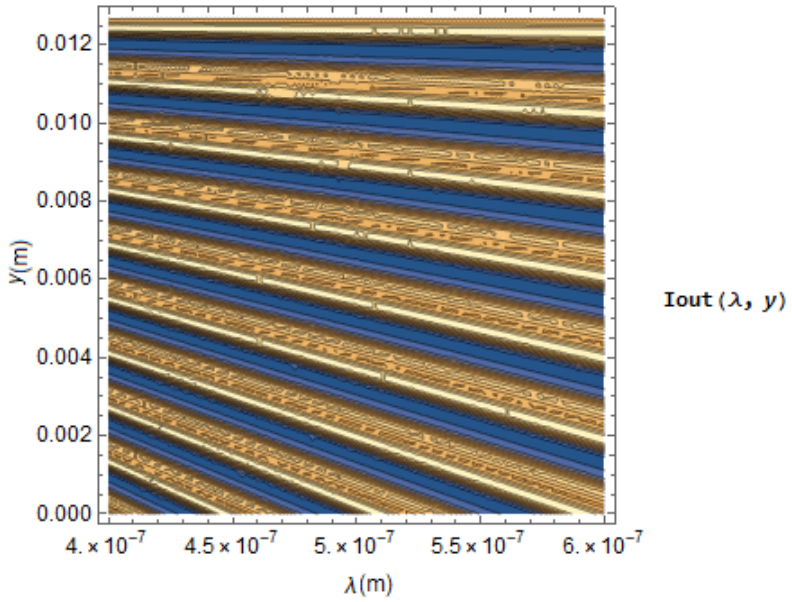


Figure 6.8: $I_{out}(\lambda, y)$ pattern for a totally polarized beam

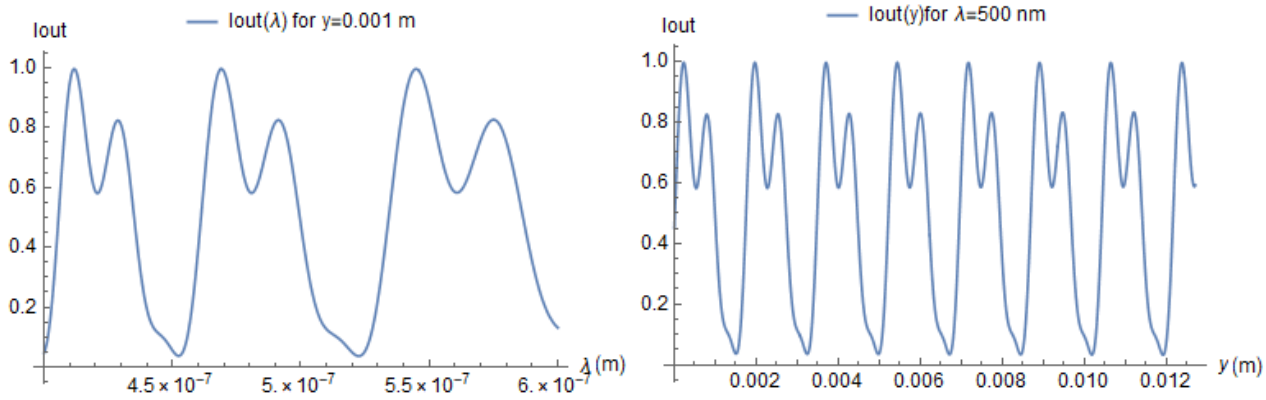


Figure 6.9: $I_{out}(\lambda)$ (left) and $I_{out}(y)$(right) in the same case of a totally polarized beam. The units of I_{out} are the units of I (most often $\frac{W}{m^2}$)

We can extend this comparison also to the circular polarization. Thus, for an incoming normal ray with $S = [1, 0, 0, 0.5]$ we will obtain the green curve from the graphics below (Figure 6.13).

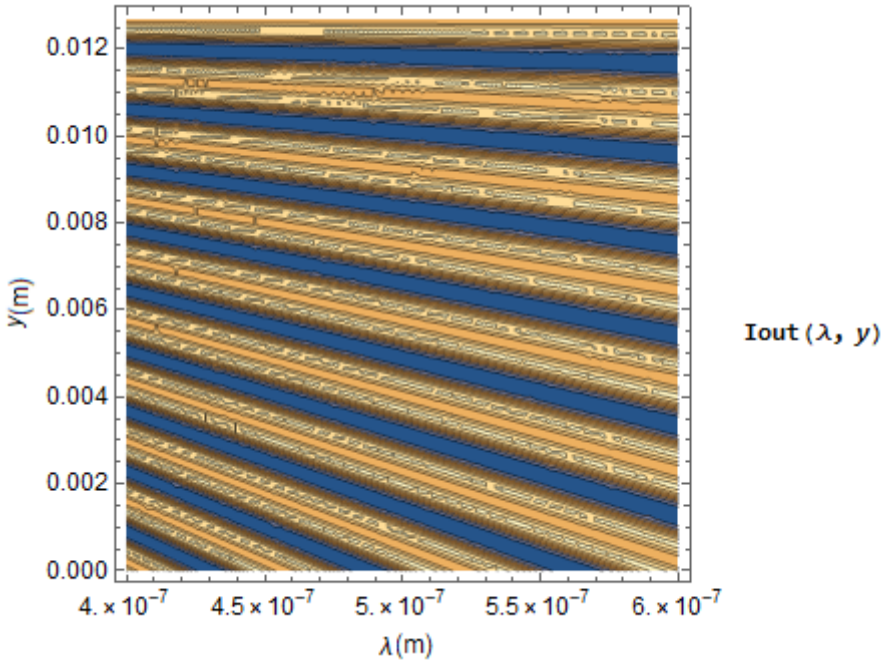


Figure 6.10: Pattern of the received intensity for a partially linearly polarized beam

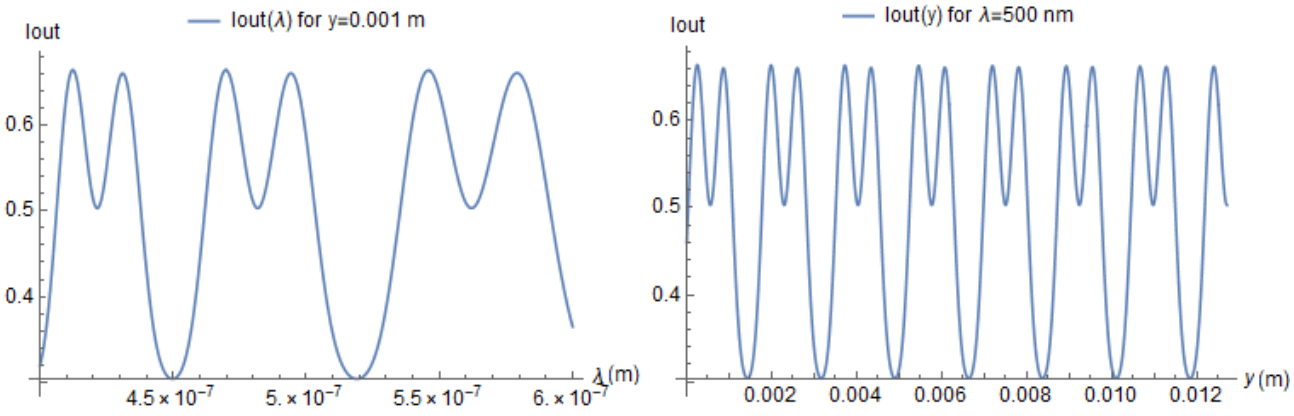


Figure 6.11: $I_{out}(\lambda)$ (left) and $I_{out}(y)$(right) in the same case of a partially linearly polarized beam. The units of I_out are the units of I (most often $\frac{W}{m^2}$)

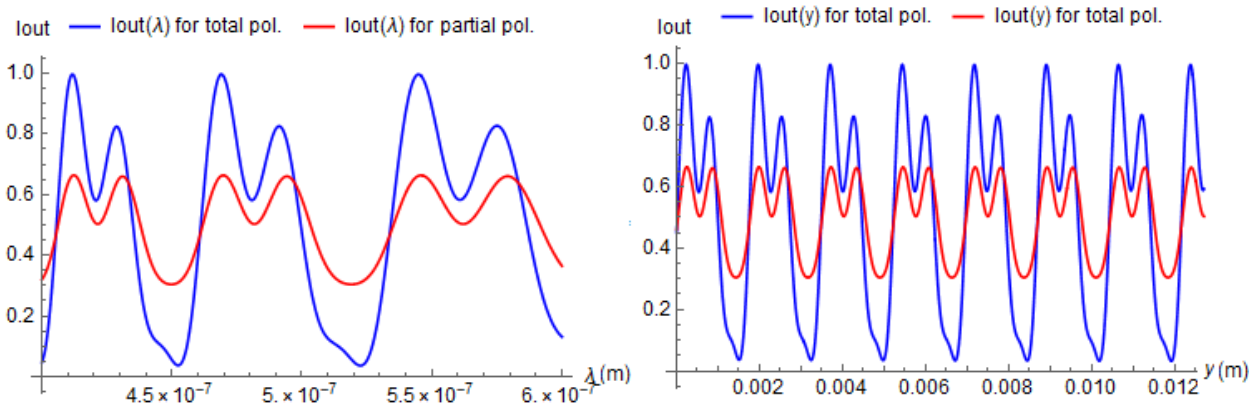


Figure 6.12: Comparison between the received intensity of the totally polarized beam and partially linearly polarized beam under normal incidence condition.

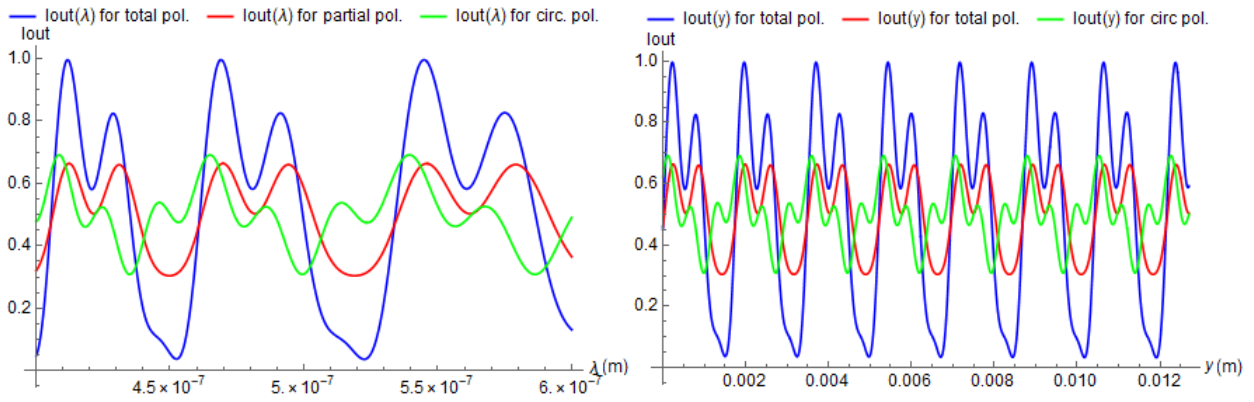


Figure 6.13: Comparison between the three studied cases: totally polarized light (blue), partially linearly polarized (red) and partially circularly polarized (green)

6.3 Mathematical optimization

In the most general way, the intensity received by the detector, depends on the position along y axis, on the apex angle ξ, on the wavelength λ, the analyzer angle θ and on the birefringence of the material, $\Delta n(\lambda)$.

Hence, for a given wavelength range and a certain material, $\Delta n(\lambda)$, it will be useful to check if there are optimal values for ξ and θ angles. At this point of our study we will try to figure out which are the best values for the angle of the analyzer, θ.

In order to do that, we have to come back to the general form of the intensity function, but this time written as:

$$I_{out}(\Delta\phi_1, \Delta\phi_2, \theta) = \frac{1}{2}(I + Q \cdot m(\Delta\phi_1, \Delta\phi_2, \theta) + U \cdot n(\Delta\phi_1, \Delta\phi_2, \theta) + V \cdot p(\Delta\phi_1, \Delta\phi_2, \theta)) \quad (6.3.1)$$

In order to simplify the notations we will consider that the variation with regard to $\Delta\phi_1, \Delta\phi_2$ is implied:

$$I_{out}(\theta) = \frac{1}{2}(I + Q \cdot m(\theta) + U \cdot n(\theta) + V \cdot p(\theta)) \quad (6.3.2)$$

The question about optimal value of θ can be translated then into a more precise one: for which values of θ the received intensity reaches the maximal value?

That means we have to look for the values of θ so that:

$$\frac{dI_{out}}{d\theta} = 0 \quad (6.3.3)$$

Developing this relation in just a few steps we may find that we are looking for an angle theta such as:

$$\tan(2\theta) = \frac{U \sin(\Delta\phi_1) \sin(\Delta\phi_2) - Q \cos(\Delta\phi_2) - V \cos(\Delta\phi_1) \sin(\Delta\phi_2)}{U \cos(\Delta\phi_1) + V \sin(\Delta\phi_1)} \quad (6.3.4)$$

This identity tells us that θ will depend all the time on the type of polarization. At a first sight we can say that if the polarization vector is of the type $S = [I, Q, 0, 0]$ then the received intensity will be maximized only if $\theta = \frac{\pi}{4}$. Nevertheless, this is a problematic value, because it will turn the m function into $m = 0$, which will erase any attempt to find the Q parameter.

For a better view of the situation we may try to represent the variation of θ that can be computed from the relation (6.3.4) as a function of λ and y for different types of polarizations.

In order to find the values of θ that are optimizing the functioning we will have to look for the points of intersection of these curves. Unfortunately, there is no such a point for all the situations. The most we can do is to find the angles that are maximizing simultaneously the reception of two states of polarization. Thus, by an inspection of the graphics 6.14 and 6.15 we may find:

- elliptical (total: $S = [1, 0.55, 0.703, 0.45]$) and linear polarization ($S = [1, 0.3, 0.26, 0]$):
  - -16.04°
  - 23.49°
  - 24.64°

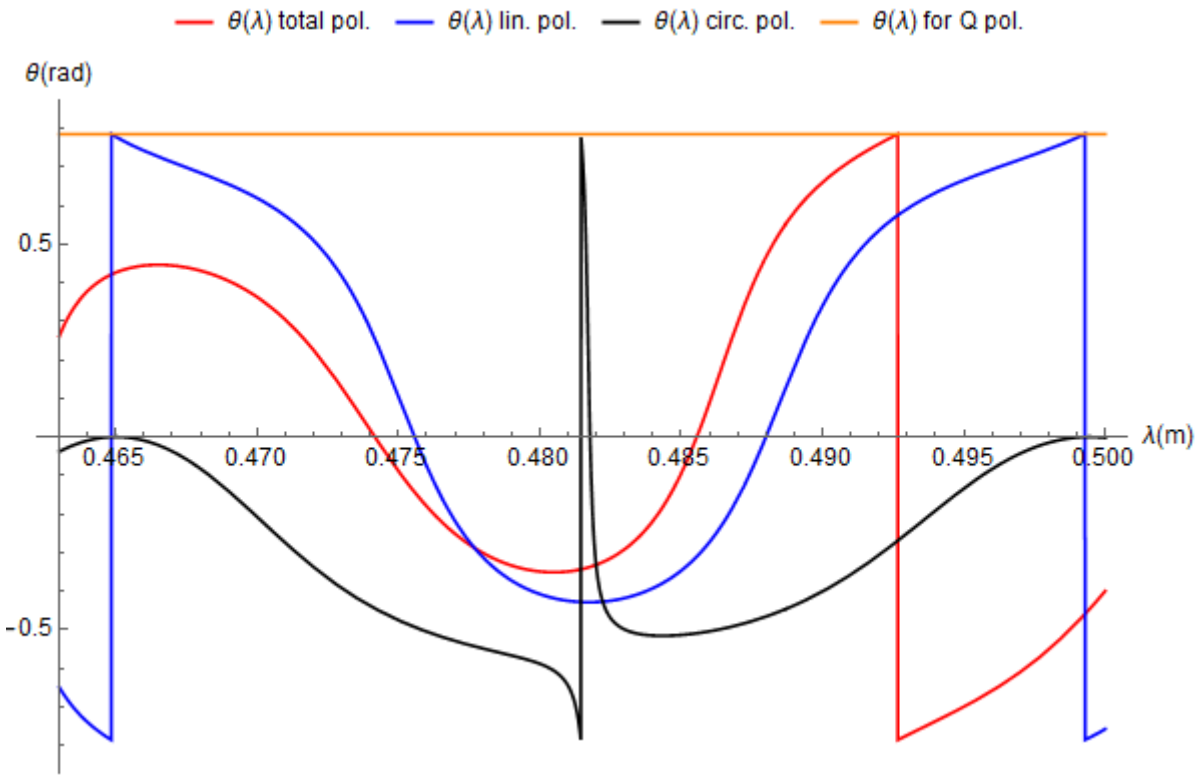


Figure 6.14: Variation of θ (radians) for four configurations of the incoming polarization: total polarization ($S = [1, 0.55, 0.703, 0.45]$), linear ($S = [1, 0.3, 0.26, 0]$), linear($S = [1, 0.55, 0, 0]$) and circular ($S = [1, 0, 0, 0.5]$) with regard to the wavelength (0.4-0.5)μm

  - 32.65°

- elliptical (total: $S = [1, 0.55, 0.703, 0.45]$) and circular polarization ($S = [1, 0, 0, 0.5]$):
  - -19.84°
  - -18.90°
  - -17.76°
- linear ($S = [1, 0.3, 0.26, 0]$) and circular polarization ($S = [1, 0, 0, 0.5]$):
  - -24.64°
  - -9.74°
  - -8.02°

Given the form of the "function" θ, it can be easily observed that is strongly dependent upon the Stokes parameters. Therefore, the values presented here are true only for the considered Stokes vectors. In order to have a reasonable choice of the angle of the analyzer we have thus to relate to practical purpose of the instrument: we need to have an idea about the expected values of the polarization.

The optimization of the device about ξ, the apex angle, is another issue that mast be addressed. This time, the analytic way is no more possible: the ξ angle is present at the level of the phase, as

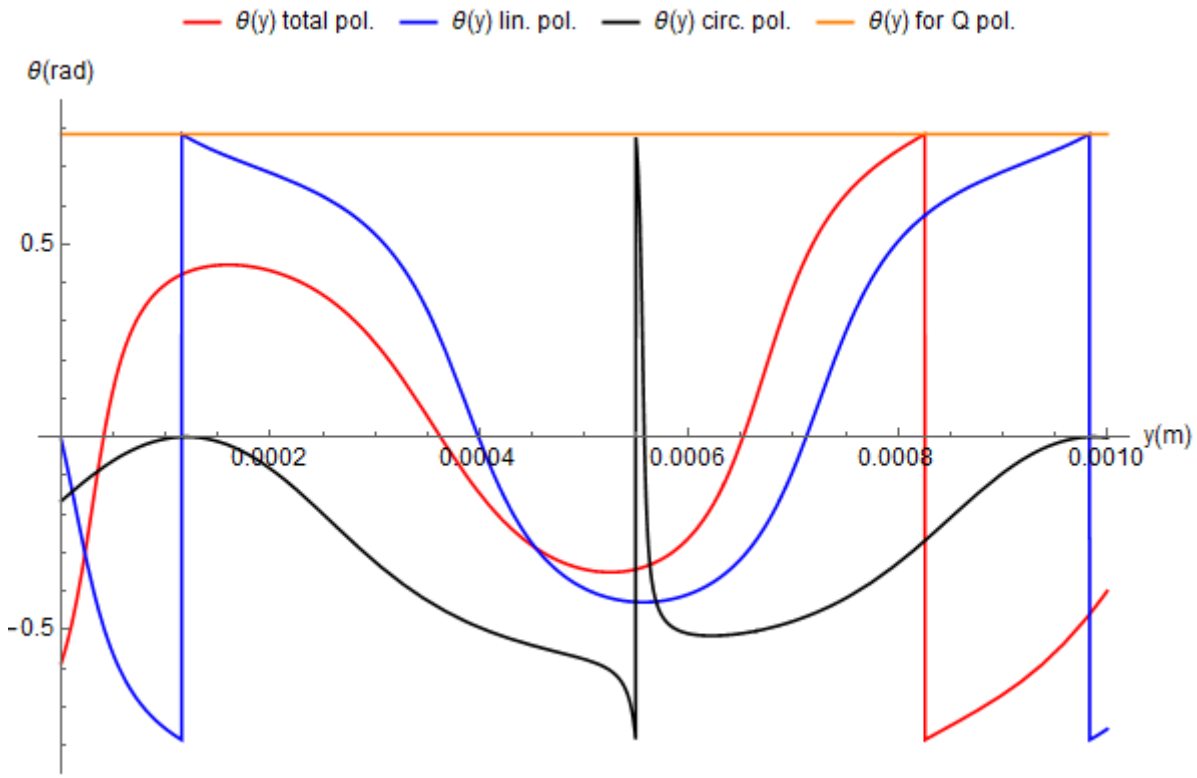


Figure 6.15: Variation of θ (radians) for four configurations of the incoming polarization: total polarization ($S = [1, 0.55, 0.703, 0.45]$), linear ($S = [1, 0.3, 0.26, 0]$), linear($S = [1, 0.55, 0, 0]$) and circular ($S = [1, 0, 0, 0.5]$) with regard to the position along y (0-1)mm

$\tan(\xi)$ and $\tan(2\xi)$. Because of this, the derivation of the intensity function will not be able to provide an easy and direct answer. Consequently, the choice of this angle will be based mostly on physic criteria. First, we saw in the previous chapter that even for the normal incidence an incoming ray will be split by this device because of the angles ξ and 2ξ coupled with the effect of the orientation of the optical axis. In order to minimize this effect of splitting, the best strategy is to consider small angles so that the separation surfaces could be assimilated, at the limit, to parallel planes.

On the other hand, the best scenario will correspond also to an angle ξ that has constant impact over the entire wavelength range.

Plotting the outgoing intensity as a function of ξ and λ we can notice that an almost constant profile is obtained for small apex angles (somewhere below 0.03 radians), as can be seen from the Figure 6.16. And this behavior is symptomatic for the entire range of y. From these observations we may conclude that the best idea it will be to work with small values of ξ if we want to have a constant efficiency of the polarimeter all along the spectral range. The value of 1.5° used until is thus in agreement with these remarks.

Such an assumption, will give us the opportunity to approximate

$$\tan(2\xi) \approx 2 \cdot \tan(\xi)$$

and, as a direct consequence we will have

$$\Delta\phi_2(\xi) = 2\Delta\phi_1(\xi)$$

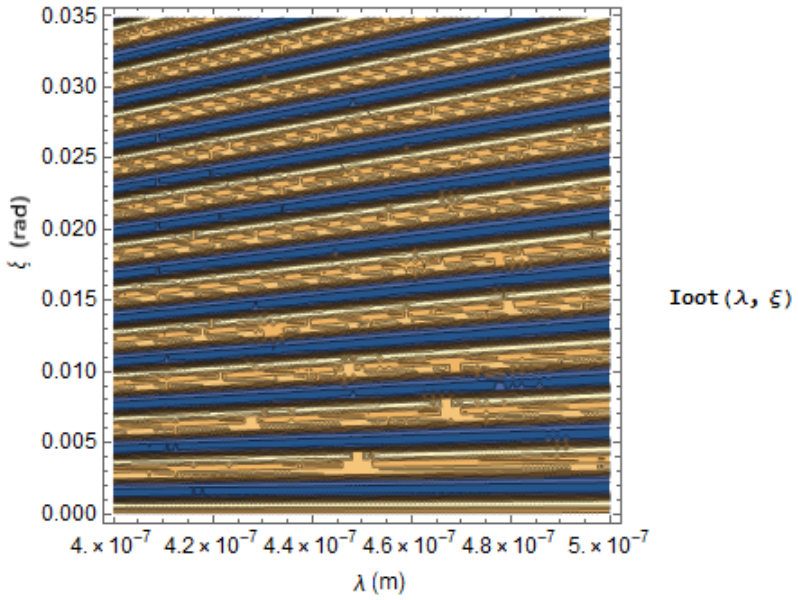


Figure 6.16: Variation of the outgoing intensity, I_{out} with regard to λ and ξ (radians), in a contour plot. The assumption of the incoming elliptical polarization was used: $S = [1, 0.55, 0.703, 0.45]$

Based on this simplifications, the expressions from (6.2.9) will become:

$$\begin{cases} m(\theta, \Delta\phi_1) = \cos(2\theta)\cos(2\Delta\phi_1) \\ n(\theta, \Delta\phi_1) = \sin(2\theta)\cos(\Delta\phi_1) - \cos(2\theta)\sin(\Delta\phi_1)\sin(2\Delta\phi_1) \\ p(\theta, \Delta\phi_1) = \cos(\Delta\phi_1)\cos(2\theta)\sin(2\Delta\phi_1) + \sin(2\theta)\sin(\Delta\phi_1) \end{cases} \tag{6.3.5}$$

Or for a given value of θ, the I_{out} function will be modulated according to three frequencies, which can be reveled through a Fourier analysis. Thus, for the terms $\cos(\Delta\phi_1)$ or $\sin(\Delta\phi_1)$ the specific period will be, obviously, 2π, for $\cos(2\Delta\phi_1)$ or $\sin(2\Delta\phi_1)$ we have a period of π, while the terms in $\sin(\Delta\phi_1)\sin(2\Delta\phi_1)$ or $\cos(\Delta\phi_1)\sin(2\Delta\phi_1)$ will introduce a period of $\frac{2\pi}{3}$. But in order to properly use this new set of information, the concept of birefringence should be thoroughly explained in a new chapter.

6.4 Conclusion

The problematic studied in this chapter, based only on the hypothesis of the normal incidence, gives us the possibility to formulate several ideas extremely important for our research:

- for a ”classical” polarimeter the need of rotation comes from the ”incapacity” of the device to provide simultaneously access to all the Stokes parameters;
- the model based on three birefringent prisms can ensure a spatial and spectral modulation of the received intensity and a full determination of the Stokes vector without using turning parts;
- for any given wavelength this type of system is univocal, being impossible to have multiple different entries with the same output;
- the optimization with regard to the analyzer angle depends on the scientific goal: this choice is governed by the type of polarization that interests us, and by the mean values of this polarization;
- the optimization with regard to the apex angle ξ suggest the use of values below 1.72°;
- the signal of the outgoing intensity exhibits three periods, which are $\frac{2\pi}{3}$,π and 2π (if we use the approximation $\Delta\phi_2(\xi) = 2\Delta\phi_1(\xi)$).

Chapter 7

The impact of the chromaticity

In the previous chapter we saw that the main physical characteristic of the proposed spectropolarimeter, that ensure its functioning, is the capacity to entail a certain phase difference between the orthogonal components of the incoming beam, $\Delta\phi$. But this capacity is based on the birefringence of the medium, Δn (i.e. the variation of the refraction index with respect to the position towards the optical axis):

$$\Delta n = |n_o - n_e|$$

where n_o and n_e are the ordinary and extraordinary indices of refraction.

Yet, for all the material used in optics, the indices of refraction are varying with the wavelength, property which is called chromaticity. Because of this, the correct expression of the birefringence is:

$$\Delta n(\lambda) = \left|n_o(\lambda) - n_e(\lambda)\right|$$

Therefore, in the study of the behavior of the spectropolarimeter we must pay attention also to the chromaticity. Withal, the transmittance of the medium is also of paramount importance: we must be sure that for the desired wavelength range the media exhibit a good transmittance.

Based on this last feature, a good candidate for our device could be Magnesium Fluoride (MgF_2). This material shows a steady and high transmittance for 0.12 until $7\mu m$ (Figure 7.1). In the same time, is very hard and rugged, resistant to mechanical and thermal shocks [3].

According to data available today [13] [6] the ordinary (n_o) and extraordinary (n_e) indexes of refraction of MgF_2 varies like in the Figure 7.2.

Thus, the birefringence $\Delta n(\lambda)$ can be accurately calculated for any wavelength from the available spectral range. The result for MgF_2 is plotted in the Figure 7.3.

The useful span of the birefringence is then comprised between the cut-off wavelength of about $100nm$ and $9\mu m$ which means that MgF_2 material can cover most of the UV spectrum, the visible and mid infrared. However, we must pay attention to the fact that at $\lambda = 0.12\mu m$, the birefringence of the material is zero, which means the modulation function will be equally zero. In order to avoid this situation, we can propose to start at $0.125\mu m$.

Nevertheless, in order to fully profit of the capacities of this fluoride, we have to find a good mathematical description of the experimental data of the birefringence. And two methods are at hand: a polynomial interpolation and the Sellmeier model. Let's start with the first.

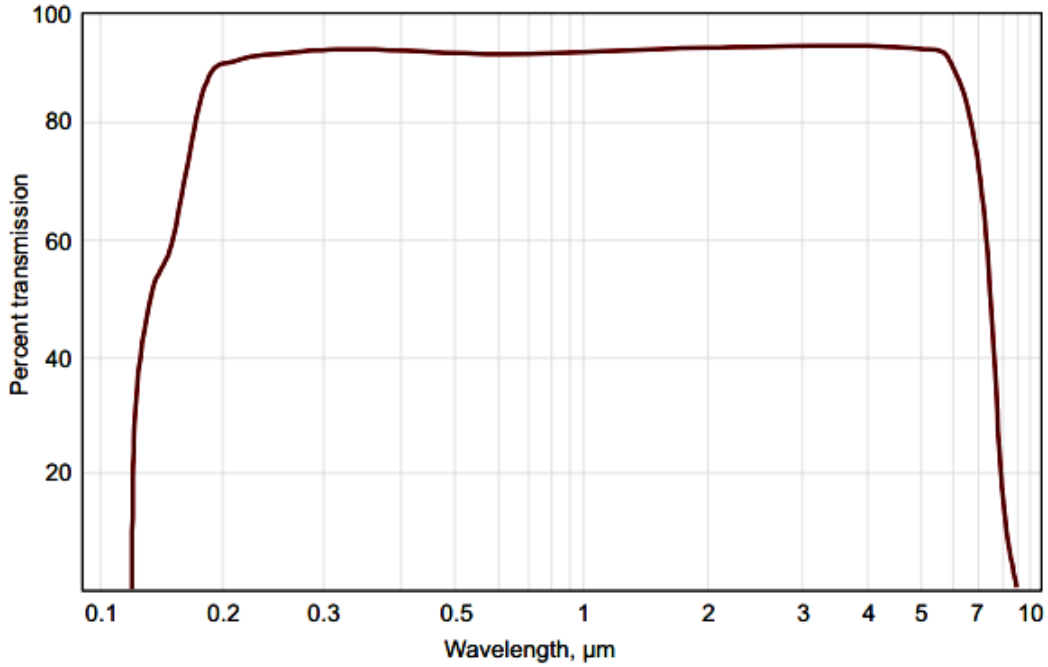


Figure 7.1: Magnesium Fluoride (MgF_2) has a transmission range of 0.12-7μm at a thickness of 10mm. Source: eksmaoptics.com[3]

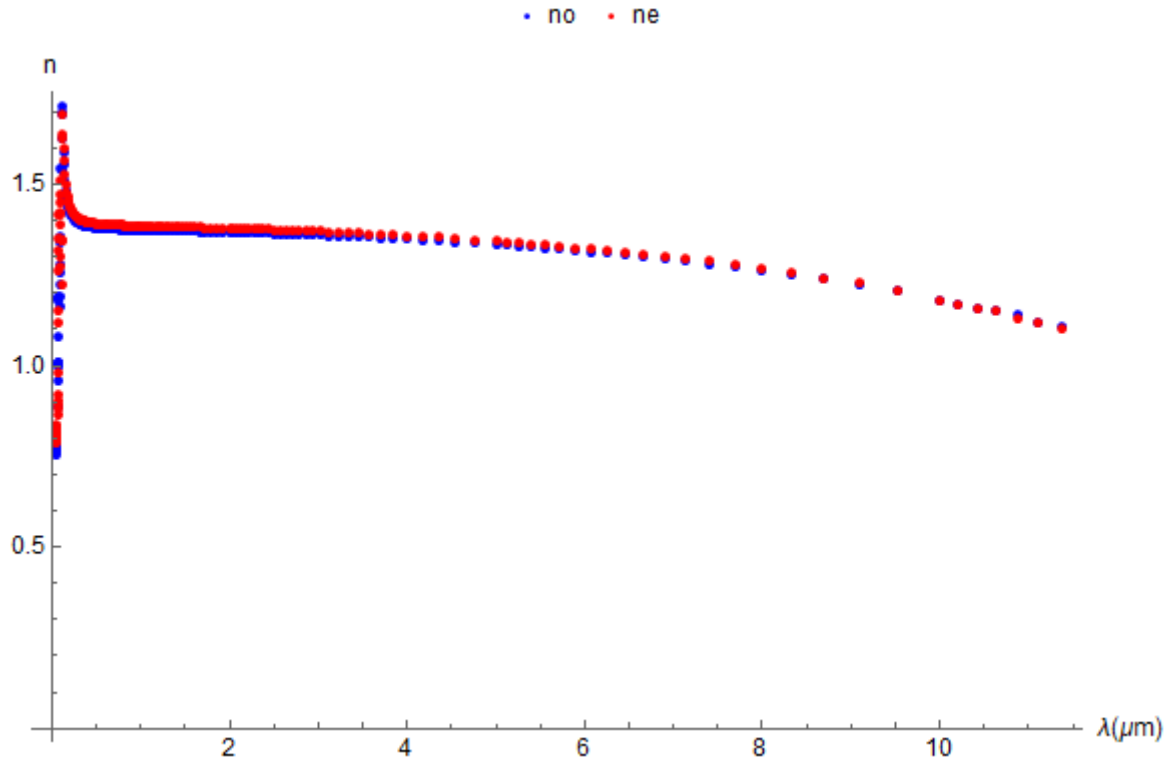


Figure 7.2: n_o and n_e of MgF_2 for $\lambda \in [0.0459; 11.364]\mu m$ [5]

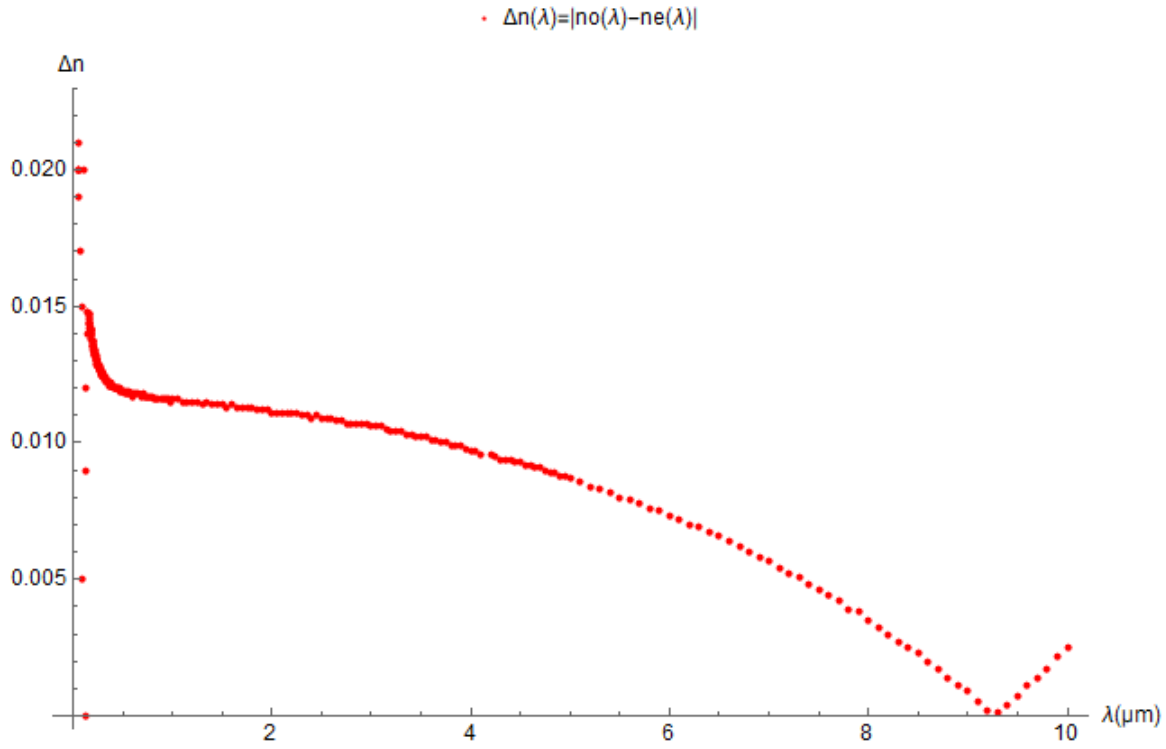


Figure 7.3: Birefringence of the Magnesium Fluoride:$\Delta n(\lambda) = \left|n_o(\lambda) - n_e(\lambda)\right|$

7.1 Interpolation of the birefringence data

Given the allure of the graphic for the variation of the birefringence, the best way to interpolate is to split the wavelength range in several intervals. Then, for each region we will search for the appropriate polynomial interpolation based on the least square method. Looking closely to the available data, which can be retrieved in Cotter [13], Dodge [6] and also in Appendix A of this thesis, three regions of interest can be dissociated between $0.125\mu m$ and $7\mu m$, according to fitting possibilities:

- 0.125 - 0.17 μm;
- 0.17 - 0.7 μm;
- 0.7 - 7 μm;

The best polynomial results for this intervals are (for the same order):

$$\begin{cases} 366087\lambda^5 - 275758\lambda^4 + 83003\lambda^3 - 12482.8\lambda^2 + 938.235\lambda - 28.1908 \\ -0.376181\lambda^5 + 0.95257\lambda^4 - 0.956983\lambda^3 + 0.481885\lambda^2 - 0.124377\lambda + 0.0254522 \\ -8.51 \cdot 10^{-7}\lambda^5 + 1.7075 \cdot 10^{-5}\lambda^4 - 1.3253 \cdot 10^{-4}\lambda^3 + 3.6696 \cdot 10^{-4}\lambda^2 - 8.116 \cdot 10^{-4}\lambda + 0.0121 \end{cases} \tag{7.1.1}$$

In the Figure 7.4 we can see the variation of the birefringence according to experimental data and the interpolation with the retrieved polynomials between 0.125 - $7\mu m$.

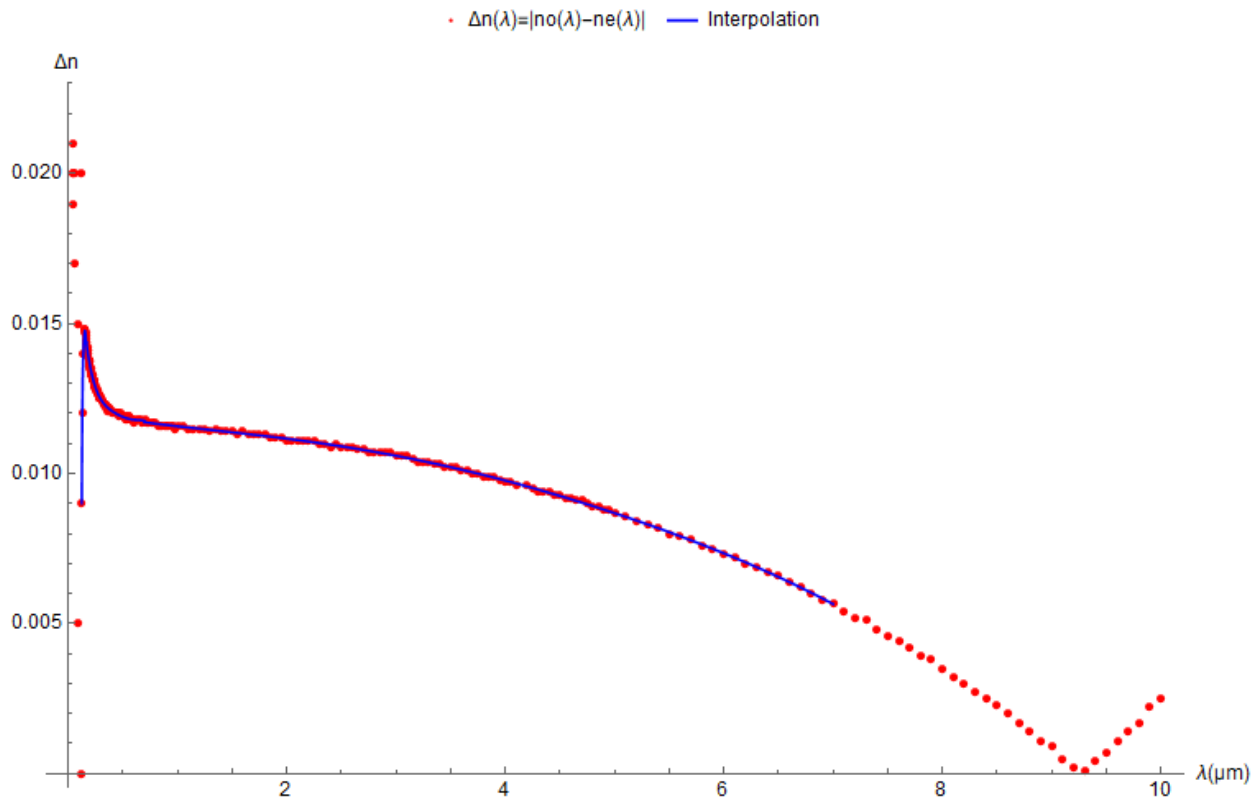


Figure 7.4: Birefringence data interpolated with three polynomials over the wavelength range 0.125-7μm

In the end of the last chapter we saw that the modulation of the signal is fallowing three periods, $\frac{2\pi}{3}$, π, 2π and. Consequently, according to Shannon theorem, the information should be sampled over an interval twice as long as the longest period (which means 4π) and the sample frequency should be at least twice the highest frequency, which in our case is $\frac{3}{2\pi}$, thus $\frac{3}{\pi}$. That means the signal should be sampled over 4π, with a minimum phase step of $\frac{\pi}{3}$. Meanwhile, because the phase is expressed as

$$\Delta\phi(y,\lambda,\xi) = \frac{2\pi}{\lambda}\Delta n(\lambda)(h-y)\tan(\xi) \tag{7.1.2}$$

it results that for a total phase difference of 4π, corresponding to the minimal sampling distance we will need, at an apex angle $\xi = 1.5^{\circ}$ at least $0.51mm$ in UV, at $125nm$, while observing at $1.4\mu m$, for example, will require a wedge with a high of at least $9.32mm$, according to the experimental data.

Using the polynomial interpolation described above, we can compute the variation of the minimal sampling distance with regard λ for different values of the apex angle, ξ. A sample of this evolution is expressed in Table 7.1, while in Figures 7.5 and 7.6 we can see the general evolution of the sampling distance with the wavelength.

Even though for certain interpolations and simulations we will still need a large wavelength range, for calculations meant to provide a realistic view about the studied device we will restrain the interval of wavelength to 0.125-1.4μm. The main reason for this is that this wavelength range is the most desirable from an astronomical point of view, the most important polarization phenomenon taking place into this interval, and because there is no detector able to cover a longer waveband.

Nevertheless, the same computation and simulations can be conducted as well for another spectral range within the limitations of MgF_2 and of the detectors.

	$\lambda = 125nm$	$\lambda = 1.4\mu m$
$\xi = 1.5^{\circ}$	$0.513mm$	$9.32mm$
$\xi = 3^{\circ}$	$0.256mm$	$4.66mm$

Table 7.1: Minimal sampling distances for two values of the apex angle ξ and two values of the wavelength, λ

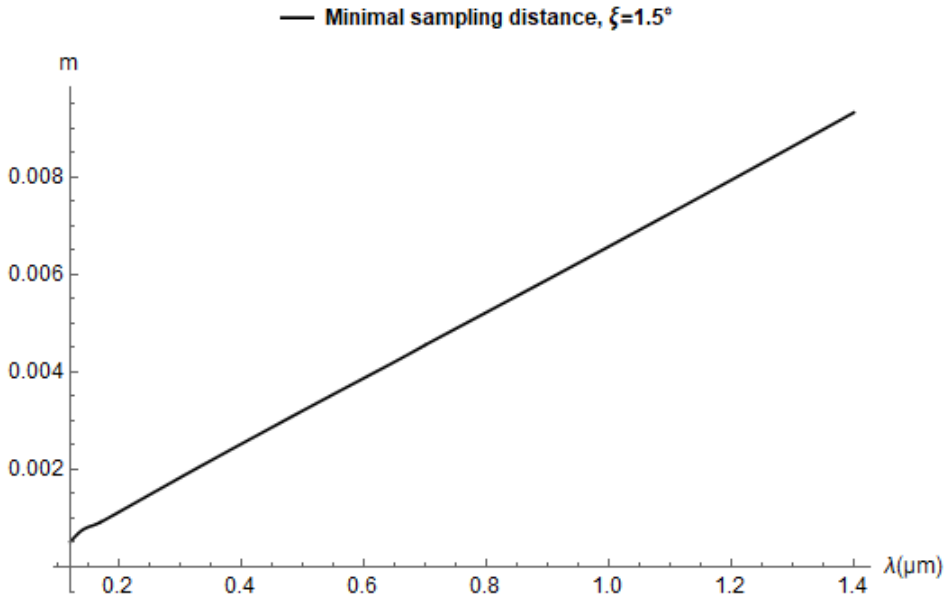


Figure 7.5: Minimal sampling distance as a function of the wavelength for an apex angle $\xi = 1.5^{\circ}$

This general variation of the minimal distance required for retrieving the signal gives us an important information about the dimension of the optical device, because is nothing else but the minimal height of the wedge of MgF_2. Thus, we may conclude that, in a strong dependence with the apex angle and the desired range of the wavelength, the minimal height of the wedge is comprised between $0.2mm$ and $9.3mm$.

7.2 Sellmeier method

Instead of using a polynomial interpolation of the birefringence we can use the empirical method of Sellmeier, which states that the refraction index of a material varies with the wavelength after a relation like:

$$n^2 - 1 = \sum \frac{A_i \cdot \lambda^2}{\lambda^2 - B_i^2}$$

or, sometimes:

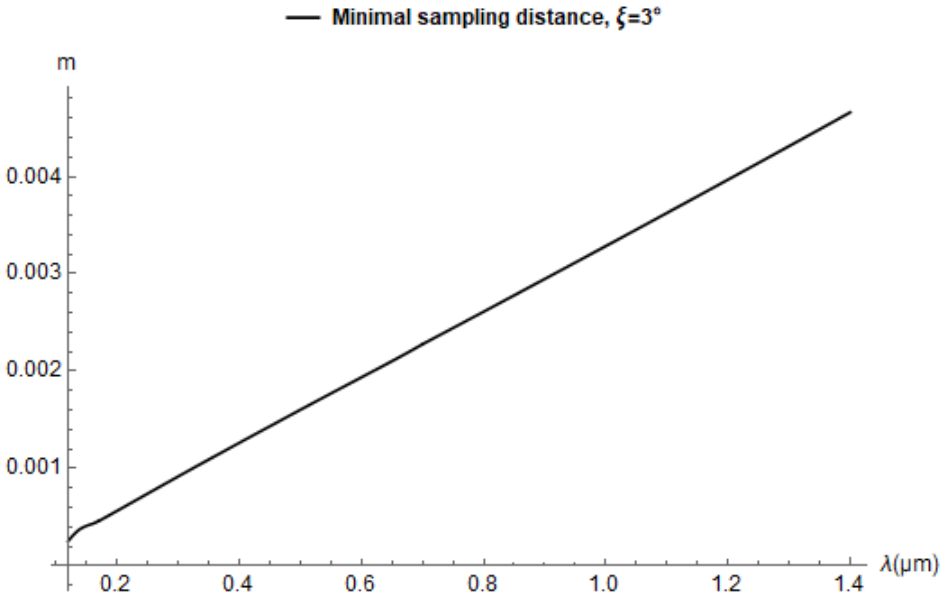


Figure 7.6: Minimal sampling distance as a function of the wavelength for an apex angle $\xi = 3^{\circ}$

$$n^2 - 1 = C + \sum \frac{A_i \cdot \lambda^2}{\lambda^2 - B_i^2}$$

The best fitting using this method and for longest wavelength range $(0.14 - 7.5\mu m)$ was provided by Li[27]:

$$n_o^2 - 1 = \frac{0.60967\lambda^2}{\lambda^2 - 0.08636^2} + \frac{0.008\lambda^2}{\lambda^2 - 18^2} + \frac{2.14973\lambda^2}{\lambda^2 - 25^2} + 0.2762 \tag{7.2.1}$$

For the extraordinary index we have:

$$n_e^2 - 1 = \frac{0.66405\lambda^2}{\lambda^2 - 0.08504^2} + \frac{1.0899\lambda^2}{\lambda^2 - 22.2^2} + \frac{0.1816\lambda^2}{\lambda^2 - 24.4^2} + \frac{2.1227\lambda^2}{\lambda^2 - 40.6^2} + 0.25385 \tag{7.2.2}$$

where the wavelength is expressed in μm.

In the Figure 7.7 we can see the plot of this fitting.

Both models seem to fit very well to the experimental data, the only difference being represented by the wavelength range. Nevertheless, a comparative analysis based on the residual values (real data - interpolated data) shows that there is a small difference in between.

Thus, the comparison of the residuals (see Figures 7.8, 7.9 and 7.10) shows that the fitted model using the Sellmeier approach works better for λ below $1.9\mu m$. Above this value, the polynomial interpolation gives the best results. Of course, because the Sellmeier method work only for the interval $0.14 - 7\mu m$ the comparison concerned only this wavelength range.

Consequently, for the proposed interval of study, $0.125\mu m$ - $1.4\mu m$, the recommendation will be to use the Sellmeier method with Li coefficients on $0.14 - 1.4\mu m$ and the polynomial interpolation (with the first polynomial) for $0.125 - 0.14$ wavelength range.

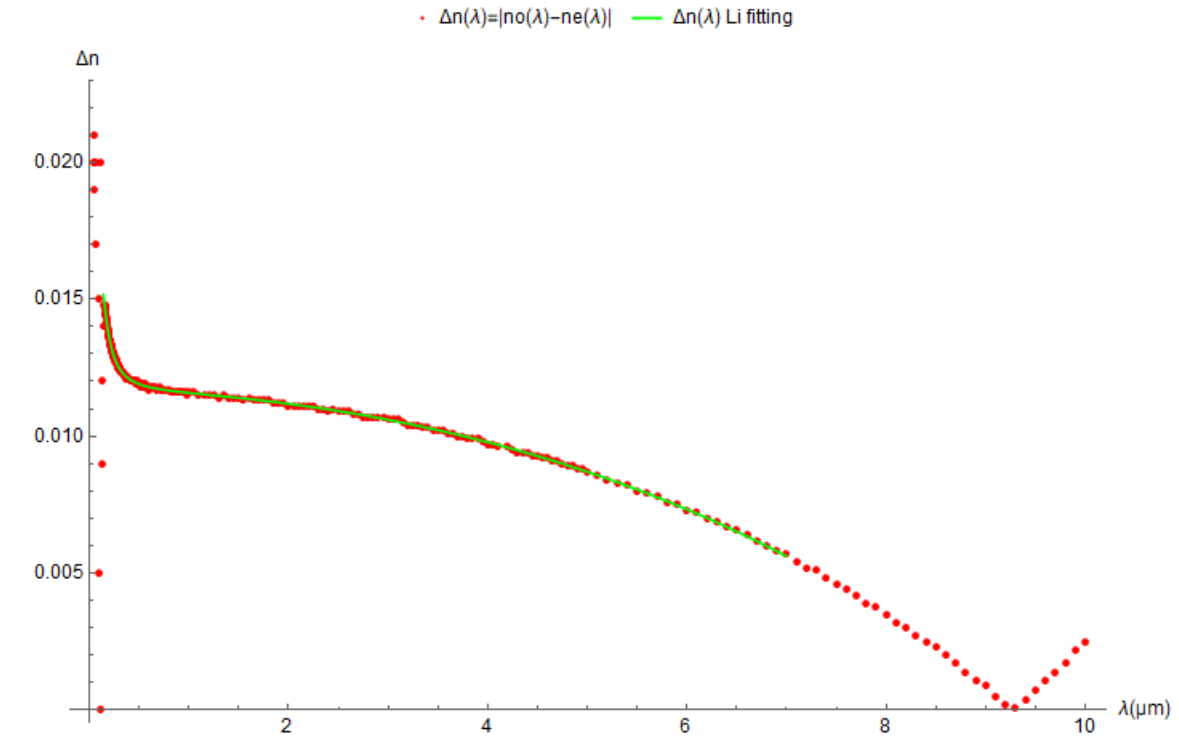


Figure 7.7: Fitting of the experimental data with the Sellmeier method between $0.14 - 7\mu m$ using Li polynomials

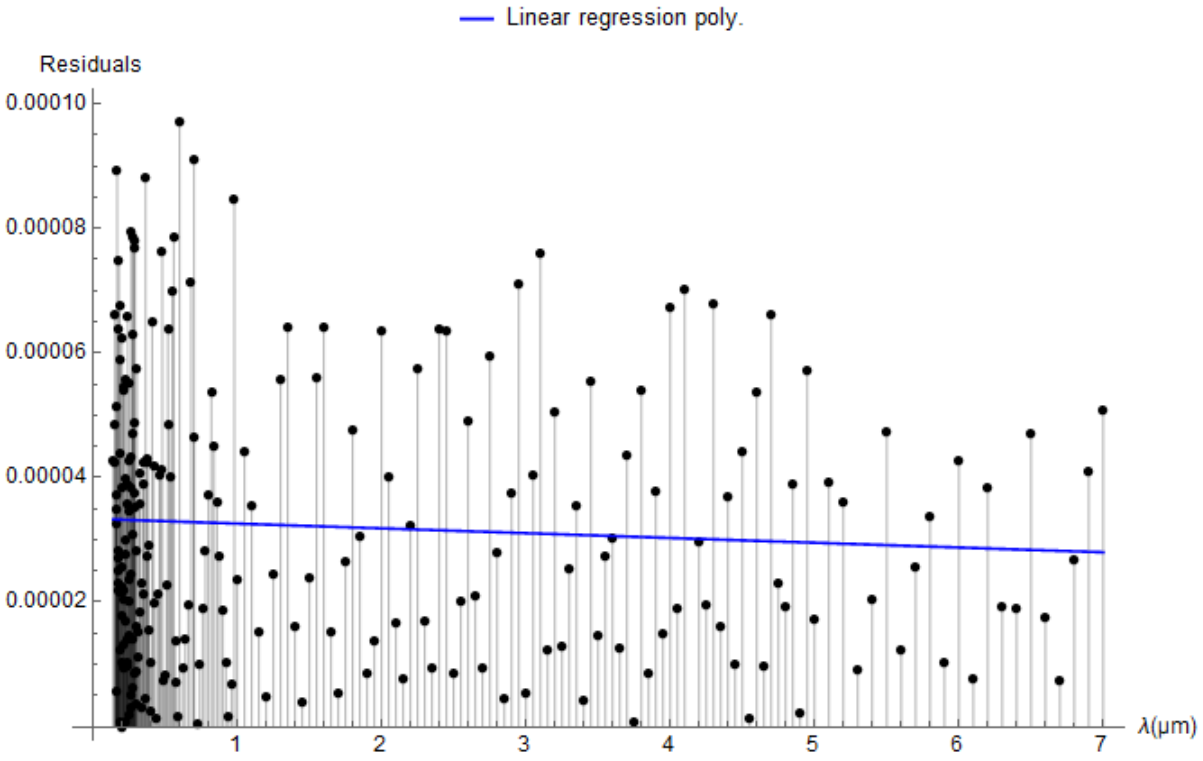


Figure 7.8: Residual values (blue dots) for $\left|\Delta n(experimental) - \Delta n(polynomial)\right|$ between $0.14 - 7\mu m$. The blue line corresponds to a linear regression for this set of points

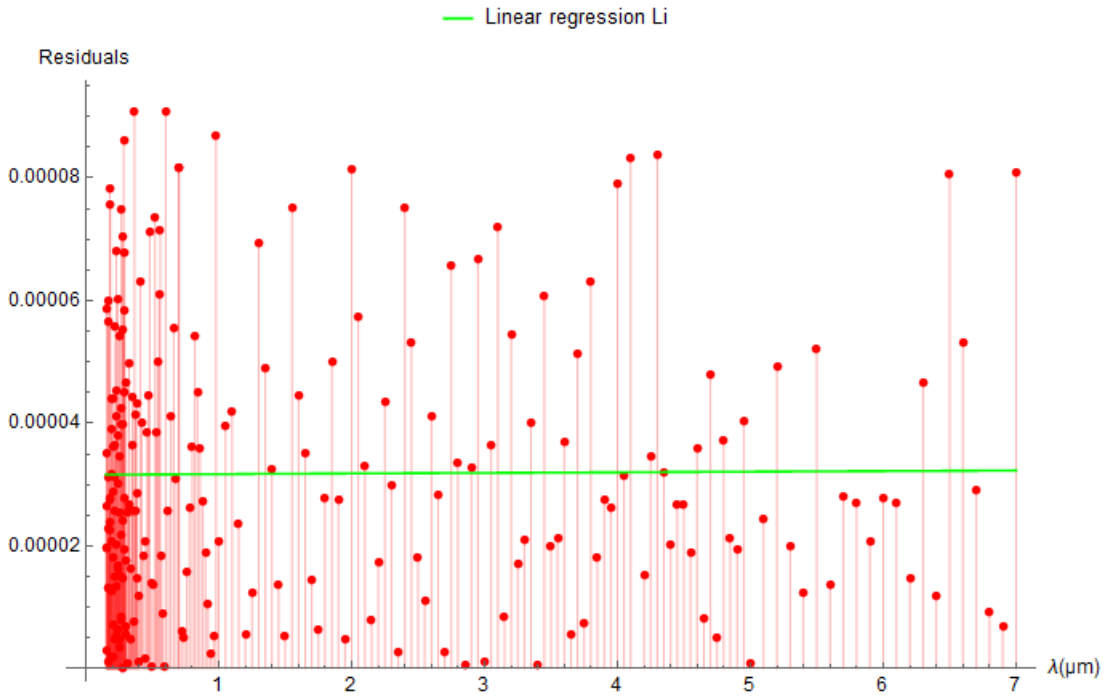


Figure 7.9: Residual values (blue dots) for $\left|\Delta n(experimental) - \Delta n(Sellmeier)\right|$ between $0.14 - 7\mu m$. The green line corresponds to a linear regression

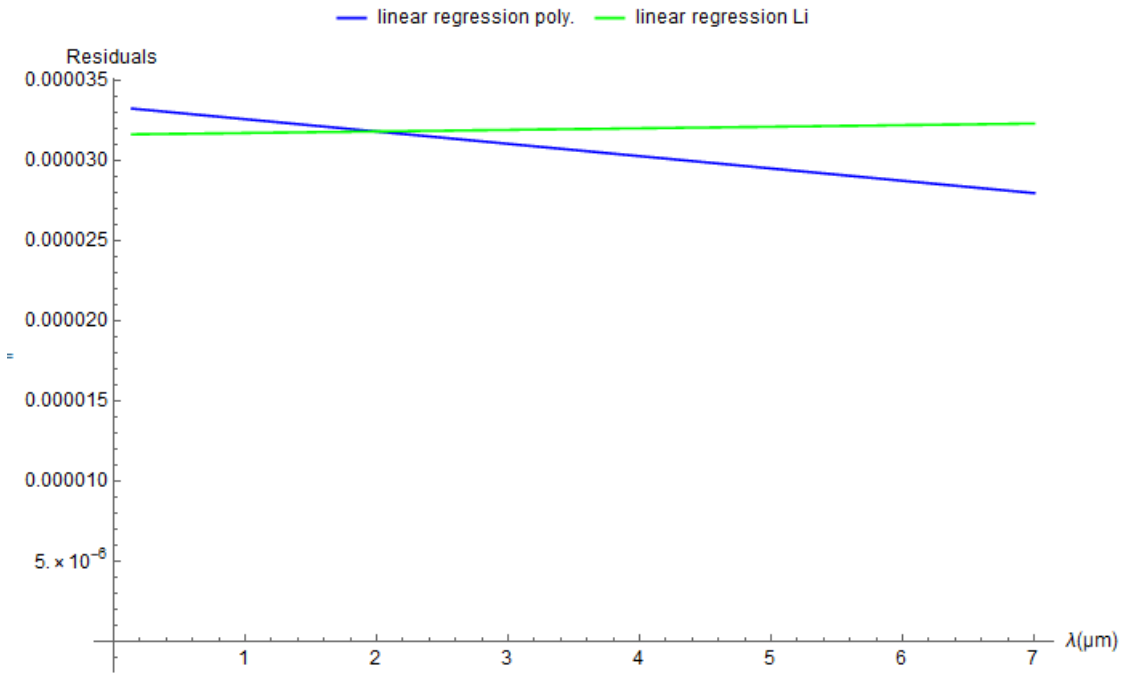


Figure 7.10: Comparison between the linear regressions in the two cases. The green line corresponds to the linear regression for the Sellmeier residuals while the blue stays for the polynomial method. Below $\lambda = 1.9\mu m$ the model Sellmeier model with the Li parameters has the lowest value of residuals.

7.3 Impact of the chromaticity on the modulation function

We established that between $\lambda = 0.14\mu m$ and $\lambda = 1.4\mu m$ the best model able to describe the birefringence variation is given by the Sellmeier method and Li coefficients. Meanwhile, because the Sellmeier method works only for $\lambda \geq 0.14\mu m$, for the range $0.125 - 0.14\mu m$ we should use the first polynomial from (7.1.1). We have to see now how this variation of the birefringence is translated at the level of the modulated intensity, $I_{out}(y, \lambda)$. Because the term $\Delta n(\lambda)$ is present at the level of the phase difference it would be expected to alter the periodicity of the outgoing intensity and the linearity with the wavelength. Both phenomena are certified by the plots from Figure 7.11 and 7.12, where the same state of polarization was represented by the mean of the received intensity with respect to the wavelength and with the position on vertical, y. The change of the periodicity is translated also into the change of sampling distance. According to the new values of the birefringence we may adjust the Table 7.1.

	$\lambda = 125nm$	$\lambda = 1.4\mu m$
$\xi = 1.5^{\circ}$	$0.51mm$	$9.38mm$
$\xi = 3^{\circ}$	$0.255mm$	$4.68mm$

Table 7.2: Minimal sampling distances for two values of the apex angle ξ and two values of the wavelength, λ in the case of variable birefringence

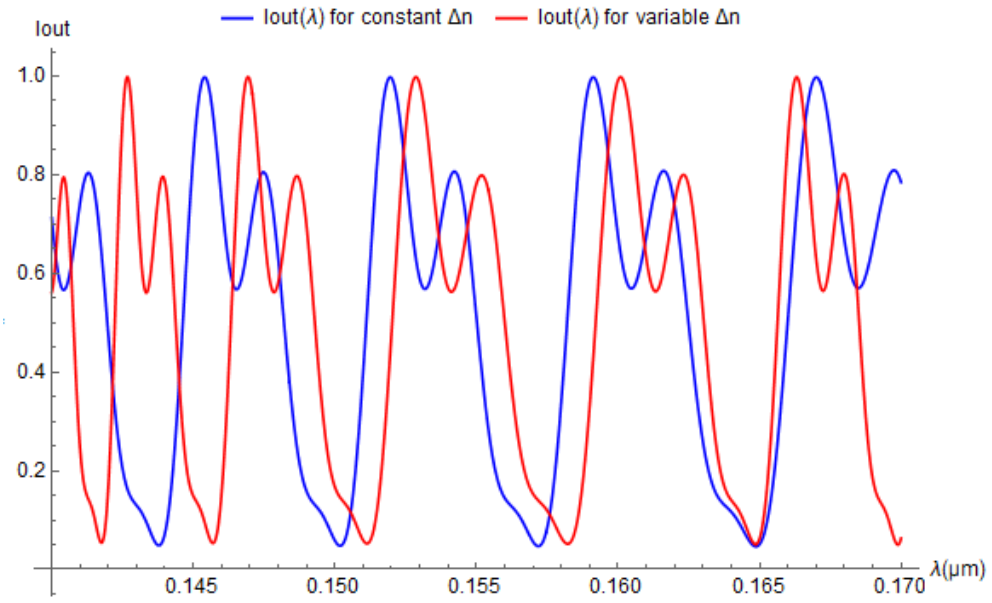


Figure 7.11: Comparison between the outgoing intensity in the case of a constant Δn and variable Δn for $\lambda \in [0.14 - 0.17]\mu m$. The hypothesis of normal incidence and elliptical polarization was used again.

Overall, the effect of the chromaticity is quite negligible. It introduces a slight variation of the periodicity of the signal but without any effect to the level of the received intensity. This one is affected only by the values of the Stokes parameters of the incoming light.

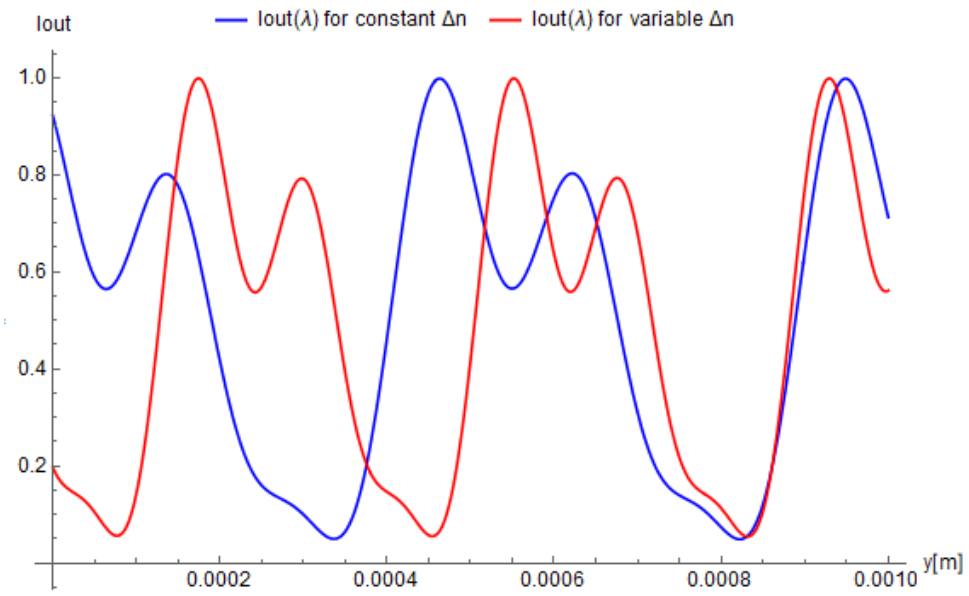


Figure 7.12: The same comparison as in the Figure 7.11, but as a function of y, for $y \in [0 - 0.001]m$

7.4 Conclusion

This short study of the variation of the birefringence with the wavelength brought several results that are extremely important for a laboratory implementation of the design.

- A good candidate for the material of the polarimeter is MgF_2, with a steady transmittance between 0.12 - 7μm;
- Nevertheless, the variation of the birefringence with the wavelength shows that MgF_2 is suitable for $\lambda \geq 0.125\mu m$;
- The birefringence of the material can be simulated through a polynomial fitting or a Sellmeier method;
- Based on a residual analysis we may conclude that the Sellmeier method follows better the experimental data below $\lambda = 4\mu m$;
- depending on the apex angle and the desired range of the wavelength, the minimal height of the wedge is comprised between 0.2mm and 9.3mm;
- Because 0.125 - 7μm wavelength range is too long for any detector, a shorter interval may be chosen. Based on practical and theoretical arguments one can propose 0.125 -1.4μm. In this case, for the spectral range $[0.125 - 0.14]\mu m$ the birefringence will be simulated with the help of the first polynomial from (7.1.1), while between $[0.14 - 1.4]\mu m$ the Sellmeier method will be preferred.
- The effect of the variation of the birefringence with the wavelength is quite negligible and is translated in a small variation of the periodicity of the signal.

Chapter 8

Uncertainty calculus and error treatment

A spectropolarimeter is a very sensitive instrument. Even though the model presented here overcomes the problems related to the presence of rotating components, it is still subject to errors and imprecision.

One part of this errors is generated by the very nature of the device. At the base, this instrument counts photons, and consequently is exposed to the specific errors of counting processes.

Meanwhile, we saw that in order to infer upon the value of the Stokes parameters we have to make a fitting of the experimental values with the theoretical expression of the intensity. This fitting is aimed to provide us the I, Q, U and V parameters, the functions m, n and p being known in the case of the normal incidence. Or because the measured intensity can exhibit a certain deviation with regard to the theoretical value, we may apply the theory of the linear adjustment. This will help us to estimate the error characterizing each parameter as a function of the number of photons in the context in which we do not know the true values of the Stokes parameters.

During the first part of our analysis from this chapter, which follows closely a development presented by Sparks[40], we will assume that the received signal deviates from the expected value with an amount ϵ, that has a normal distribution of mean zero and variance σ. Physically, ϵ is nothing else but the noise affection the signal. Therefore, a series of simulations will be conducted in the second part of this chapter, regarding the behavior of the device at different values of the noise.

8.1 Theoretical error estimation

Generally speaking we saw that the outgoing intensity of the optical system can be expressed as:

$$I_{out}(\lambda, y) = \alpha(\lambda, y)I + m(\lambda, y)Q + n(\lambda, y)U + p(\lambda, y)U \tag{8.1.1}$$

where $\alpha(\lambda, y), m(\lambda, y), n(\lambda, y), p(\lambda, y)$ are the modulation functions derived through the Mueller calculus. These functions are depending on the phase difference induced by the system and on the position of the analyzer axis. Because in our case the phase varies with the wavelength and with y, the position on vertical, we kept only this variables as main variables of the system. Meanwhile, the angle θ of the analyzer remains implied. It is a "secondary" type of variable, because is constant during the measurement of I_{out}. Just like the apex angle, ξ, this angle influences the overall functioning of the device.

The last relation can be expressed also as:

$$I_{out}(\lambda, y) = \begin{bmatrix} \alpha(\lambda, y) & m(\lambda, y) & n(\lambda, y) & p(\lambda, y) \end{bmatrix} \cdot \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \tag{8.1.2}$$

or simply stated

$$I_{out}(\lambda, y) = \Lambda \cdot S$$

where

$$\Lambda = \begin{bmatrix} \alpha(y, \lambda) & m(y, \lambda) & n(y, \lambda) & p(y, \lambda) \end{bmatrix}, \qquad S = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$$

and S is the Stokes vector to be determined. In our case, $\alpha(\lambda, y)$ it was simply given by $\frac{1}{2}$, while the m, n, p functions represented the modulations endured by each of the Stokes parameters. Now, because there is no perfect detector, able to register precisely the corresponding mathematical value of I_{out}, the intensity will be always measured with a certain fluctuation ϵ which can take either a positive or negative value, varying from one point of reading to another:

$$I_d(\lambda, y) = I_{out}(\lambda, y) + \epsilon$$

where I_d is the detected intensity.

The situation is depicted in the Figure 8.1 where we considered that the signal corresponding to an arbitrary incoming Stokes vector $S_{in} = [1, 0.23, 0.34, 0.15]$ is altered by a Gaussian noise with a variance of 0.06, also arbitrary chosen.

For simplification, we will consider that ϵ_i follows a normal distribution of mean 0 and variance σ_i. The index i suggest that we are working with a set of values read all along y. As a consequence, $I_d(\lambda, y_i) - I_{out}(\lambda, y_i) = \epsilon_i$ will follow also a normal distribution. Thus, the sum

$$\sum_{i=1}^{n} (\frac{\epsilon_i}{\sigma_i})^2$$

will follow necessarily a χ^2 distribution.

Based on this hypothesis we may apply now a linear adjustment by least square method. The chi-squared distribution associated to this error, for a given lambda will be:

$$\chi^2 = \sum_{i=1}^{n} \frac{(I_d(\lambda, y_i) - I_{out}(\lambda, y_i))^2}{\sigma_i^2}$$

$$= \sum_{i=1}^{n} \frac{(I_d(\lambda, y_i) - \alpha(\lambda, y_i)I - m(\lambda, y_i)Q - n(\lambda, y_i)U - p(\lambda, y_i)U)^2}{\sigma_i^2}$$

the summation taking place over the entire range of sampling points along y, because it is only there that we can measure I_d. Because we want to reach a maximum of likelihood between the theoretical and the fitted model, the derivatives of the chi-squared distribution with regard to Stokes parameters of the incoming light should be set to zero (for more details about this method see Magain [28]):

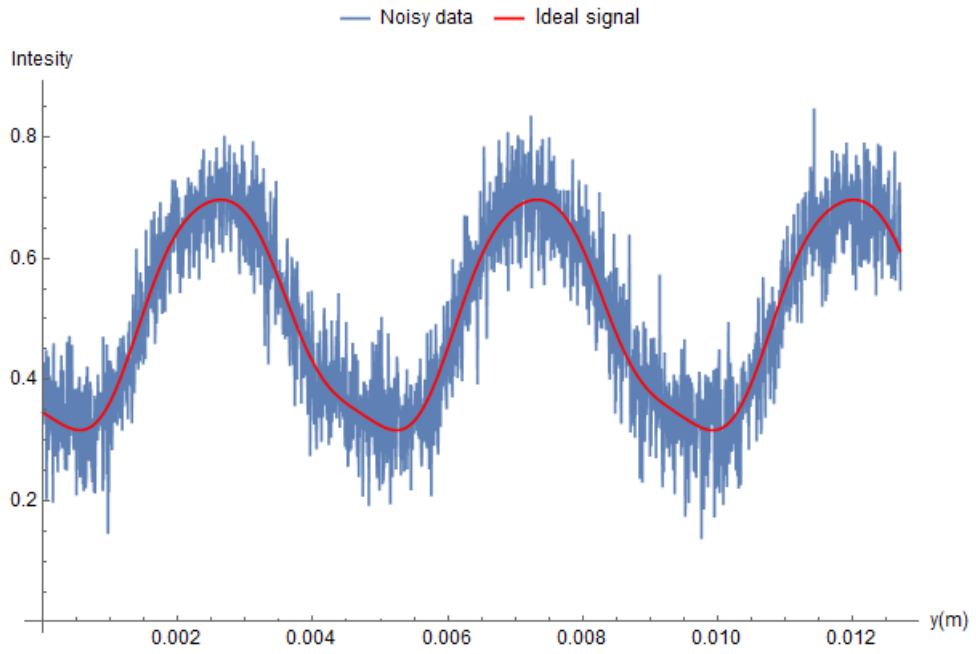


Figure 8.1: Noisy data and perfect signal (in the absence of noise), corresponding to I_{out}, for $\lambda = 0.14\mu m$. Here again we use arbitrary units for the intensity

$$\begin{cases} \frac{\partial\chi^2}{\partial I} = -2\sum_{i=1}^{n}\alpha(\lambda,y_i)\cdot\frac{(I_d(\lambda,y_i)-\alpha(\lambda,y_i)I-m(\lambda,y_i)Q-n(\lambda,y_i)U-p(\lambda,y_i)U)}{\sigma_i^2} = 0 \\ \frac{\partial\chi^2}{\partial Q} = -2\sum_{i=1}^{n}m(\lambda,y_i)\cdot\frac{(I_d(\lambda,y_i)-\alpha(\lambda,y_i)I-m(\lambda,y_i)Q-n(\lambda,y_i)U-p(\lambda,y_i)U)}{\sigma_i^2} = 0 \\ \frac{\partial\chi^2}{\partial U} = -2\sum_{i=1}^{n}n(\lambda,y_i)\cdot\frac{(I_d(\lambda,y_i)-\alpha(\lambda,y_i)I-m(\lambda,y_i)Q-n(\lambda,y_i)U-p(\lambda,y_i)U)}{\sigma_i^2} = 0 \\ \frac{\partial\chi^2}{\partial V} = -2\sum_{i=1}^{n}p(\lambda,y_i)\cdot\frac{(I_d(\lambda,y_i)-\alpha(\lambda,y_i)I-m(\lambda,y_i)Q-n(\lambda,y_i)U-p(\lambda,y_i)U)}{\sigma_i^2} = 0 \end{cases}$$

And as a consequence, opening the parenthesis and rearranging the therms, we will have the next system of equations:

$$\begin{cases} \sum_{i=1}^{n}\frac{I_d(\lambda,y_i)\cdot\alpha(\lambda,y_i)}{\sigma_i^2} = \sum_{i=1}^{n}\frac{I\alpha(\lambda,y_i)^2}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Qm(\lambda,y_i)\alpha(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Un(\lambda,y_i)\alpha(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Vp(\lambda,y_i)\alpha(\lambda,y_i)}{\sigma_i^2} \\ \sum_{i=1}^{n}\frac{I_d(\lambda,y_i)\cdot m(\lambda,y_i)}{\sigma_i^2} = \sum_{i=1}^{n}\frac{I\alpha(\lambda,y_i)m(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Qm(\lambda,y_i)^2}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Un(\lambda,y_i)m(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Vp(\lambda,y_i)m(\lambda,y_i)}{\sigma_i^2} \\ \sum_{i=1}^{n}\frac{I_d(\lambda,y_i)\cdot n(\lambda,y_i)}{\sigma_i^2} = \sum_{i=1}^{n}\frac{I\alpha(\lambda,y_i)n(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Qm(\lambda,y_i)n(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Un(\lambda,y_i)^2}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Vp(\lambda,y_i)n(\lambda,y_i)}{\sigma_i^2} \\ \sum_{i=1}^{n}\frac{I_d(\lambda,y_i)\cdot p(\lambda,y_i)}{\sigma_i^2} = \sum_{i=1}^{n}\frac{I\alpha(\lambda,y_i)p(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Qm(\lambda,y_i)p(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Un(\lambda,y_i)p(\lambda,y_i)}{\sigma_i^2} + \sum_{i=1}^{n}\frac{Vp(\lambda,y_i)^2}{\sigma_i^2} \end{cases}$$

In order to simplify the notations, we will let aside (λ, y_i) parenthesis, but without forgetting that only I,Q,U,V are independent of the summation index. Passing then into a matrix form, we will have:

$$\begin{pmatrix} \sum_{i=1}^{n} \frac{I_{di}\cdot\alpha_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{I_{di}\cdot m_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{I_{di}\cdot n_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{I_{di}\cdot p_i}{\sigma_i^2} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \frac{\alpha_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i\alpha_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i\alpha_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i\alpha_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{\alpha_i m_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i m_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i m_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{\alpha_i n_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i n_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i n_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{\alpha_i p_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i p_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i p_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i^2}{\sigma_i^2} \end{pmatrix} \cdot \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \tag{8.1.3}$$

This is a general result characterizing any spectropolarimeter able to modulate the incoming signal through functions of the type of α, m, n and p. Nevertheless, because we follow the development of Sparks[40], we will adopt also its notation at this level. Thus, we will consider that :

$$S_I = \begin{pmatrix} \sum_{i=1}^{n} \frac{I_{di}\cdot\alpha_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{I_{di}\cdot m_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{I_{di}\cdot n_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{I_{di}\cdot p_i}{\sigma_i^2} \end{pmatrix}$$

S_I representing the distribution of the detected intensity, as a function of α, m, n, p. Meanwhile,

$$\mathbf{B} = \begin{pmatrix} \sum_{i=1}^{n} \frac{\alpha_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i\alpha_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i\alpha_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i\alpha_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{\alpha_i m_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i m_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i m_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{\alpha_i n_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i n_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i n_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{\alpha_i p_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{m_i p_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{n_i p_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{p_i^2}{\sigma_i^2} \end{pmatrix}$$

is the covariance (or curvature matrix), and

$$S = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}$$

is the Stokes vector of the incoming light. Based of this notations, the system (8.1.3) can be expressed as:

$$s_I = \mathbf{B} \cdot S$$

Assuming that **B** is invertible, we have then:

$$S = \mathbf{B}^{-1} \cdot s_I$$

the diagonal elements of the $\mathbf{B}^{-1}$ matrix being the variances associated to the Stokes parameters[28]. We have to remember now that in our case α, m, n, p are given by:

$$\begin{cases} \alpha = \frac{1}{2} \\ m = \frac{1}{2}\cos(2\theta)\cos\rho \\ n = \frac{1}{2}\cdot(\cos\phi\sin(2\theta) - \cos(2\theta)\sin\rho\sin\phi) \\ p = \frac{1}{2}\cdot(\cos\phi\cos(2\theta)\sin\rho + \sin(2\theta)\sin\phi) \end{cases}$$

In order to simplify the computation of the **B** matrix, we may chose a particular case with $\theta = \frac{\pi}{4}$, even if this corresponds to a scenario in which Q is undetermined. This value will drastically simplify the expressions of m, n and p functions.

Thus, the previous relations will become:

$$\begin{cases} \alpha = \frac{1}{2} \\ m = 0 \\ n = \frac{1}{2}\cos\phi \\ p = \frac{1}{2}\sin\phi \end{cases}$$

It should be stressed here that an analyzer of $\theta = \frac{\pi}{4}$ will cancel the effect introduced by the last wedge of apex 2ξ upon the Q parameter, because the m function is reduced to zero.

For simplification we may consider also that the summation covers an entire number of periods for n and p. Meanwhile, we may suppose that the variance of the distribution of ϵ does not varies from one point to another, but is constant and specific to a Poissonian process of counting:

$$\sigma = \sqrt{\frac{N}{n}}$$

where N is the total number of incoming photons and n the number of sampling points. Based on this assumptions, the first line of the curvature matrix will be:

$$\begin{cases} \sum_{i=1}^{n} \frac{\alpha_i^2}{\sigma_i^2} = \sum_{i=1}^{n} \frac{(\frac{1}{2})^2}{\sigma^2} = \frac{1}{4\sigma^2} n \\ \sum_{i=1}^{n} \frac{m_i\alpha_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{0\frac{1}{2}}{\sigma} = 0 \\ \sum_{i=1}^{n} \frac{n_i\alpha_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{\frac{1}{2}\cos\phi\frac{1}{2}}{\sigma^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\cos\phi \\ \sum_{i=1}^{n} \frac{p_i\alpha_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{\frac{1}{2}\sin\phi\frac{1}{2}}{\sigma^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\sin\phi \end{cases}$$

On the same way, for the next lines we have:

$$\begin{cases} \sum_{i=1}^{n} \frac{\alpha_i m_i}{\sigma_i^2} = 0 \\ \sum_{i=1}^{n} \frac{m_i^2}{\sigma_i^2} = 0 \\ \sum_{i=1}^{n} \frac{n_i m_i}{\sigma_i^2} = 0 \\ \sum_{i=1}^{n} \frac{p_i m_i}{\sigma_i^2} = 0 \end{cases} \quad \begin{cases} \sum_{i=1}^{n} \frac{\alpha_i n_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{\frac{1}{2}\cos\phi\frac{1}{2}}{\sigma^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\cos\phi \\ \sum_{i=1}^{n} \frac{m_i n_i}{\sigma_i^2} = 0 \\ \sum_{i=1}^{n} \frac{n_i^2}{\sigma_i^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\cos^2\phi \\ \sum_{i=1}^{n} \frac{p_i n_i}{\sigma_i^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\cos\phi\sin\phi \end{cases} \quad \begin{cases} \sum_{i=1}^{n} \frac{\alpha_i p_i}{\sigma_i^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\sin\phi \\ \sum_{i=1}^{n} \frac{m_i p_i}{\sigma_i^2} = 0 \\ \sum_{i=1}^{n} \frac{n_i p_i}{\sigma_i^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\cos\phi\sin\phi \\ \sum_{i=1}^{n} \frac{p_i^2}{\sigma_i^2} = \frac{1}{4\sigma^2}\sum_{i=1}^{n}\sin^2\phi \end{cases}$$

and consequently, the covariance matrix will be:

$$\mathbf{B} = \frac{1}{4\sigma^2} \cdot \begin{pmatrix} n & 0 & \sum_{i=1}^{n} \cos\phi & \sum_{i=1}^{n} \sin\phi \\ 0 & 0 & 0 & 0 \\ \sum_{i=1}^{n} \cos\phi & 0 & \sum_{i=1}^{n} \cos^2\phi & \sum_{i=1}^{n} \cos\phi \sin\phi \\ \sum_{i=1}^{n} \sin\phi & 0 & \sum_{i=1}^{n} \cos\phi \sin\phi & \sum_{i=1}^{n} \sin^2\phi \end{pmatrix}$$

Omitting the Q component from the system we may also write:

$$\frac{1}{2\sigma^2} \begin{pmatrix} \sum_{i=1}^{n} I_{di} \\ \sum_{i=1}^{n} I_{di} \cos\phi \\ \sum_{i=1}^{n} I_{di} \sin\phi \end{pmatrix} = \frac{1}{4\sigma^2} \cdot \begin{pmatrix} n & \sum_{i=1}^{n} \cos\phi & \sum_{i=1}^{n} \sin\phi \\ \sum_{i=1}^{n} \cos\phi & \sum_{i=1}^{n} \cos^2\phi & \sum_{i=1}^{n} \cos\phi \sin\phi \\ \sum_{i=1}^{n} \sin\phi & \sum_{i=1}^{n} \cos\phi \sin\phi & \sum_{i=1}^{n} \sin^2\phi \end{pmatrix} \cdot \begin{pmatrix} I \\ U \\ V \end{pmatrix} \tag{8.1.4}$$

Because the summation covers a large amount of data (each pixel or each sampling point), then according the Riemann definite integral expression, we have:

$$\int_a^b f(x)dx = \lim_{\|\Delta x_i\| \to 0} \Delta x_i \cdot \sum_{i=0}^{n} f(x_i)$$

If we consider that the readings of I_d take place on a total height of h, over a total number of points n covering an integer number of periods, then we may write:

$$\int_0^h \cos\phi(y)dy \approx \frac{h}{n} \sum_{i=1}^{n} \cos\phi$$

and so:

$$\sum_{i=1}^{n} \cos\phi \approx \frac{n}{h} \int_0^h \cos\phi(y)dy$$

Meanwhile, using the fact that

$$\int_0^h \cos^2\phi d\phi = \frac{1}{2}\Big[\frac{1}{2}\sin(2\phi) + \phi\Big]_0^h$$

we infer that

$$\sum_{i=1}^{n} \cos^2\phi = \sum_{i=1}^{n} \sin^2\phi = \frac{n}{2}$$

With these results, the previous system (8.1.4) becomes:

$$\frac{1}{2\sigma^2}\begin{pmatrix}\sum\limits_{i=1}^{n} I_{di}\\ \sum\limits_{i=1}^{n} I_{di}\cos\phi\\ \sum\limits_{i=1}^{n} I_{di}\sin\phi\end{pmatrix} = \frac{1}{4\sigma^2}\cdot\begin{pmatrix}n & 0 & 0\\ 0 & \frac{n}{2} & 0\\ 0 & 0 & \frac{n}{2}\end{pmatrix}\cdot\begin{pmatrix}I\\ U\\ V\end{pmatrix} \tag{8.1.5}$$

And then

$$\begin{pmatrix}I\\ U\\ V\end{pmatrix} = \frac{2}{n}\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}\cdot\begin{pmatrix}\sum\limits_{i=1}^{n} I_{di}\\ \sum\limits_{i=1}^{n} I_{di}\cos\phi\\ \sum\limits_{i=1}^{n} I_{di}\sin\phi\end{pmatrix}$$

Directly, we have:

$$\begin{cases} I = \frac{2}{n}\cdot\sum\limits_{i=1}^{n} I_{di} = \frac{2N}{n}\\ U = \frac{4}{n}\cdot\sum\limits_{i=1}^{n} I_{di}\cos\phi\\ V = \frac{4}{n}\cdot\sum\limits_{i=1}^{n} I_{di}\sin\phi\end{cases}$$

Here we used the idea that because we are in a Poisson statistic in which the number of photons counts, then the sum of intensities can be replaced by the total number of received photons, N. Because $\mathbf{B}^{-1}$ is a correlation matrix between the Stokes parameters and the received intensities, we may have access then to the variance and the standard deviation for each of the Stokes parameters:

$$\begin{cases}\sigma^2(I) = 4\sigma^2\cdot\frac{1}{n} = \frac{4\frac{N}{n}}{n} = \frac{4N}{n^2}\\ \sigma^2(U) = \frac{8N}{n^2}\\ \sigma^2(V) = \frac{8N}{n^2}\end{cases} \tag{8.1.6}$$

Coming back to the system (8.1.5), for simplification we may replace the vector from the left side with $S = [S1, S2, S3]$, where

$$\begin{pmatrix}\sum\limits_{i=1}^{n} I_{di}\\ \sum\limits_{i=1}^{n} I_{di}\cos\phi\\ \sum\limits_{i=1}^{n} I_{di}\sin\phi\end{pmatrix} = \begin{pmatrix}S1\\ S2\\ S3\end{pmatrix}$$

Consequently, the system (8.1.5) becomes:

$$\begin{pmatrix}S1\\ S2\\ S3\end{pmatrix} = \frac{1}{2}\cdot\begin{pmatrix}n & 0 & 0\\ 0 & \frac{n}{2} & 0\\ 0 & 0 & \frac{n}{2}\end{pmatrix}\cdot\begin{pmatrix}I\\ U\\ V\end{pmatrix} \tag{8.1.7}$$

In order to find the normalized parameters u and v we divide the last system by I, where $I = (2/n)S_1$:

$$\begin{cases} S_2 = \frac{S_1}{2} \cdot u \\ S_3 = \frac{S_1}{2} \cdot v \end{cases} \tag{8.1.8}$$

Therefore:

$$\begin{cases} \sigma^2(u) = \frac{2}{S_1} \\ \sigma^2(v) = \frac{2}{S_1} \end{cases}$$

Or the therm S_1 is nothing else than the total received energy, thus, N, the number of photons:

$$\begin{cases} \sigma^2(u) = \frac{2}{N} \\ \sigma^2(v) = \frac{2}{N} \end{cases} \Rightarrow \begin{cases} \sigma(u) = \sqrt{\frac{2}{N}} \\ \sigma(v) = \sqrt{\frac{2}{N}} \end{cases} \tag{8.1.9}$$

From the system (8.1.8) we may infer also immediately the standard deviation for the degree of polarization :

$$\sigma(p = \sqrt{u^2 + v^2}) = \sqrt{\frac{2}{N}} \tag{8.1.10}$$

It must be stressed here that the results from (8.1.9) and (8.1.10) are true only for the cases considered. In a general way, if we keep only the assumption that the covariance matrix is invertible, all what we can say is that the standard deviation for the normalized parameters is proportional to $\frac{1}{\sqrt{N}}$.

In this approach of the errors from the system we considered a compressed scenario, in which all the sources of errors are in a certain way related only to the counting of photons, without any question about the source of these photons.

About ϵ, the variation regarding the expected value of the signal, we considered that is a fluctuation of mean 0 and variance σ.

Nevertheless, ϵ encapsulates the noise affecting the system and can be related to a very important property characterizing the detector, which is the signal to noise ratio (SNR):

$$SNR = \frac{S}{N} \approx \frac{\sigma^2_{signal}}{\sigma^2_{noise}} \tag{8.1.11}$$

where S is the signal and N is the noise, while in the approximate expression σ_{signal} is the amplitude of the signal in the case of a sinusoidal modulation. Generally, when we calculate the noise, we have to take care of the dark current noise (N_{DC}), the signal noise (N_S) and read-out noise (N_{RO}):

$$N = \sqrt{N_S^2 + N_{DC}^2 + N_{RO}^2} \tag{8.1.12}$$

The signal noise (N_S) is a statistical noise and is given by:

$$N_S = \sqrt{S} \tag{8.1.13}$$

The dark current noise (N_{DC}) which represents the detected signal when no source is observed, and which is originated in the electronics of the detector, is given by:

$$N_{DC} = \frac{I_{DC}}{e} t_i \tag{8.1.14}$$

where I_{DC} is the dark current, e is the charge of the electron and t_i is the integration time. N_{RO}, the read-out noise is the noise associated to the transfer of the charge from the detector, or to the conversion into a digital signal.

It is worthy to be mentioned here that the intensity "perceived" by the detector should be expressed like:

$$I_{out}(\lambda) = t(\lambda) \cdot I_0(\lambda) \tag{8.1.15}$$

where $I_0(\lambda)$ is the theoretical received intensity, and $t(\lambda)$ stays for the transmission of the instrument. Thus, any real implementation of this design should consider this last formula in order to properly infer upon the true value of received intensity.

8.2 Numeric simulation

During the first part of this chapter we considered that the true value of the intensity I_{out} is an unknown and that we have access only to a "distorted" value I_d which varies in every point of detection, i, with an amount ϵ relative to the ideal value I_{out}. And we supposed also that this ϵ follows a normal distribution of constant variance during the integration time of the detector.

Another strategy, even more helpful, that can show us the true behavior of the optical device, is to consider that I_{out} is known, together with the Stokes vector of the incoming light, S_{in}. Adding then different values of noise ϵ we can generate a number of points n representing the detected signal, I_d. Fitting this points with a curve given by (8.1.1) we will try to infer the values of $S_{in} = [I, Q, U, V]$. And of course, because in this case we know which the true values are, we may compute all the time the relative error corresponding to each parameter, the impact of the level of noise over this error and the impact of the other parameters of the system over the relative error.

8.2.1 Impact of noise

The level of noise has a huge impact over the quality and the capacity to extract I, Q, U and V parameters of the incoming light, and the numeric analysis suggest that this impact is not homogeneous and also is not equally distributed among the four parameters. In fact, each configuration of the optical device (certain value of θ, the analyzer angle, and ξ, the apex angle) will affect more or less the capacity to extract the Stokes parameters.

In order to properly represent the consequences of the noise we may consider the ideal case of a normal incident beam having an elliptical state of polarization described by[1]:

$$S_{in} = \begin{bmatrix} 1 \\ 0.55 \\ 0.703 \\ 0.45 \end{bmatrix}$$

More than that we will consider in the beginning that the analyzer is optimized for linear (Q,U) and circular polarization (V), all-together. According to the equation (6.3.4), for a very narrow

[1] This is only an arbitrary choice, characteristic for a total polarization.

wavelength range 0.12 - 0.125 μm we will find that the most suitable values of the analyzer angle are:

- 24.06°
- -5.7°
- -12.6°
- -22.34°

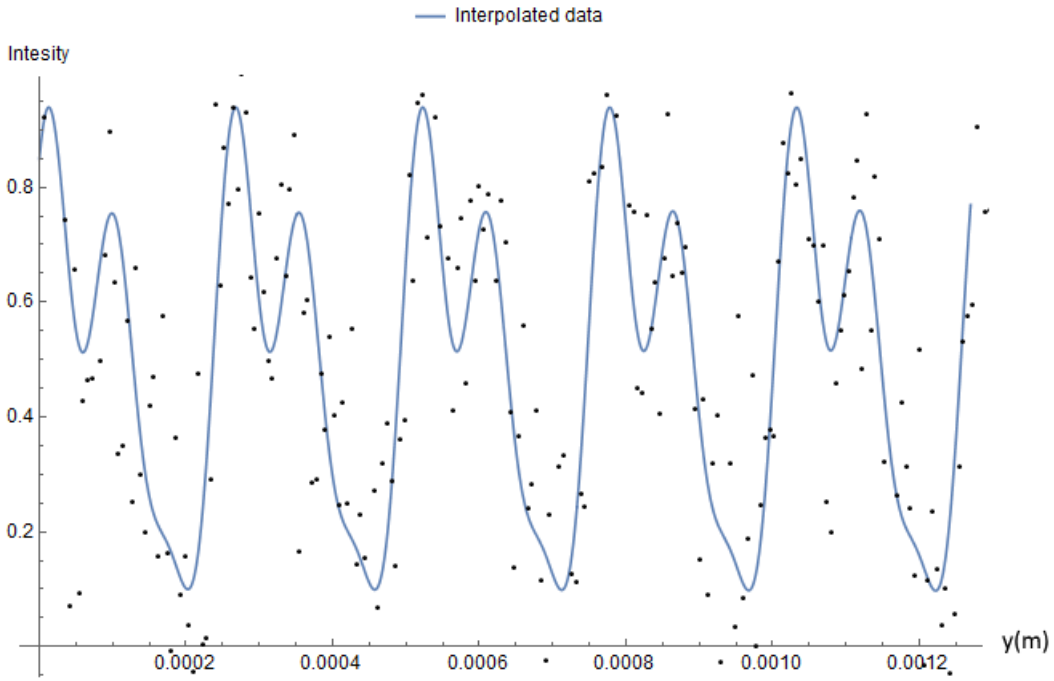


Figure 8.2: Simulated data for $\sigma = 0.2$ on a $1.2mm$ range, with a step of $6\mu m$ and for $\lambda = 0.125\mu m$

Considering that this values are symptomatic for the entire spectrum $0.125 - 1.4\mu m$, we may fix θ, the analyzer angle at, let's say, 24.06°. Meanwhile, because at $\lambda = 0.125\mu m$ the minimum sampling distance is of about 0.51 mm, we may chose a wedge of height 1.27cm.

Thus, after ten simulations of the extraction of Stokes parameters of an incoming beam characterized by the vector S_{in} we will find that the relative error $(I_{measured} - I)/I\%, (Q_{measured} - Q)/Q\%, etc.$ will vary like in the Figure 8.3. The graphic suggest a relatively slow increase of the error for Q and V with the level of noise and an almost unacceptable situation for U which is undetermined in this configuration.

Only for proving the very important role played by the angle of the analyzer we may consider a second scenario: this time we will chose $\theta = 18.33^{\circ}$, corresponding to an optimization for elliptic and circular polarizations. The result is plotted in the Figure 8.4. Again, the U parameter stays above the limit of a reasonable error.

If we change again the angle θ to a value of 139°, which is an arbitrary point on the curve of $\theta(y)$ for $\lambda = 0.125\mu m$, curve corresponding to the total polarization described by the vector S_{in}, then the fluctuation of the relative error with σ will have the allure from the Figure 8.5.

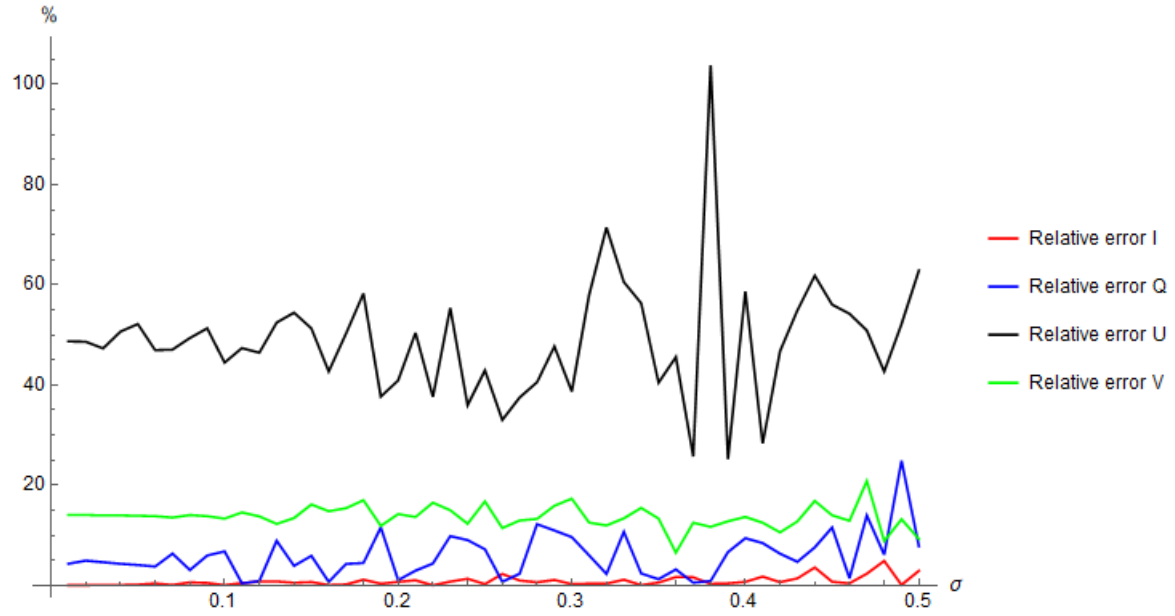


Figure 8.3: Relative error variation for the fitted Stokes parameters with regard to the level of noise, for $\lambda = 0.12\mu m$ and a system optimized for the linear and circular polarization.

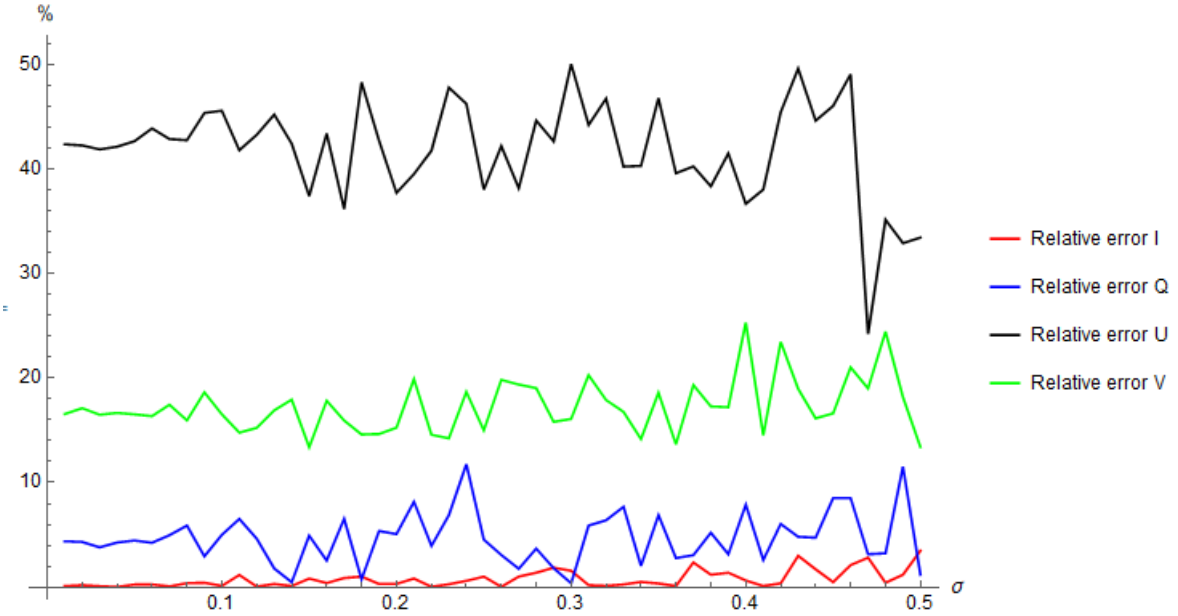


Figure 8.4: Relative error variation for the fitted Stokes parameters with regard to the level of noise, for $\lambda = 0.12\mu m$ and a system optimized for the elliptical and circular polarizations.

This evolution suggests that the impact of noise upon the extraction of the Stokes parameters is not homogeneous. Changing the angle θ will give us the opportunity to optimize the extraction of a parameter or another.

As an example, we may take a look to the Figure 8.6, where the relative error for each parameter was plotted as a function of θ, under the assumption of a normal incidence with the Stokes vector S_{in}.

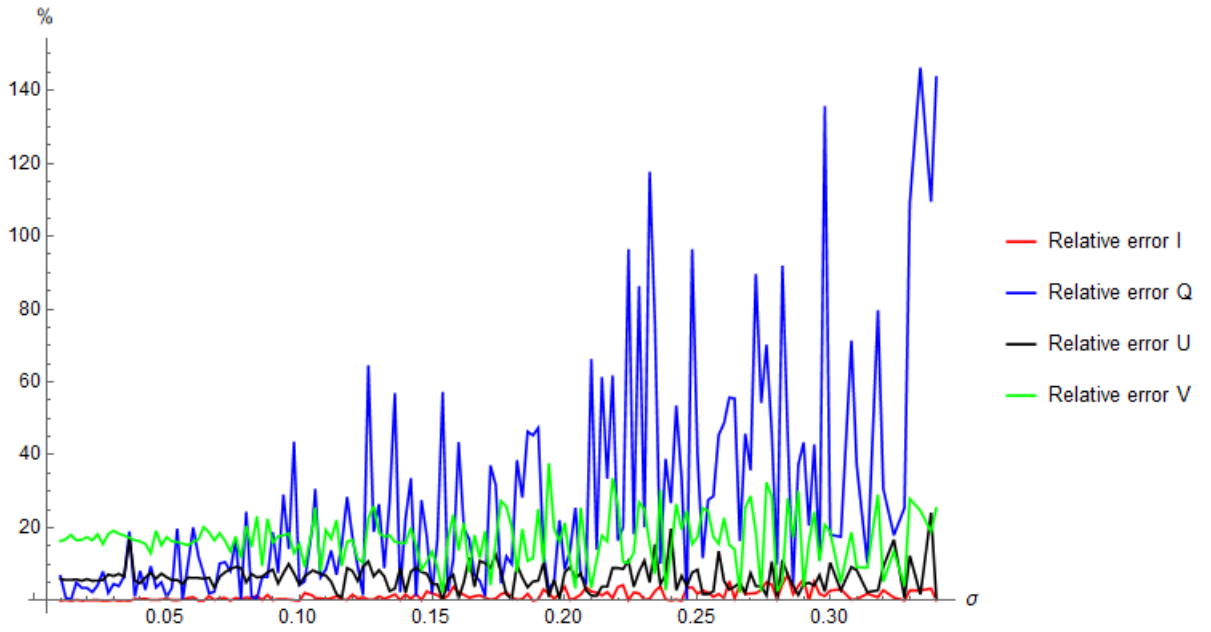


Figure 8.5: Relative error variation for $\lambda = 0.12\mu m$ and a system optimized for the elliptical polarization given by S_{in}.

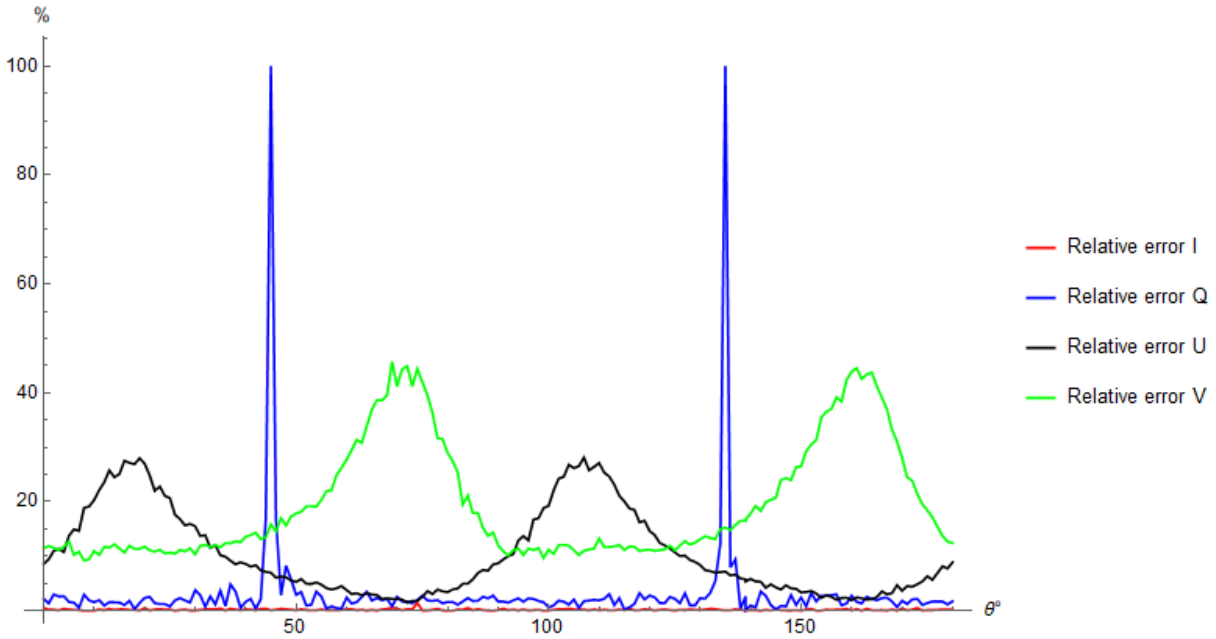


Figure 8.6: Error variation for the fitted Stokes parameters with regard to the analyzer angle, θ.

Nevertheless, a detailed study about the Stokes parameters and the analyzer angle, θ, is beyond the purpose of this thesis.

Based on the limitations imposed by the noise we may also simulate a minimal height of the wedges. Thus, during the previous chapter we saw that the total sampling distance corresponding to $\lambda = 0.125\mu m$, was of 0.51 mm, the minimum distance between two points being 0.04 mm. Now, if we consider the hypothetical case of $\sigma = 0.06$ and a maximal relative error of 50%, we will obtain

the next situation:

- readings at each pixel (pas=$6\mu m$) $\Rightarrow$ minimum height = 0.51 mm;
- at 2 pixels $\Rightarrow$ 2 mm;
- at 3 pixels $\Rightarrow$ 10 cm

It is an almost exponential evolution of the needed total sampling distance if we increase the distance between the readings. This behavior is imposed only be the noise.

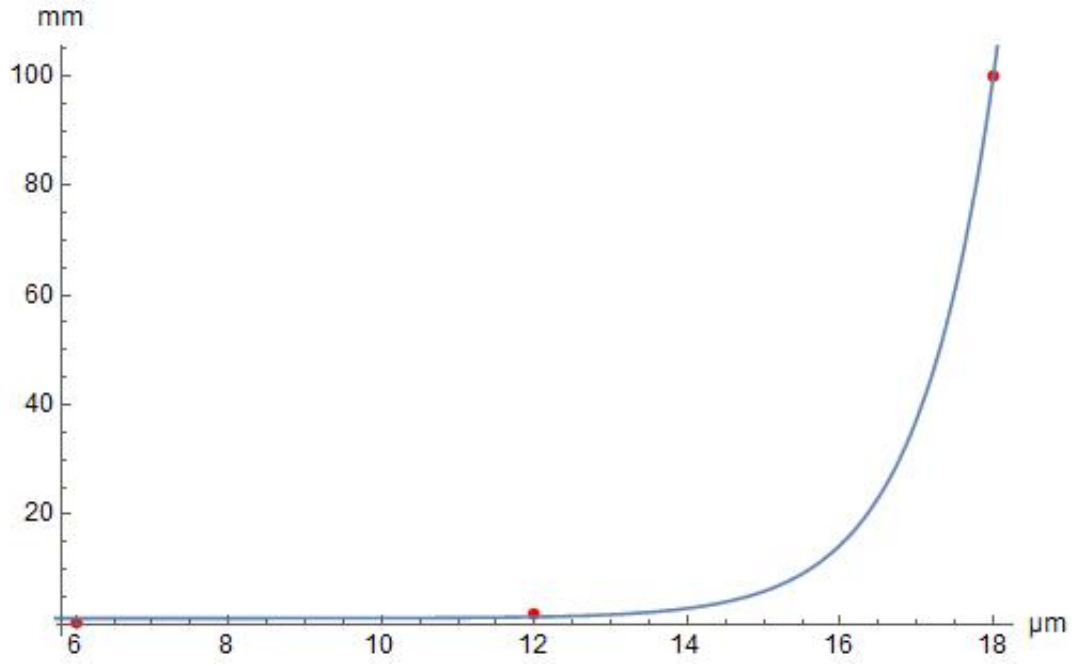


Figure 8.7: More the readings are rare, more the need of a higher prism increases. The relation is almost exponential: horizontal - sampling distance; vertical - height of the prism

In the same way that the variation of the sampling distance affects the needed total sampling distance, the value of the later will affect the level of the error. Shorter will be the total sampling distance, higher will be error and vice versa.

Of paramount importance for the functioning of this spectropolarimeter is also the eventual correlation between the error and the wavelength. A series of simulations were conducted to check this correlation for the same dimensions of the instrument as before and again under the assumption of normal incidence of a ray elliptically polarized (S_{in}). The results for $\sigma = 0.04$ (corresponding to an SNR of about 100[2]) are plotted in the Figure 8.8.

We may notice a strong dependency of V and U parameters over the wavelength, while for I and Q the plot is inconclusive. In order to see if I and Q are also related to λ an hypothesis test should be conducted. For this test, a series of 28 simulations (a number that was arbitrary chosen) at a steady $\lambda = 0.14\mu m$ were realized. In addition, a simulation for n=72 values (again, arbitrary chosen) of λ between $0.125 - 1.4\mu m$ was also computed. The entire statistical procedure is represented in the Figure 8.9.

For the V parameter, the average value during this last simulation was:

[2]During these simulations we chose values of σ corresponding to levels of SNR currently encountered in practice.

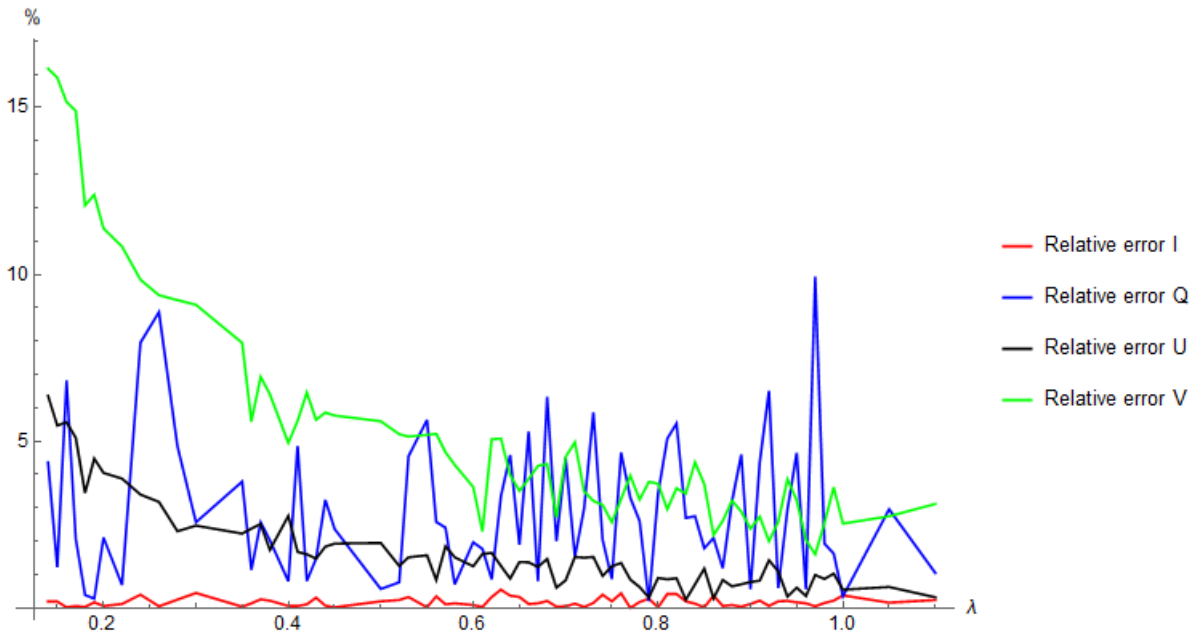


Figure 8.8: The spectropolarimeter exhibits a stronger sensibility with regard to the noise at lower wavelengths. The simulations were conducted for $\sigma = 0.04$

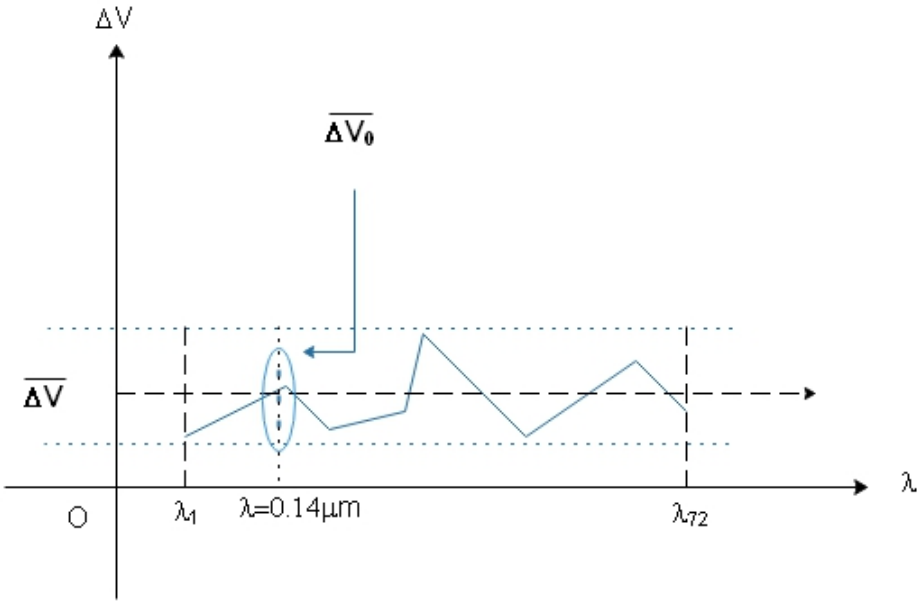


Figure 8.9: The statistical test for ΔV (intuitive representation): $\overline{\Delta V}$ is the mean value of ΔV for 72 values of λ, while $\overline{\Delta V_0}$ is the mean value of ΔV over 28 readings at $\lambda = 0.14\mu m$

$$\overline{\Delta V} = \frac{1}{n}\sum_{i=1}^{n}(\Delta V_i) = 7.53$$

while the variance is estimated by:

$$s_n^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(\Delta V_i - \overline{\Delta V})^2 = 23.16$$

The statistic of the test for a risk $\alpha = 0.05$ (considered as a common value for this type of test) is given then by:

$$z = \sqrt{n}\frac{\overline{\Delta V} - \overline{\Delta V_0}}{s_n^*} = -10.34$$

where $\overline{\Delta V_0}$ is the average value of ΔV for constant wavelength, in this case, 16.93.

Or, because

$$|z| = 10.34 > 1 - \frac{\alpha}{2} = 0.97$$

it results that the null hypothesis should be rejected, and so we can say that the available data suggest a dependency of the relative error of V over the wavelength. The test confirms thus our observations.

The same hypothesis test conducted for I parameter will give a $|z| = 0.05 < 0.97$, which can be translated into an independence of the error with regard to the wavelength.

This result is obvious if we remember that I is modulated by a constant in the final expression of the outgoing intensity and has no dependency about the wavelength.

Meanwhile, the statistic for Q will be 3.43, and for U, 10.49, which proves again that the error on Q and U parameters is correlated with λ.

The same result about the correlation was obtained also for $\sigma = 0.06$ or $\sigma = 0.08$, proving that this behavior characterizes any ratio between the signal and the noise, the only difference being that for lower SNR the trend of the variation becomes visible mostly for V and U. Moreover, there is no reason to suppose that this result is dependent over the type of polarization.

Using the same type of simulation as before, with $\sigma = 0.03$, we can analyze also the variation of the relative error on the Stokes parameters with the apex angle, ξ (Figure 8.10).

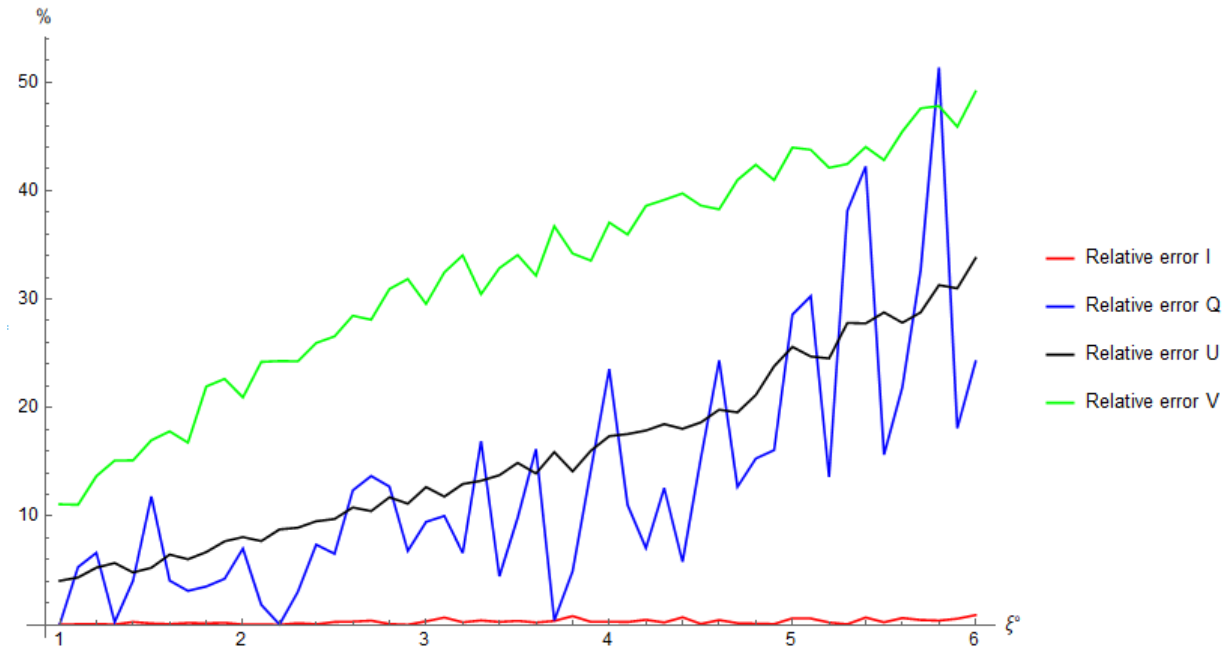


Figure 8.10: Variation of the relative error on I,Q,U and V with regard to the angle of the wedge, ξ

Considering an inferior limit of this angle of about 1°, dictated by the fabrication capabilities, and $\lambda = 0.14\mu m$, we will retrieve the behavior exposed in the Figure 8.10, suggesting a strong dependence over this angle. More than that, we see that at values of only about 6° we already reach an imprecision level of 50% for the V parameter, V, Q, and U being again subject of the strongest variation. And this fact gives us a supplementary reason to choose an apex angle as small as possible. Again, this result can be more precisely described if multiple simulations, for different types of incoming polarization are considered.

8.2.2 Uncertainty for partial polarization: the limits of the design

An extremely important aspect in the study of such a device aimed for an astronomical space use is also its ability to provide access to the Stokes parameters under the assumption of a low degree of polarization. And this because generally the values of the degree of polarization (p) encountered in astronomy are very small. Thus, we saw before that a planet like Venus exhibits a linear polarization below 11%, Mars is characterized also by a linear degree of polarization less than 3%, while Jupiter reaches a maximum of 7%. Overall, these values suggest that our spectropolarimeter should provide a very good access to rates between 0.01 and 0.1.

In addition, any attempt to use the spectropolarimeter for the study of exoplanets must consider a degree of polarization of about 10^{-5} or even 10^{-6}...

The biggest hindrance for reaching this precision is the noise. Several numerical simulations were conducted in order to point out the impact of noise over the minimum degree of linear polarization that can be retrieved, and the results can be seen in the figures below. At this point we considered that the V component is 0, and the degree of polarization was calculated using the relation (4.3.10):

$$p = \frac{\sqrt{Q^2 + U^2}}{I}$$

For the degree of circular polarization we will use the relation (4.3.11):

$$p = \frac{V}{I}$$

All these simulations show that in order to work below a degree of polarization of 0.1 at a wavelength of $0.14\mu m$, the noise should be contained below $\sigma = 0.008$ (SNR=625).

Concerning the circular polarization, which in astronomy reaches very small degrees (even 100 times smaller than the linear polarization), the device will require even a higher sensibility. Fortunately, the numerical simulations suggest that the detection of V parameter is not so exposed to the noise as the rest: at $\sigma = 0.01$ we may have a precision on V of more than 70% for a circular degree of polarization of only 0.001.

8.3 Instrumental polarization

A possible source of error for a space polarimeter is the polarization induced by the telescope itself. The best option is to use a Cassegrain structure, which, generally introduces a polarization of about 0.01%. Otherwise, any other type of telescope that will require tilted mirrors in front of the polarimeter will induce polarization even until 10%[26][p.55].

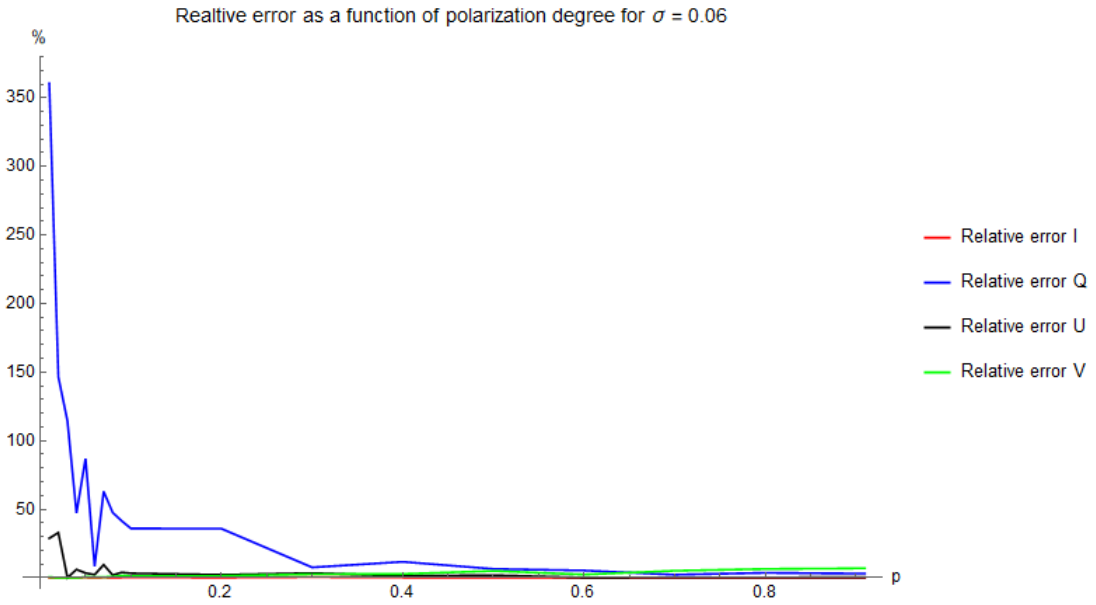


Figure 8.11: Error variation for the fitted Stokes parameters with regard to the polarization degree for $\sigma = 0.06$ and linear incoming polarization

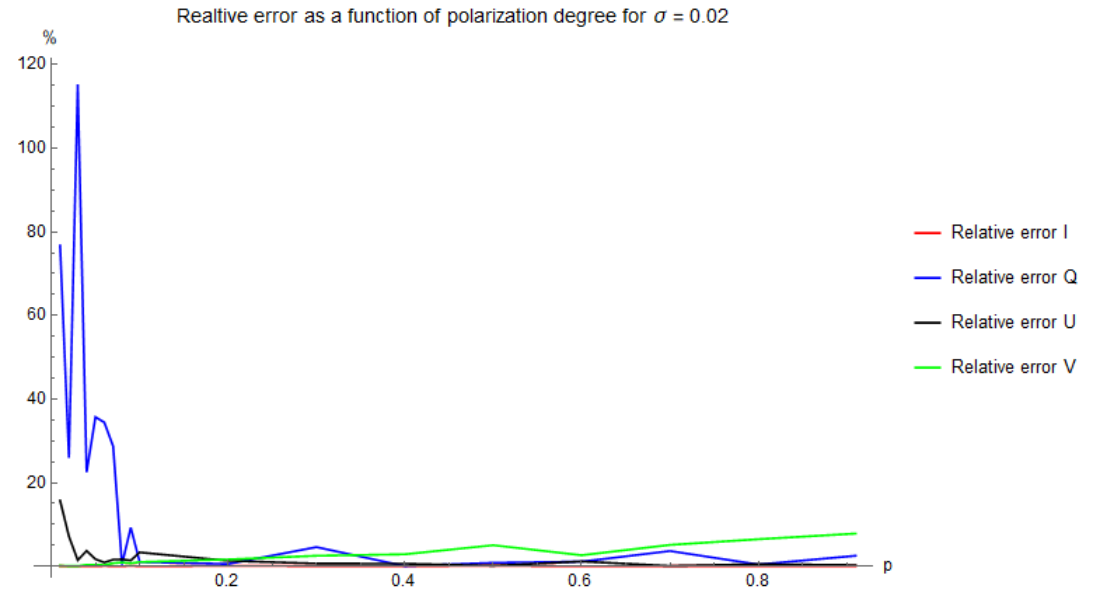


Figure 8.12: Error variation for the fitted Stokes parameters with regard to the polarization degree for $\sigma = 0.02$ and linear incoming polarization

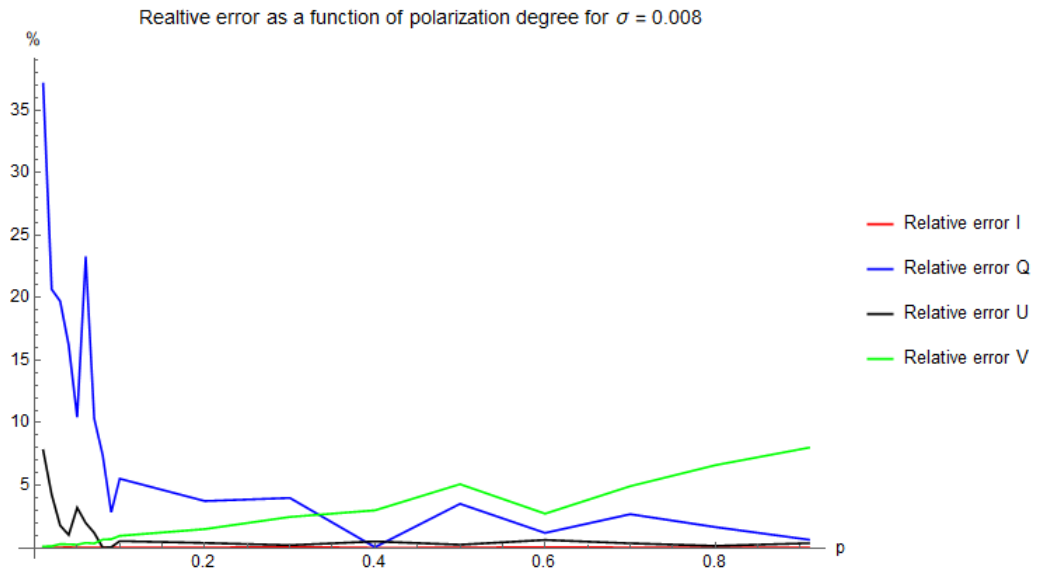


Figure 8.13: Error variation for the fitted Stokes parameters with regard to the polarization degree for $\sigma = 0.008$ and linear incoming polarization

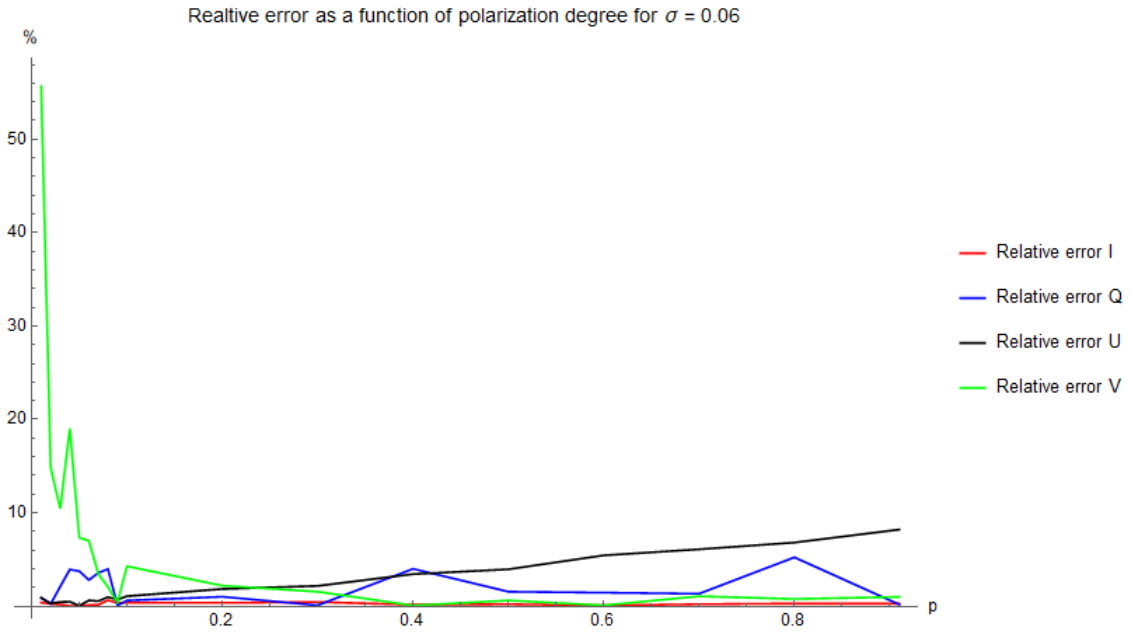


Figure 8.14: Error variation for the fitted Stokes parameters with regard to the polarization degree for $\sigma = 0.06$ and incoming circular polarization

8.3.1 Conclusions

- For a given polarization, the level of noise affects differently the extraction of the Stokes parameters.
- The capacity to infer upon the values of Stokes parameters depends also on the characteristics of the device: the analyzer angle (θ) and the apex angle (ξ).
- Concerning the analyzer angle, the simulations suggest that there is no value able to minimize simultaneously the errors for all Stokes terms. But the analyzer can be optimized for the extraction of at least two parameters.
- The relative error increases at low values of the wavelength.
- The relative error decreases at small values of the apex angle, ξ.
- A degree of linear polarization below 0.1 is achievable for SNR around 625.
- For the circular polarization, with an SNR=700 we may reach a degree of polarization of 10^{-3}.
- In order to reduce the polarization induced by the telescope, a Cassegrain structure is recommended.
- Further studies are needed in order to describe the impact of θ, ξ upon the capacity of the device. Also the tests for the achievable degree of polarizations should be carried on.

Chapter 9

Off-axis problem

Even if this kind of instrument is intended mostly for stars observations, with a very precise positioning, and parallel rays in the entry, the possibility to encounter off-axis rays should not be neglected. Can be determined by a misalignment of the polarimeter or, in the case of use of a collimator, by the chromatic aberrations.

Because of this, we have to understand what is happening with the modulation of the intensity in this case and if this situation determines a limitation of the system. The study of the off-axis scenario is even more important in the perspective to convert this instrument into an imaging spectropolarimeter, where the capability to observe extended bodies will depend on its performance in off-axis conditions.

9.1 Normal incidence revisited

Until now we used the hypothesis that the rays entering the system are normal to the incidence surface and consequently they travel without deviation through the birefringent bloc of MgF_2. For a better accuracy of the theoretical analysis this situation must be properly represented with the help of Huygens construction.

The ray tracing for the first two parts of the system was already explained in the Chapter 4 and is reproduced in the Figure 9.1.

The most important observation coming with this representation is the fact that in the second medium the light undergoes a splitting. The ordinary (o) and extraordinary (e) ray will have no longer the same path. Thus, the hypothesis that the normal incident ray will pass as a single beam trough the system is not valid anymore. For an undetermined incoming state of polarization there will be a splitting along y axis (see orientation of the axis in Figure 9.1).

For the third wedge the geometrical representation is a little bit more complicated (Figure 9.2). Because of the fact that the optical axis is tilted to 45° in the (xy) plane, each of the previous e and o-wave could be now splitted in new e and o-vibrations. If the incidence would be horizontal on the third surface, the o-wave would propagate without deviation while the e-wave would be deviated as a consequence of passage from a medium of index n_o to a medium of index n_e. The true trajectory of the rays in the third medium is very hard to determine analytically because the corresponding refraction indices must be obtain by "cutting" the ellipsoid of indices, which is tilted to 45°, with the (yz) plane, or the plane of incidence. Instead of using the relation (5.1.3) for the determination of n_e we have to use now the system (5.1.1).

Simulations conducted with Mathematica have shown that the separation between the coupled

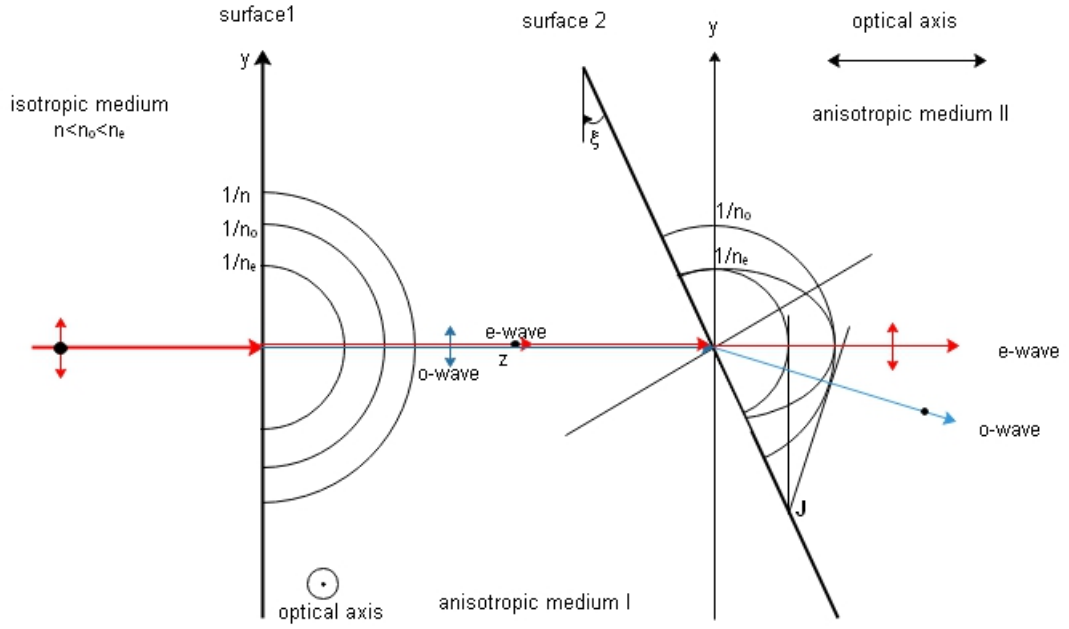


Figure 9.1: Ray tracing for the first and second bloc

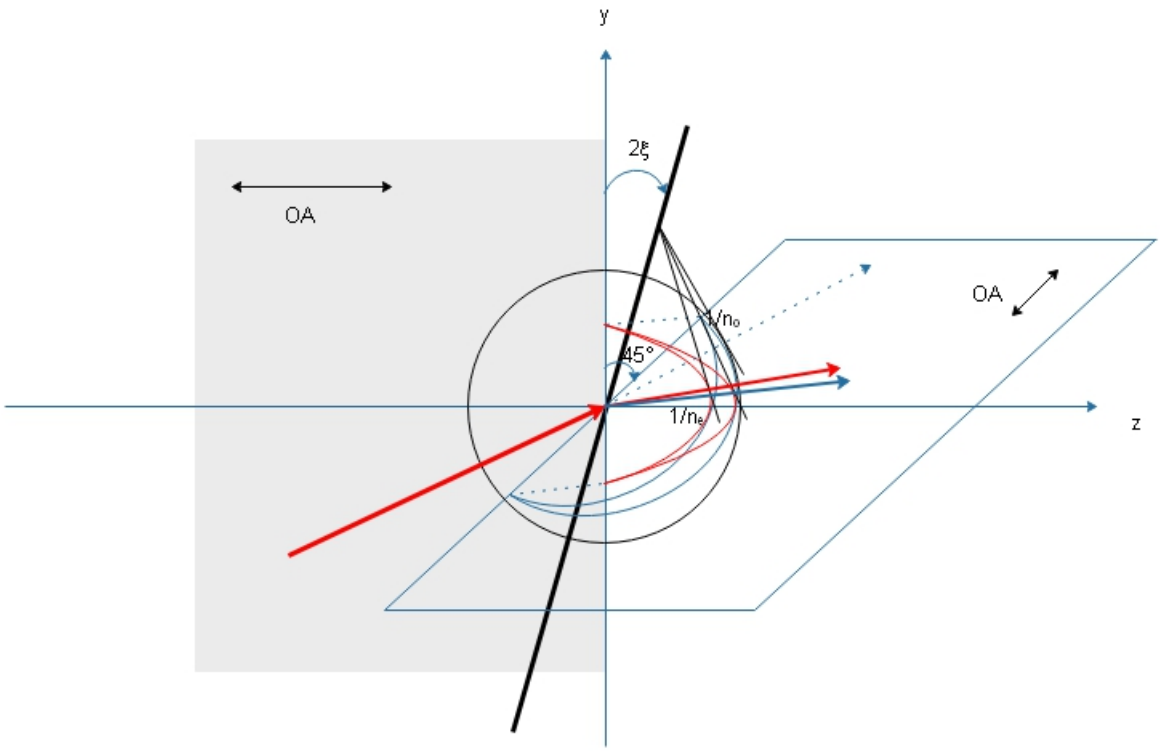


Figure 9.2: Ray tracing for the third bloc

outgoing points of exit from the third wedge is below $10^{-9}m$ for any incidence. Because of this we can neglect the separation of rays in the third wedge.

Overall, for normal incidence the rays' path inside the spectropolarimeter will follow the de-

scription from the Figure 9.3, where we neglected the separation between the rays occurring the last wedge.

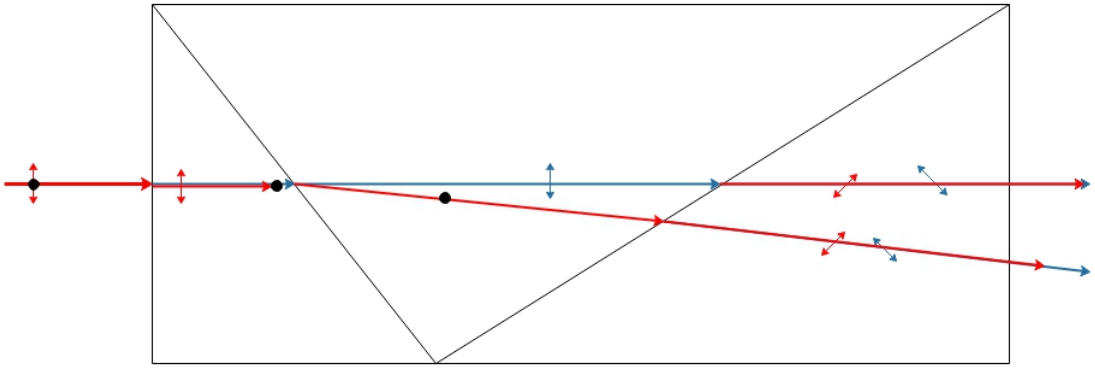


Figure 9.3: Ray tracing for the entire spectropolarimeter under normal incidence

The first conclusion of this detailed description that must be pointed out is that the assumption of a single ray path for the normal incidence is contradicted by the theory. Thus, the middle part of the instrument cannot be considered as a "neutral" zone while the last wedge will require a different approach for the phase calculation. Following its "activity", the middle part will introduce a non-zero phase difference between the ordinary and extraordinary rays. In the end, this will affect the Mueller calculus, because the corresponding matrix cannot be considered any longer as an identity matrix. Or changing the value of the Mueller matrix means that the outgoing intensity will be differently modulated.

Further we will try to calculate the exact deviation of light for the normal incidence in order to properly determine the phase along the instrument.

For the first wedge, exactly how it was considered before, the length of the ray path for normal incidence is:

$$z_1 = (h - y) \cdot \tan \xi$$

and the phase difference along this path:

$$\Delta\phi_1(y, \lambda) = \frac{2\pi}{\lambda} \cdot \Delta n(\lambda)(h - y) \tan \xi \tag{9.1.1}$$

In the second medium, the path length of the e-ray is (see Figure 9.4):

$$z_{2e} = x_C - x_B = y(\tan \xi + \tan 2\xi) \tag{9.1.2}$$

The o-wave from the second medium corresponds to the e-wave from the first. Thus, by using the Snell-Descartes law:

$$n_e \sin \alpha_1 = n_o \sin ro_2$$

Then we will have:

$$ro_2 = \arcsin\left(\frac{n_e}{n_o} \sin \alpha_1\right) \tag{9.1.3}$$

In this case of a normal incidence the angle α_1 is the angle between the e-ray from the first bloc and the normal to the surface 2, thus being equal to ξ. On the other hand, ro_2 is the corresponding angle of refraction in the second medium. Therefore, the angle between the e-ray and the horizontal is:

$$\beta = ro_2 - \xi$$

The entire geometry of this problem is depicted in the Figure 9.4.

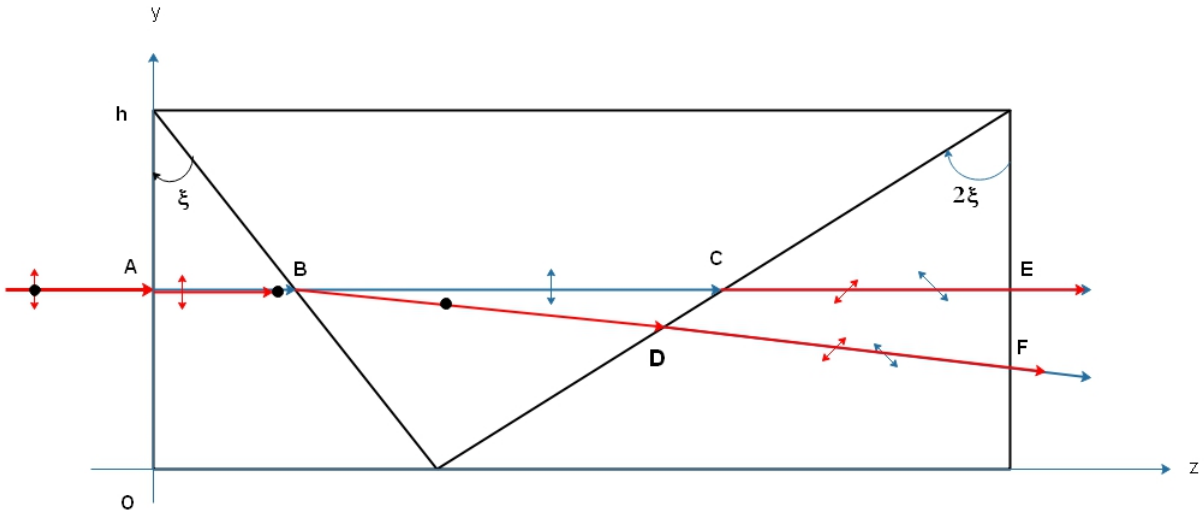


Figure 9.4: Geometry of the normal incidence

In order to compute the coordinates x_D, y_D we use the system:

$$\begin{cases} \tan(2\xi) = \frac{z_C - z_D}{y - y_D} \\ \tan\beta = \frac{y - y_D}{z_D - z_B} \end{cases} \tag{9.1.4}$$

Solving this system for (z_D, z_D) we find easily:

$$\begin{cases} z_D = h\tan\xi + y\tan(2\xi)\frac{1-\tan\beta\tan(\xi)}{1+\tan\beta\tan(2\xi)} \\ y_D = y\cdot\left(\frac{1-\tan\beta\tan(\xi)}{1+\tan\beta\tan(2\xi)}\right) \end{cases} \tag{9.1.5}$$

With the help of this values we may compute now the distance DB and then the phase difference induced by the second element of the bloc for the case of a normal incidence:

$$\Delta\phi_2(y,\lambda) = \frac{2\pi}{\lambda}(n_e(\lambda)\cdot BC - n_o(\lambda)\cdot BD) \tag{9.1.6}$$

In the last wedge the rays follow the same path as in the middle, without deviation. Thus, the distance CE will be:

$$CE = (h-y)\tan(2\xi) \tag{9.1.7}$$

Meanwhile, for DF we have:

$$\begin{cases} x_F = L = h(\tan\xi + \tan(2\xi)) \\ y_F = y(1 - \tan\beta\tan(\xi)) - h\tan\beta\tan(2\xi) \end{cases} \tag{9.1.8}$$

Nevertheless, in the third wedge each of the incoming e and o wave is split in two. But the geometry of this part of the instrument, with an angle of 45° between the optical axis and the plane of incidence, makes that the exit points E and F to be separated by less than $0.01\mu m$, all along the y axis, for any λ. The same thing is true on horizontal.

Because of the small angle between the two rays in each case, we decided then to neglect the geometrical effect of the splitting. But for each path (CE,DF) we have a phase difference. Accordingly, we will consider that the total phase difference is the average of the two situations:

$$\Delta\phi_3(y,\lambda) = \frac{\Delta\phi_{31}(y,\lambda) + \Delta\phi_{32}(y,\lambda)}{2} \tag{9.1.9}$$

where

$$\Delta\phi_{31}(y,\lambda) = \frac{2\pi}{\lambda} n_o(\lambda) CE \tag{9.1.10}$$

$$\Delta\phi_3(y,\lambda) = \frac{2\pi}{\lambda} n_o(\lambda) DF \tag{9.1.11}$$

Overall, the phase difference from the second medium and the new expression of the phase difference for the last wedge will alter the shape of the outgoing intensity. A comparison between the ideal case, of a single beam and a split beam in normal incidence can be seen in the Figure 9.5.

9.2 Tilted rays in (yz) plane

Using the notations from the Figure 9.6 let us consider an incident ray making an angle α with the normal, the incidence plane being parallel to the (yz) plane. Just as before we will try to calculate the path length of each ray inside the optical instrument in order to find an expression for the phase difference corresponding to each element. Thus, in the first wedge the extraordinary ray will follow an angle r_e inside the anisotropic medium, while for the ordinary the refraction angle will be r_o.

Because the incoming ray is seeing an isotropic medium regardless the value of α (e-wave being everywhere parallel to the optical axis and o-wave perpendicular), the Snell-Descartes low can be applied without further considerations:

$$\begin{cases} r_o = \arcsin\left[\frac{n}{n_o} \cdot \sin\alpha\right] \\ r_e = \arcsin\left[\frac{n}{n_e} \cdot \sin\alpha\right] \end{cases} \tag{9.2.1}$$

Just as before we may compute again the phase difference characterizing the first wedge:

$$\Delta\phi_1(y,\lambda) = \frac{2\pi}{\lambda}\left(\frac{\sqrt{1+\tan^2 r_e}}{\tan r_e + \cot\xi} \cdot n_e(\lambda) - \frac{\sqrt{1+\tan^2 r_o}}{\tan r_o + \cot\xi} \cdot n_o(\lambda)\right) \cdot (h - y) \tag{9.2.2}$$

Is easy to see that this expression is tending towards (9.1.1) for the normal incidence.

In the second medium, the situation is a little bit more complicated and will require a detailed examination. Taking at first the case of the incoming o-ray that appears as e-ray for the second

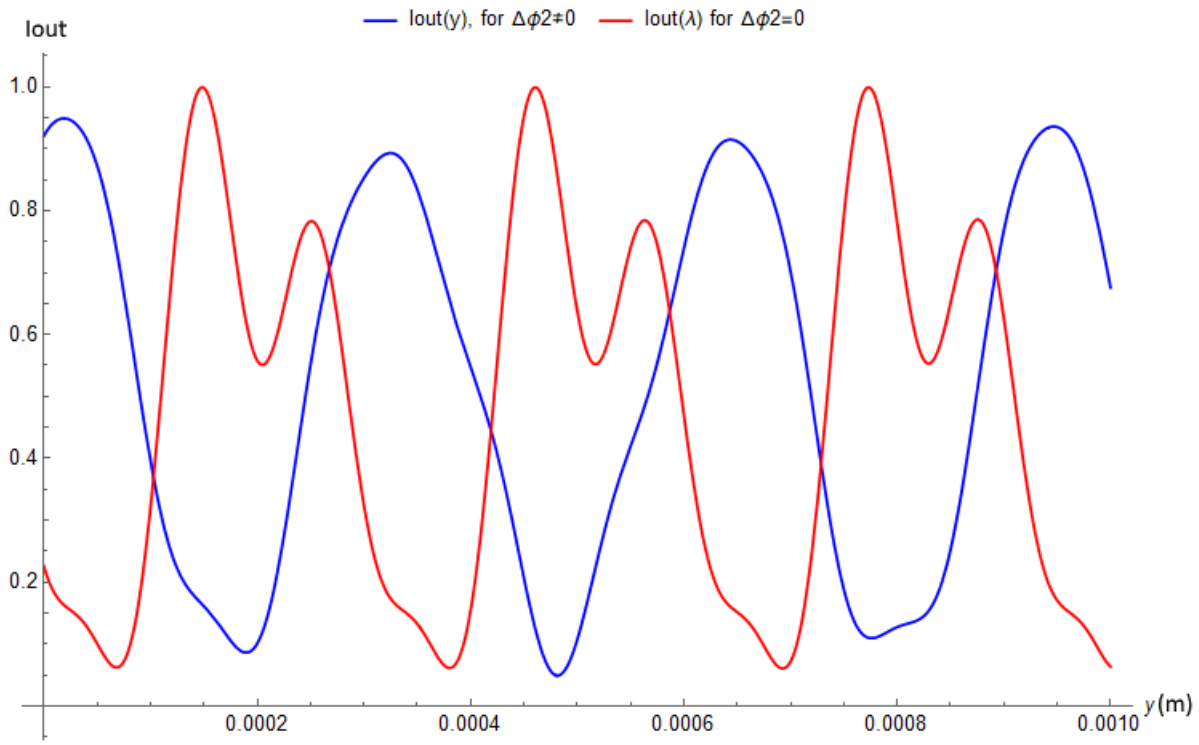


Figure 9.5: Received intensity along y for normal incidence, $\lambda = 0.125\mu m$, for the case $\Delta\phi_2 = 0$ (red) and for the case $\Delta\phi_2 \neq 0$ (blue). The fact that the second prism induces a phase difference changes the shape of the outgoing intensity.Just as before, the units of I_{out} are the arbitrary units of intensity (most often W/m^2 or V)

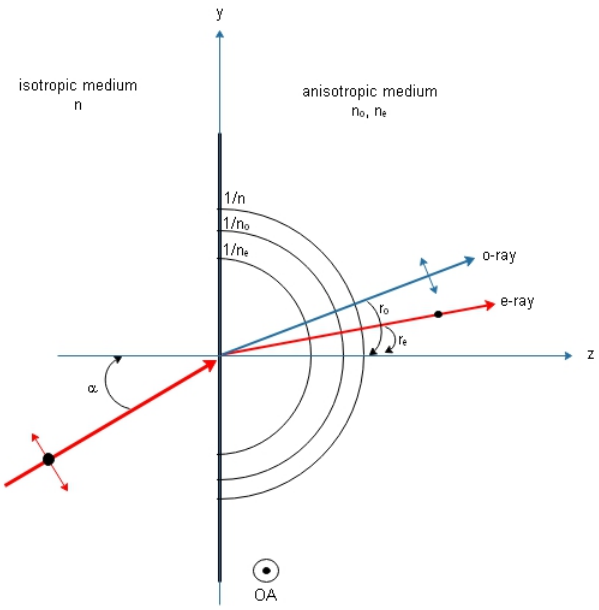


Figure 9.6: Huygens construction for the tilted scenario in the first medium

medium given the orientation of the optical axis, we will note with $r_e^{'}$ the corresponding refraction angle. Therefore, the total angle between the refracted ray and the horizontal will be $r_e^{'} + \xi$ (see

Figure 9.7). The generalized Snell-Descartes low tells us that:

$$n_o \cdot \sin(r_o - \xi) = n(r_o) \cdot \sin r_e^{'} \tag{9.2.3}$$

where θ_e is the angle between the refracted ray and the optical axis, thus $r_e^{'} + \xi$. In order to compute $n(r_o)$ we have to use the relation (5.1.4). In this way we infer:

$$n(r_o) = \frac{n_o n_e}{\sqrt{\sin^2 r_o n_o^2 + \cos^2 r_o n_e^2}} \tag{9.2.4}$$

Coming back to (9.2.3) we can build the equation for $r_e^{'}$ which can be solved by knowing the values of r_o and ξ:

$$\sin(r_o - \xi) = \frac{n_e}{\sqrt{\sin^2 r_o n_o^2 + \cos^2 r_o n_e^2}} \sin r_e^{'} \tag{9.2.5}$$

Then it will be enough to calculate $\theta_e = r_e^{'} + \xi$ to have the angle with the horizontal.

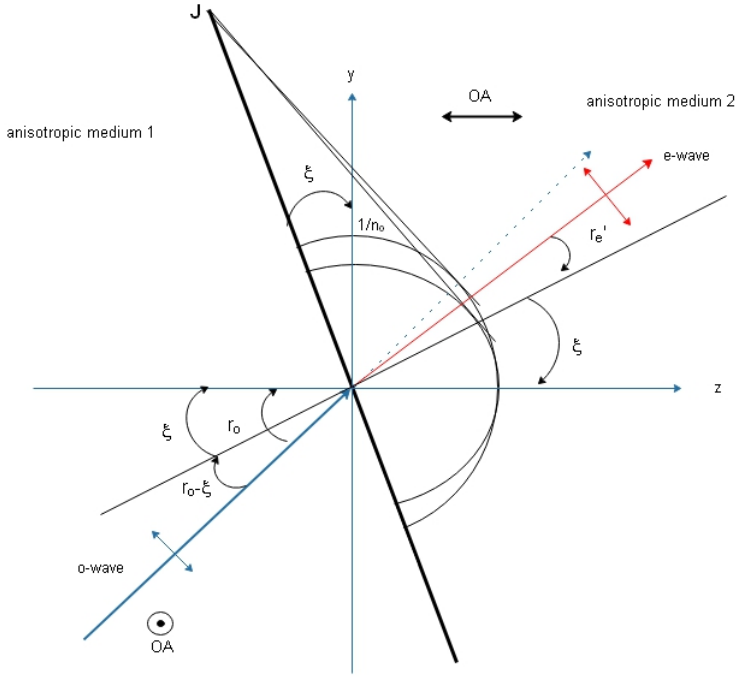


Figure 9.7: Huygens construction for the tilted scenario in the second medium

For the second ray we saw that the extraordinary ray from the first medium becomes ordinary here, with an electrical field vibrating perpendicular to optical axis everywhere in the plane (yz). Thus, this ray is passing from an isotropic medium into another, from an index of refraction n_e, to an index n_o. Simply by using the Descartes low we have:

$$n_e \cdot \sin(r_e - \xi) = n_o \cdot \sin(r_o^{'}) \tag{9.2.6}$$

Thus, the angle with the horizontal will be:

$$\theta_o = \xi + \arcsin\left(\frac{n_e}{n_o} \sin(r_e - \xi)\right) \tag{9.2.7}$$

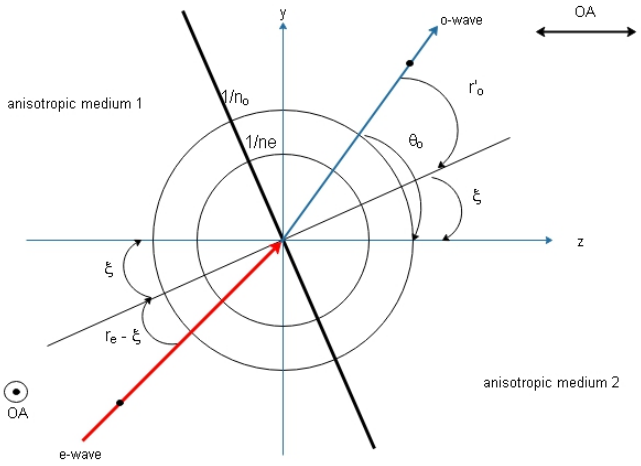


Figure 9.8: Schematic view of the ordinary ray in the second medium

A schematic representation of this situation can be retrieved in the Figure 9.8.

Having the values of the angles for each ray we may compute the path length for the ordinary and extraordinary beam: l_{e2}, l_{o2}.

Then the phase difference corresponding to the second body of the system is:

$$\Delta\phi_2(y,\lambda) = \frac{2\pi}{\lambda}(n(\lambda)l_{e2} - n_o(\lambda)l_{e2}) \tag{9.2.8}$$

where $n(\lambda)$ should be determined according to the angle r_o.

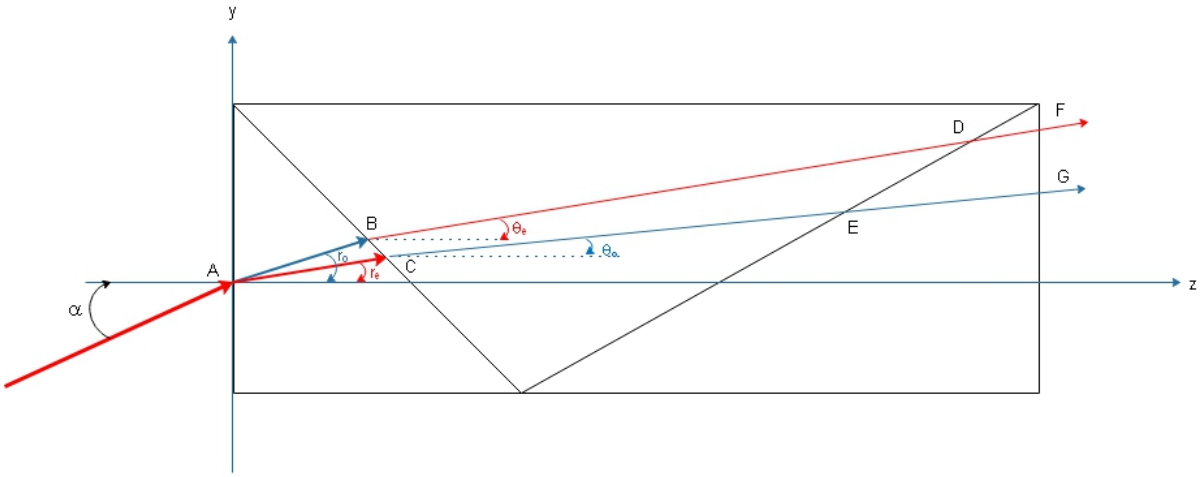


Figure 9.9: General reprezentation of the wedges

In the last wedge, because we neglect the splitting of each of the incoming rays, we can consider that the phase difference is expressed again by the relation (9.1.9),(9.1.10) and (9.1.11).

Having now these new values of the phase difference for every element of the optical device we may properly compute the outgoing intensity (Figure 9.10 and 9.11).

Therefore, we may observe that the main consequence of the off-axis incidence is the shifting of the intensity pattern in the image plane. Simultaneously, the ”irregularities” exhibited by the

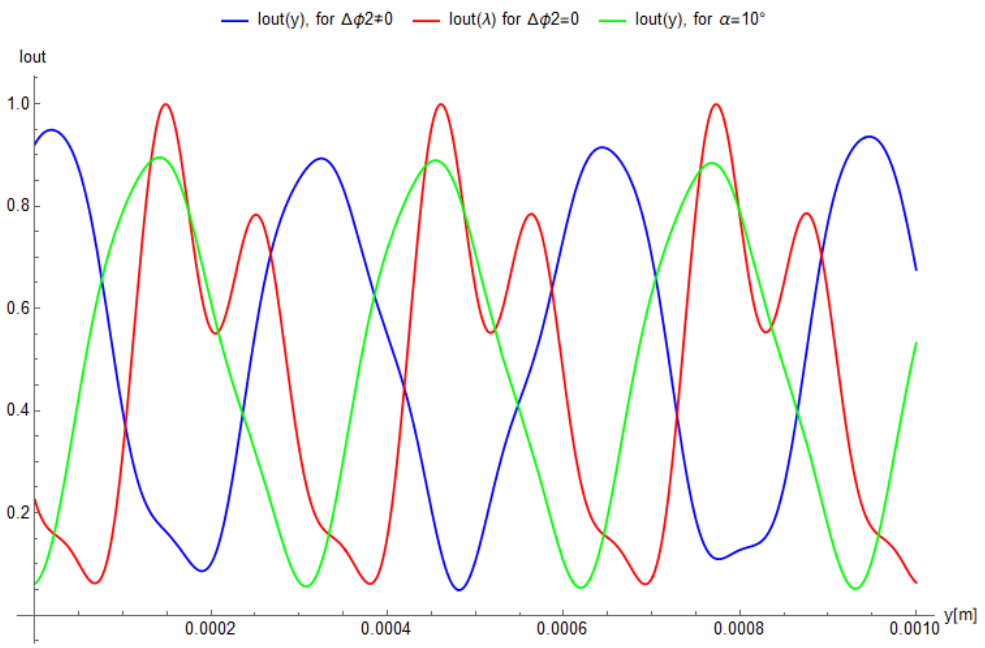


Figure 9.10: Outgoing intensity for $\lambda = 0.12\mu m$ for normal incidence without splitting inside the optical device (red) and with splitting (blue). In green we have the intensity for an incidence $\alpha = 10^{\circ}$

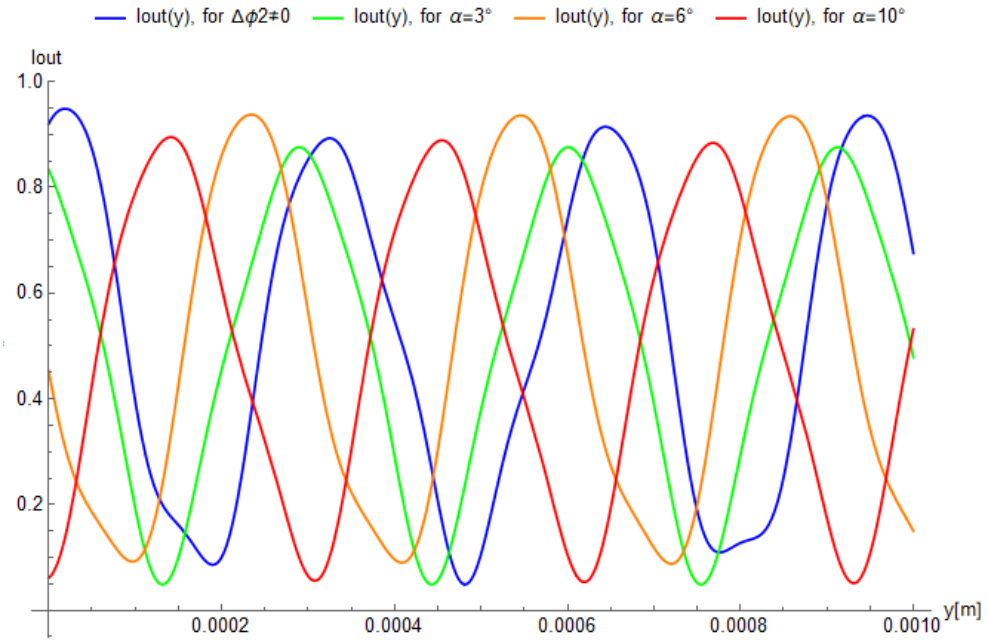


Figure 9.11: Outgoing intensity for $\lambda = 0.125\mu m$ for normal incidence with splitting (blue), and for tilted incoming rays: $\alpha = 3^{\circ}$ (green), $\alpha = 6^{\circ}$ (orange) and $\alpha = 10^{\circ}$ (red)

modulation in the normal incidence are erased by a big incidence angle (more than 10°). Because of the this change of the pattern, the retrieval algorithm used before to compute the value of the Stokes parameters must be adapted to take into consideration the presence of tilted rays. Moreover, the adaptation of this algorithm exceeds the objective of this thesis, where we will try only to visualize the effects of the off-axis incidence.

9.3 Tilted rays into the horizontal plane (xz)

We may consider now that the incoming ray, traveling in the plane (xz), makes an angle α with the normal to the left side of the instrument. Again, this ray will be split into an ordinary and extraordinary ray into the the first wedge. The geometry of this situation is presented in the Figures 9.12 and 9.13. In the second part of the instrument, the rays will follow the paths represented in the Figure 9.14. Meanwhile, because of the special orientation of the optical axis in the last wedge, we will neglect again the deviation that happen here (Figure 9.15). Nevertheless, because each ray entering the last wedge undergoes a new splitting, with a very small angle that can be neglected, we will consider that over each path the phase difference can be expressed via:

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta n(\lambda) l_i \tag{9.3.1}$$

where l_i is the distance traveled in the third block by each ray (i: ordinary and extraordinary). Consequently, the phase difference from the last part of the optical instrument is the average of the two phase differences existing here. Because the computation of the phase difference is based on the same type of geometrical and optical considerations as before, we will skip this time the details.

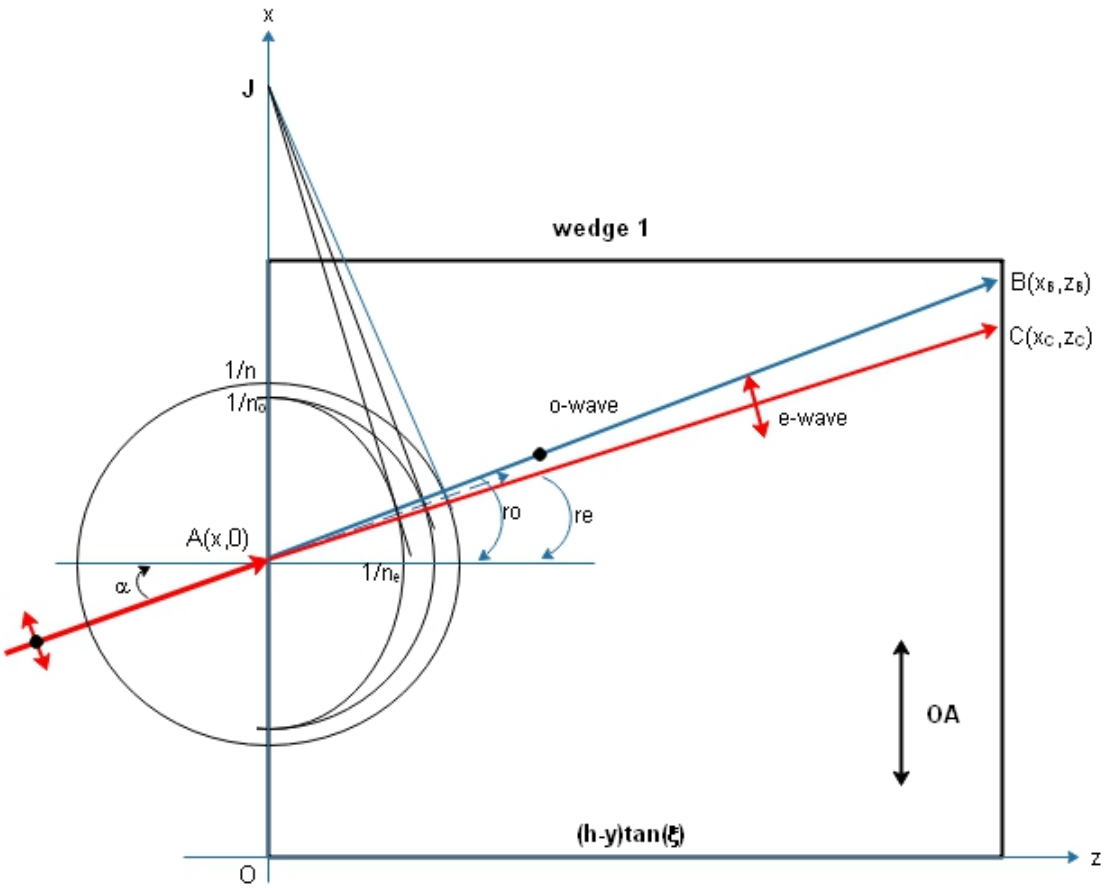


Figure 9.12: Ray tracing and Huygens construction for the first block for an incidence at angle α

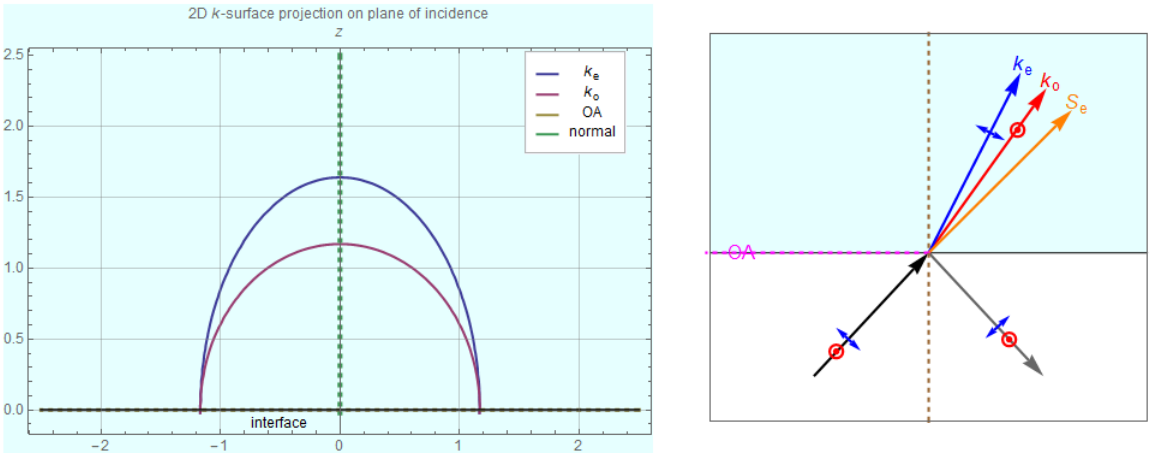


Figure 9.13: Simulation for the first wedge with Mathematica software. At left we have a 2D projection of the k surface for two indices n_e and n_o arbitrary chosen to point out the difference between the two surfaces. At left we have the vectorial reprezentation: the incoming ray (black), the reflected ray (gray), the extraordinary ray (blue), the ordinary (red) and the pointing vector (Se)[1].

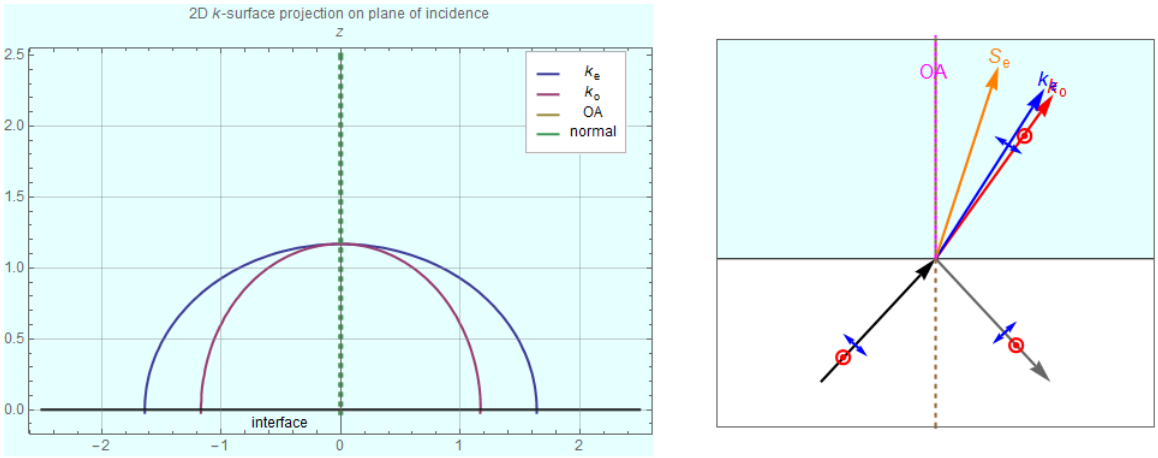


Figure 9.14: Simulation for the second wedge with Mathematica software.The notations are the same as in the previous figure. We notice now the change of the surface of indices.[1].

The numerical simulations for several values of the incidence angle α shows again a shift of the intensity pattern in the image plane, accompanied by a smoothing of the intensity profile.

The displacement of the outgoing profile of the intensity with the incidence angle requires an adjustment of the extraction algorithm previously used for the computation of the I, Q, U and V

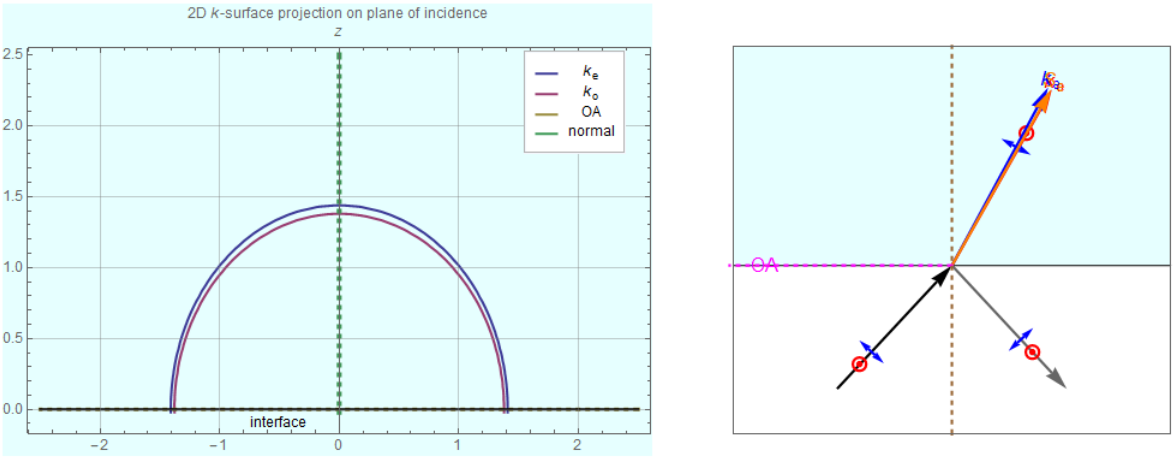


Figure 9.15: Simulation for the third wedge with Mathematica software.The extraordinary and ordinary ray are almost indistinct as well as the surfaces of indices even if we kept a birefringence of about 0.1[1].

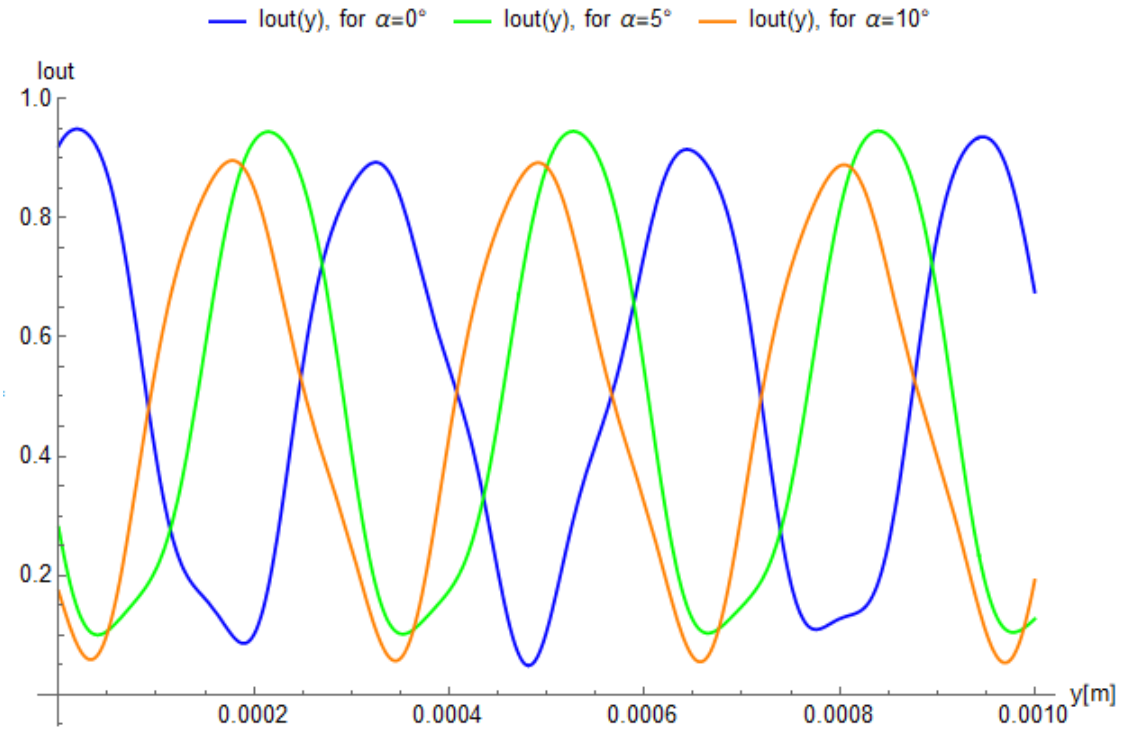


Figure 9.16: Profile of the received intensity along y with regard to the incidence angle, α

parameters. But this is a new and laborious task which exceeds the purpose of this thesis. It may as well represent the subject of further research.

9.4 Extended target

The best results of the studied spectropolarimeter are obtained under the condition of normal incidence. It is only in this case that the device maximizes its capacity and it fulfills its purpose. Once the normal incidence is put aside, the difficulties start to arise. Thus, in a very simple scenario let us consider at first the case of an uni-dimensional objected situated along y axis. Instead of having a normal view of this object, we will consider that all the rays are coming from it with a tilt α. This scenario corresponds in fact, for example, to the situation of a misalignment between the polarimeter and the collimator that may come in front, or to a chromatic aberration (Figure 9.15) or simply to the misalignment with the source. We will suppose that the extended source is homogeneous (Lambertian) and characterized by the same polarization, regardless the angle, this polarization being represented by the arbitrary state $S_{in} = [1, 0.55, 0.703, 0.45]$, corresponding to a total polarization.

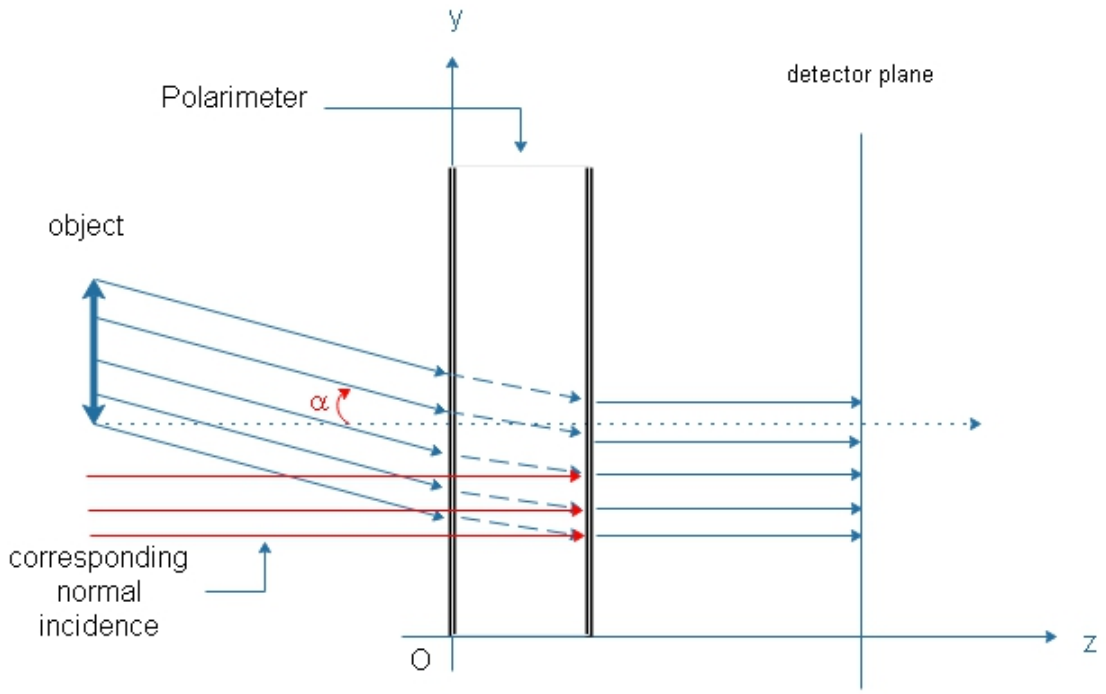


Figure 9.17: Off-axis uni-dimensional object in the (yz) plane. We consider that all the rays coming from the object are tilted to an angle α with regard to the normal position. In red we have the normal rays corresponding to the same exit positions.

The numerical simulations for different values of the angle α shows a continuous variation of the outgoing intensity with this inclination (Figure 9.18). However, the signal is not totally distorted and with an appropriate extraction algorithm, able to detect the inclination of the incoming rays, the Stokes vector of the incident light probably could be determined. Things are not the same on horizontal. Here the situation is much more easy, at least for this type of simulation. Thus, if we imagine a similar emitting object, but this time on horizontal (the same figure as 9.17, only with x instead of y), the deviation from the normal incidence will be translated into a different level of the received intensity. But this level will be the same for the entire area of reception, because there is no modulation along x axis as long as α remains constant. The situation is depicted in the figure 9.19, where aside the normal incidence (blue), an incidence at 10° was also plotted (red).

A last question that can be addressed here is what is happening if we have simultaneously an

inclination along x and y. To answer, we imagined an object in the plane (xy), making an angle of 45° with x but received by the spectropolarimeter with an inclination of 10° on x and 10° about y (see the picture 9.20). The plot of the received intensity along the inclined direction from the detector plane shows a relatively small variation regarding the normal incidence (Figure 9.21), the effect of the deviation along x being marked by the variation in amplitude.

All these simulations show an almost steady behavior of the spectropolarimeter in the presence of tilted but still parallel rays. However, the exploitation of the modulation requires, how it was mentioned before, a modification of the extraction algorithm so that the inclination of rays to be considered also.

Nevertheless, the situation seems to be more complicated if in the entry we have multiples angles. Thus, if we take the case of a simple uni-dimensional object along y seen under a FOV of 10°, then the modulation of the intensity will become unreadable (Figure 9.22). Consequently, the necessity to ensure the presence of parallel rays inside the instrument is of paramount importance. This situation is depicted in the Figure 9.23, where we see in red the very fast modulation of the tilted rays. The situation can be explained by the fact that the tilt of rays increases the speed of variation of the phase because of the longer distances traveled inside the medium.

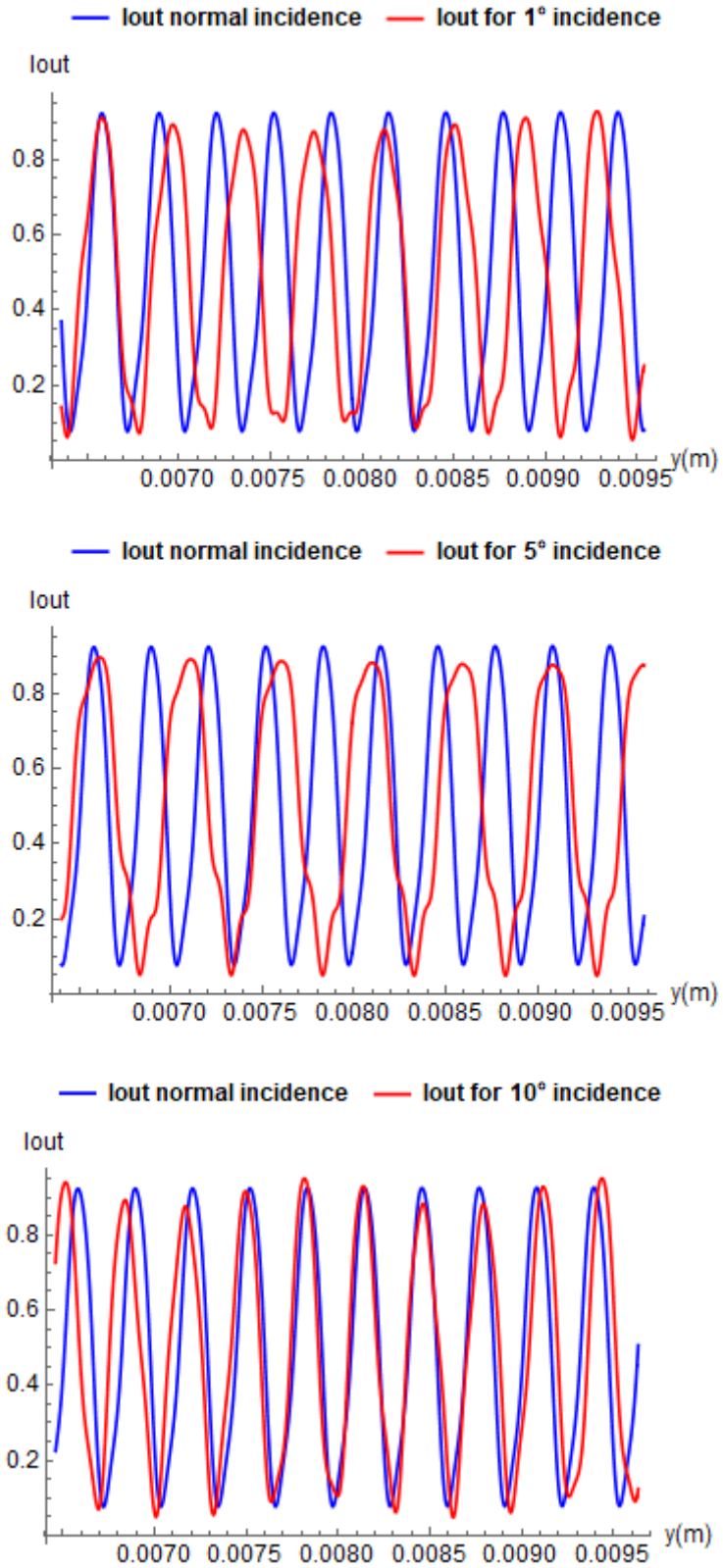


Figure 9.18: Comparison between the signal received in normal incidence from the uni-dimensional object situated along y axis, and the same object with tilted incidence, at angle α of 1°, 5° and 10°. The comparison was realized over a distance of $3mm$ (in the detector plane).

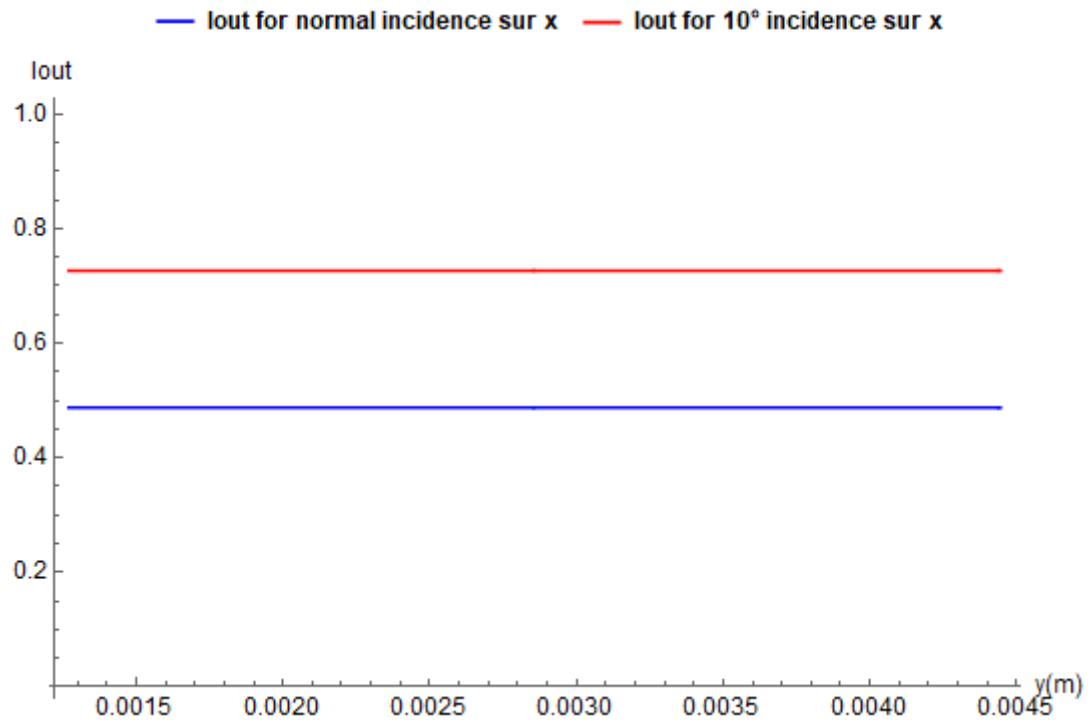


Figure 9.19: Received intensity at a high of $6mm$ (the middle of the considered prism) from a unidimensional object along x with a length of about $3mm$. Blue line is for normal incidence, red line for tilted at 10°.

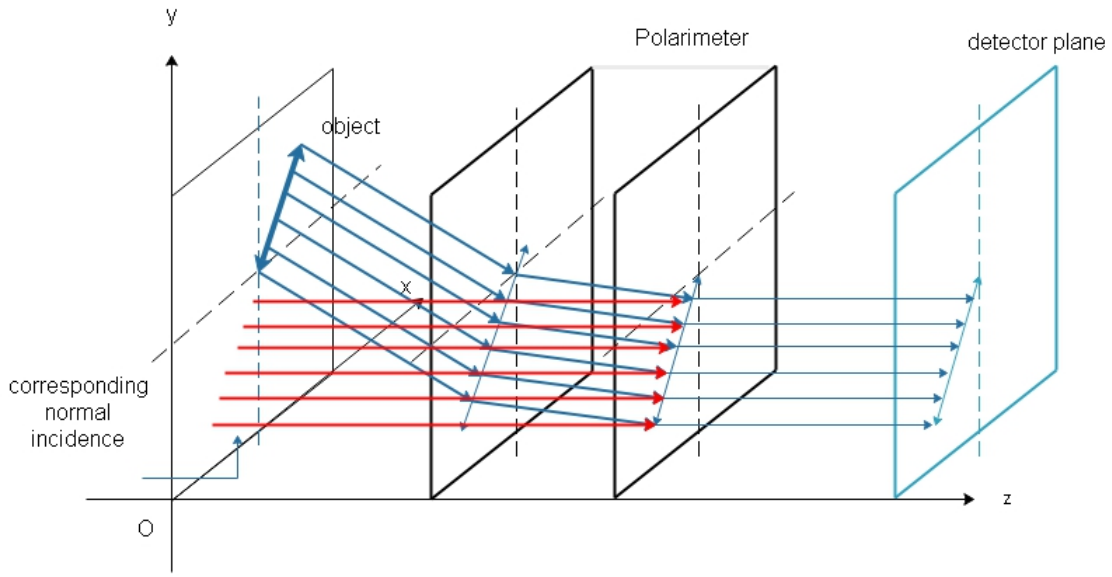


Figure 9.20: Geometry of an inclination of 10° on x and y. The object is tilted to 45° about x axis

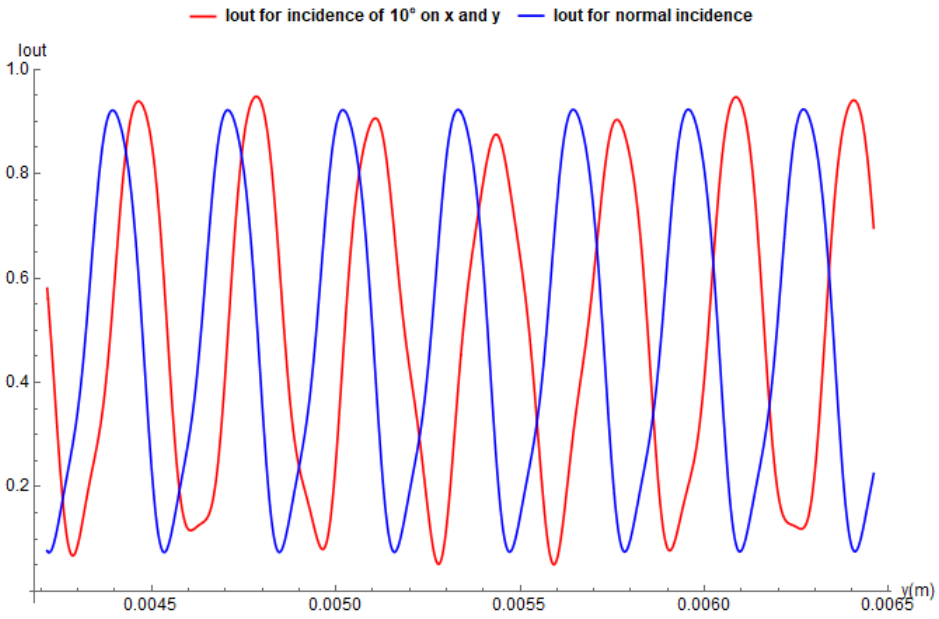


Figure 9.21: Pattern of the received intensity along the diagonal direction for tilted rays (red) and normal incidence (blue)

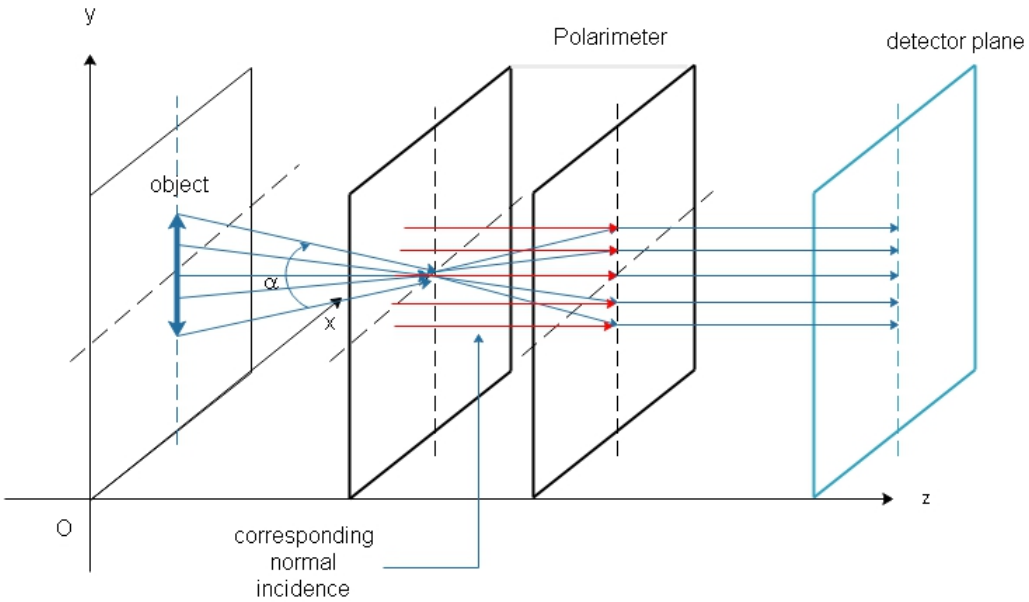


Figure 9.22: Geometry of a FOV α along y. In red we have the corresponding normal incidence. Many scenarios are possible, but here we considered that the polarimeter is placed at the focal point.

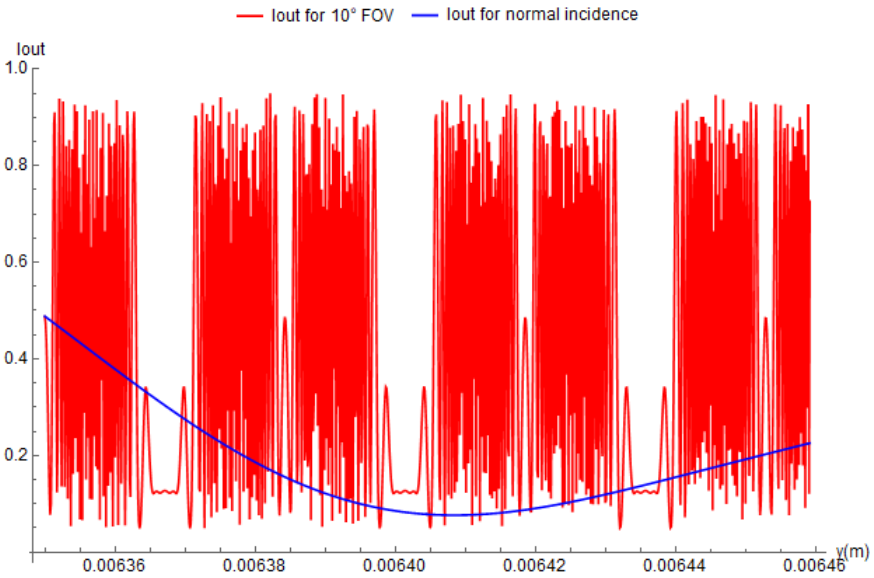


Figure 9.23: Received intensity from an object observed under a FOV of 10° (red) and for normal incidence (blue). Here again, the units of intensity are arbitrary (W/m^2 or V usually). In this case we imposed $\lambda = 0.125\mu m$.

9.5 Conclusions

- The detailed representation of the ray tracing for the normal incidence proves the splitting of the rays inside the optical device. Because of this phenomenon the phase difference corresponding to each part of the instrument is different from the approximation used in the Chapter 6. Consequently, the intensity is differently modulated by the birefringent medium.
- The phase difference is bigger than zero in all the parts of the instrument.
- Tilted rays in the (yz) plane (on the vertical) will generate shifted and smoothed profile of the intensity. The same is true for the rays in (xz) plane.
- An extended target on vertical, source of parallel rays, will generate a profile analogue to the normal target, but shifted and slightly distorted as a function of the angle of incidence. An extended target on horizontal will generate a constant signal.
- Parallel rays with the same inclination along x and y will generate also a sinusoidal modulation shifted and altered.
- Rays coming from a FOV bigger than 1° (along y or x axis) will impose a very fast variation of the signal.

Chapter 10

Conclusions and perspectives

Despite the unsophisticated structure, the spectropolarimeter presented here has proved itself to be a very complex instrument, requiring a deep and large analysis in order to understand its functioning.

Therefore, we saw during this research that every little detail can play a major role in the efficiency of the instrument. First, we observed that analyzer and apex angles (θ and ξ) are dominating the behavior of the device. Apparently, there is no "perfect" solution for these angles, able to maximize simultaneously the detection of all the types of incoming polarization. Thus, in the future, a straight classification of the best choices for the angles regarding the desired type of measurement and sensibility should conducted. Theoretically, the polarimeter can descend far below a degree of linear polarization of 10^{-2} and 10^{-3} for circular polarization. And these results qualify it for the space use.

On the other hand, in this approach we considered that lateral wedges have an apex of ξ and 2ξ respectively. Nevertheless, in the case of the off-axis incidence, this configuration determines a very fast variation of the phase difference in the exit wedge. It would be then appropriate to study a design with apexes of closer values.

Another aspect that should be investigated is also the possibility to use an isotropic medium between the two wedges or even to separate them, because in the presence of tilted rays this part of the instrument distorts the signal.

Finally, the most important part that should be studied in detail in the next future is the behavior in the presence of tilted rays. We observed that non-parallel rays entering the system can drastically alter the functioning of the device. Therefore, an optical design able to minimize this phenomenon should also be studied. Simultaneously, for the parallel and tilted rays an extraction algorithm for the Stokes parameters must be developed.

Overall, the spectropolarimeter presented here is still far for being completely understood and optimized. It requires even more and deep analyze. But the most important step was made: there is enough proof that the device can work under normal incidence, and even in certain cases of the off-axis. From now on the entire effort should be dedicated to the optimization of this spectropolarimeter.

Appendix A

Birefringence of MgF_2

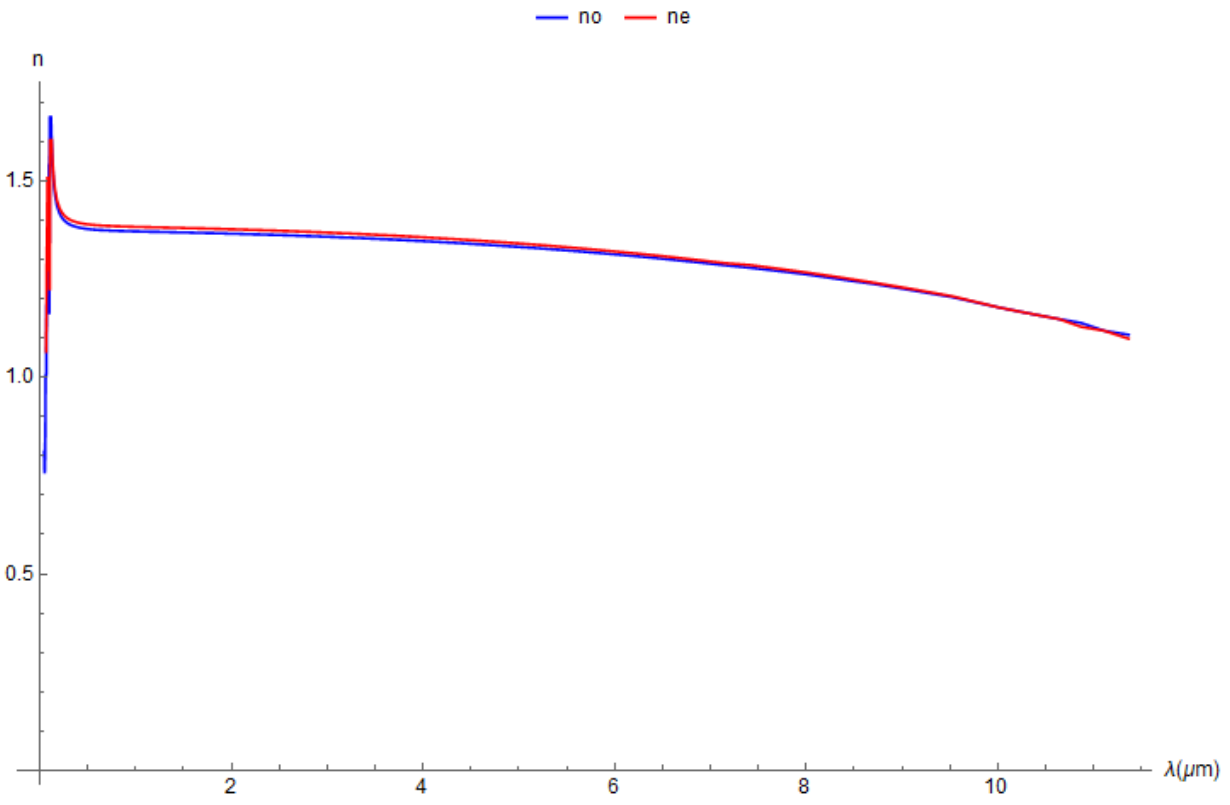


Figure A.1: Variation of the ordinary (n_o) and extraordinary (n_e) indexes of refraction of MgF_2 with regard to wavelength (μm) at $19^{\circ}C$, based on the experimental data from Table A.1

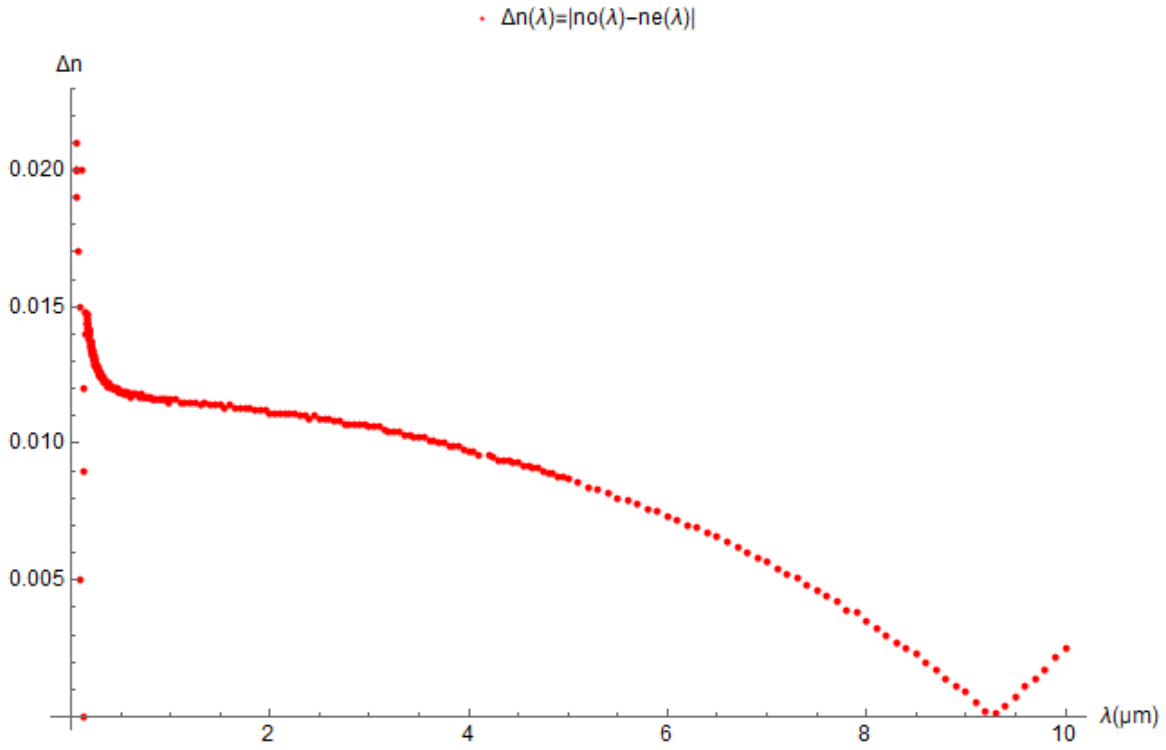


Figure A.2: Birefringence of MgF_2 between $0.04\mu m$ and $10\mu m$ at $19^{\circ}C$ according to the data presented in Table A.1

Table A.1: Birefringence of MgF_2 between $0.04\mu m$ and $10.0\mu m$ at $19^{\circ}C$. The data was extracted from [5][6][11][13][14][27]

$\lambda(\mu m)$	n_o	n_e	$\Delta n = n_e - n_o$
0.0459	0.021	0.832	0.021
0.0468	0.02	0.816	0.02
0.0477	0.019	0.803	0.019
0.0486	0.02	0.79	0.02
0.0496	0.029	0.794	0.029
0.0506	0.043	0.799	0.043
0.0517	0.044	0.807	0.044
0.0528	0.042	0.816	0.042
0.0539	0.037	0.827	0.037
0.0551	0.028	0.837	0.028
0.0577	0.02	0.866	0.02
0.059	0.078	0.881	0.078
0.0605	0.109	0.894	0.109
0.062	0.105	0.905	0.105
0.0636	0.087	0.919	0.087
0.0653	0.017	0.983	0.017
0.067	0.109	1.117	0.109
0.0689	0.071	1.151	0.071
0.0708	0.084	1.264	0.084
0.0729	0.128	1.317	0.128
0.0751	0.167	1.353	0.167
0.0775	0.232	1.416	0.232
0.08	0.321	1.513	0.321
0.0827	0.033	1.387	0.033
0.0855	0.043	1.299	0.043
0.0886	0.015	1.272	0.015
0.0918	0.005	1.348	0.005
0.0954	0.097	1.449	0.097
0.0969	0.142	1.423	0.142
0.0984	0.257	1.419	0.257
0.1	0.248	1.47	0.248
0.1016	0.119	1.346	0.119
0.1033	0.47	1.223	0.47
0.1051	0.58	1.343	0.58
0.1069	0.307	1.636	0.307
0.1088	0.077	1.911	0.077
0.1107	0.287	1.988	0.287
0.1127	0.393	1.931	0.393
0.115	0.02	1.694	0.02
0.12	0	1.626	0
0.125	0.009	1.597	0.009
0.13	0.012	1.567	0.012
0.14	0.014	1.527	0.014
0.15	1.4792	1.4941	0.0149
0.152	1.4752	1.49	0.0148
0.154	1.4715	1.4863	0.0148
0.156	1.468	1.4827	0.0147
0.158	1.4647	1.4794	0.0147
0.16	1.4616	1.4762	0.0146
0.162	1.4587	1.4732	0.0145
0.164	1.4559	1.4703	0.0144
0.166	1.4532	1.4676	0.0144
0.168	1.4507	1.4651	0.0144
0.17	1.4483	1.4626	0.0143
0.172	1.4461	1.4603	0.0142
0.174	1.4439	1.4581	0.0142
0.176	1.4419	1.456	0.0141
0.178	1.4399	1.454	0.0141
0.18	1.438	1.4521	0.0141
0.182	1.4362	1.4502	0.014
0.184	1.4345	1.4485	0.014
0.186	1.4329	1.4468	0.0139
0.188	1.4313	1.4451	0.0138
0.19	1.4298	1.4436	0.0138
0.192	1.4284	1.4421	0.0137
0.194	1.427	1.4407	0.0137
0.196	1.4256	1.4393	0.0137
0.198	1.4244	1.438	0.0136
0.2	1.4231	1.4367	0.0136
0.202	1.422	1.4355	0.0135
0.204	1.4208	1.4343	0.0135
0.206	1.4197	1.4332	0.0135
0.208	1.4187	1.4321	0.0134
0.21	1.4176	1.431	0.0134
0.212	1.4166	1.43	0.0134
0.214	1.4157	1.429	0.0133
0.216	1.4148	1.4281	0.0133
0.218	1.4139	1.4272	0.0133
0.22	1.413	1.4263	0.0133
0.222	1.4122	1.4254	0.0132
0.224	1.4114	1.4246	0.0132
0.226	1.4106	1.4238	0.0132
0.228	1.4098	1.423	0.0132
0.23	1.4091	1.4222	0.0131
0.232	1.4084	1.4215	0.0131
0.234	1.4077	1.4208	0.0131
0.236	1.4071	1.4201	0.013
0.238	1.4064	1.4194	0.013

$\lambda(\mu m)$	n_o	n_e	$\Delta n = n_e - n_o$
0.24	1.4058	1.4188	0.013
0.242	1.4052	1.4181	0.0129
0.244	1.4046	1.4175	0.0129
0.246	1.404	1.4169	0.0129
0.248	1.4034	1.4163	0.0129
0.25	1.4029	1.4158	0.0129
0.252	1.4024	1.4152	0.0128
0.254	1.4019	1.4147	0.0128
0.256	1.4014	1.4142	0.0128
0.258	1.4009	1.4137	0.0128
0.26	1.4004	1.4132	0.0128
0.262	1.3999	1.4127	0.0128
0.264	1.3995	1.4122	0.0127
0.266	1.399	1.4118	0.0128
0.268	1.3986	1.4113	0.0127
0.27	1.3982	1.4109	0.0127
0.272	1.3978	1.4105	0.0127
0.274	1.3974	1.4101	0.0127
0.276	1.397	1.4097	0.0127
0.278	1.3966	1.4093	0.0127
0.28	1.3963	1.4089	0.0126
0.282	1.3959	1.4085	0.0126
0.284	1.3956	1.4081	0.0125
0.286	1.3952	1.4078	0.0126
0.288	1.3949	1.4074	0.0125
0.29	1.3946	1.4071	0.0125
0.292	1.3942	1.4068	0.0126
0.294	1.3939	1.4064	0.0125
0.296	1.3936	1.4061	0.0125
0.298	1.3933	1.4058	0.0125
0.3	1.393	1.4055	0.0125
0.305	1.3923	1.4048	0.0125
0.31	1.3917	1.4041	0.0124
0.315	1.391	1.4034	0.0124
0.32	1.3904	1.4028	0.0124
0.325	1.3899	1.4022	0.0123
0.33	1.3893	1.4016	0.0123
0.335	1.3888	1.4011	0.0123
0.34	1.3883	1.4006	0.0123
0.345	1.3878	1.4001	0.0123
0.35	1.3874	1.3996	0.0122
0.355	1.387	1.3992	0.0122
0.36	1.3866	1.3988	0.0122
0.365	1.3862	1.3983	0.0121
0.37	1.3858	1.398	0.0122
0.375	1.3854	1.3976	0.0122

$\lambda(\mu m)$	n_o	n_e	$\Delta n = n_e - n_o$
0.38	1.3851	1.3972	0.0121
0.385	1.3848	1.3969	0.0121
0.39	1.3844	1.3965	0.0121
0.395	1.3841	1.3962	0.0121
0.4	1.3838	1.3959	0.0121
0.41	1.3833	1.3953	0.012
0.42	1.3828	1.3948	0.012
0.43	1.3823	1.3943	0.012
0.44	1.3818	1.3938	0.012
0.45	1.3814	1.3934	0.012
0.46	1.381	1.393	0.012
0.47	1.3807	1.3926	0.0119
0.48	1.3803	1.3923	0.012
0.49	1.38	1.3919	0.0119
0.5	1.3797	1.3916	0.0119
0.51	1.3794	1.3913	0.0119
0.52	1.3792	1.391	0.0118
0.53	1.3789	1.3908	0.0119
0.54	1.3787	1.3905	0.0118
0.55	1.3784	1.3903	0.0119
0.56	1.3782	1.3901	0.0119
0.57	1.378	1.3898	0.0118
0.58	1.3778	1.3896	0.0118
0.59	1.3776	1.3894	0.0118
0.6	1.3775	1.3892	0.0117
0.62	1.3771	1.3889	0.0118
0.64	1.3768	1.3886	0.0118
0.66	1.3765	1.3883	0.0118
0.68	1.3763	1.388	0.0117
0.7	1.376	1.3878	0.0118
0.72	1.3758	1.3875	0.0117
0.74	1.3756	1.3873	0.0117
0.76	1.3754	1.3871	0.0117
0.78	1.3752	1.3869	0.0117
0.8	1.375	1.3867	0.0117
0.82	1.3749	1.3865	0.0116
0.84	1.3747	1.3863	0.0116
0.86	1.3746	1.3862	0.0116
0.88	1.3744	1.386	0.0116
0.9	1.3743	1.3859	0.0116
0.92	1.3741	1.3857	0.0116
0.94	1.374	1.3856	0.0116
0.96	1.3739	1.3855	0.0116
0.98	1.3738	1.3853	0.0115
1	1.3736	1.3852	0.0116
1.05	1.3733	1.3849	0.0116

$\lambda(\mu m)$	n_o	n_e	$\Delta n = n_e - n_o$
1.1	1.3731	1.3846	0.0115
1.15	1.3728	1.3843	0.0115
1.2	1.3726	1.3841	0.0115
1.25	1.3723	1.3838	0.0115
1.3	1.3721	1.3835	0.0114
1.35	1.3718	1.3833	0.0115
1.4	1.3716	1.383	0.0114
1.45	1.3714	1.3828	0.0114
1.5	1.3711	1.3825	0.0114
1.55	1.3709	1.3822	0.0113
1.6	1.3706	1.382	0.0114
1.65	1.3704	1.3817	0.0113
1.7	1.3701	1.3814	0.0113
1.75	1.3699	1.3812	0.0113
1.8	1.3696	1.3809	0.0113
1.85	1.3694	1.3806	0.0112
1.9	1.3691	1.3803	0.0112
1.95	1.3688	1.38	0.0112
2	1.3686	1.3797	0.0111
2.05	1.3683	1.3794	0.0111
2.1	1.368	1.3791	0.0111
2.15	1.3677	1.3788	0.0111
2.2	1.3674	1.3785	0.0111
2.25	1.3671	1.3782	0.0111
2.3	1.3668	1.3778	0.011
2.35	1.3665	1.3775	0.011
2.4	1.3662	1.3771	0.0109
2.45	1.3658	1.3768	0.011
2.5	1.3655	1.3764	0.0109
2.55	1.3652	1.3761	0.0109
2.6	1.3648	1.3757	0.0109
2.65	1.3645	1.3753	0.0108
2.7	1.3641	1.3749	0.0108
2.75	1.3638	1.3745	0.0107
2.8	1.3634	1.3741	0.0107
2.85	1.363	1.3737	0.0107
2.9	1.3626	1.3733	0.0107
2.95	1.3622	1.3729	0.0107
3	1.3618	1.3724	0.0106
3.05	1.3614	1.372	0.0106
3.1	1.361	1.3716	0.0106
3.15	1.3606	1.3711	0.0105
3.2	1.3602	1.3706	0.0104
3.25	1.3598	1.3702	0.0104
3.3	1.3593	1.3697	0.0104
3.35	1.3589	1.3692	0.0103
3.4	1.3584	1.3687	0.0103
3.45	1.358	1.3682	0.0102
3.5	1.3575	1.3677	0.0102
3.55	1.357	1.3672	0.0102
3.6	1.3566	1.3667	0.0101
3.65	1.3561	1.3662	0.0101
3.7	1.3556	1.3656	0.01
3.75	1.3551	1.3651	0.01
3.8	1.3546	1.3645	0.0099
3.85	1.3541	1.364	0.0099
3.9	1.3535	1.3634	0.0099
3.95	1.353	1.3628	0.0098
4	1.3525	1.3622	0.0097
4.05	1.3519	1.3616	0.0097
4.1	1.3514	1.361	0.0096
4.2	1.3502	1.3598	0.0096
4.25	1.3497	1.3592	0.0095
4.3	1.3491	1.3585	0.0094
4.35	1.3485	1.3579	0.0094
4.4	1.3479	1.3573	0.0094
4.45	1.3473	1.3566	0.0093
4.5	1.3466	1.3559	0.0093
4.55	1.346	1.3552	0.0092
4.6	1.3454	1.3546	0.0092
4.65	1.3448	1.3539	0.0091
4.7	1.3441	1.3532	0.0091
4.75	1.3434	1.3524	0.009
4.8	1.3428	1.3517	0.0089
4.85	1.3421	1.351	0.0089
4.9	1.3414	1.3502	0.0088
4.95	1.3407	1.3495	0.0088
5	1.34	1.3487	0.0087
5.1	1.3386	1.3472	0.0086
5.2	1.3372	1.3456	0.0084
5.3	1.3357	1.344	0.0083
5.4	1.3341	1.3423	0.0082
5.5	1.3326	1.3406	0.008
5.6	1.331	1.3389	0.0079
5.7	1.3293	1.3371	0.0078
5.8	1.3277	1.3353	0.0076
5.9	1.3259	1.3334	0.0075
6	1.3242	1.3315	0.0073
6.1	1.3224	1.3296	0.0072
6.2	1.3206	1.3276	0.007
6.3	1.3187	1.3256	0.0069
6.4	1.3168	1.3235	0.0067

$\lambda(\mu m)$	n_o	n_e	$\Delta n = n_e - n_o$
6.5	1.3148	1.3214	0.0066
6.6	1.3128	1.3192	0.0064
6.7	1.3108	1.317	0.0062
6.8	1.3087	1.3147	0.006
6.9	1.3066	1.3124	0.0058
7	1.3044	1.3101	0.0057
7.1	1.3022	1.3076	0.0054
7.2	1.3	1.3052	0.0052
7.3	1.2976	1.3027	0.0051
7.4	1.2953	1.3001	0.0048
7.5	1.2929	1.2975	0.0046
7.6	1.2904	1.2948	0.0044
7.7	1.2879	1.2921	0.0042
7.8	1.2854	1.2893	0.0039
7.9	1.2827	1.2865	0.0038
8	1.2801	1.2836	0.0035
8.1	1.2774	1.2806	0.0032
8.2	1.2746	1.2776	0.003
8.3	1.2718	1.2745	0.0027
8.4	1.2689	1.2714	0.0025
8.5	1.2659	1.2682	0.0023
8.6	1.2629	1.2649	0.002
8.7	1.2598	1.2615	0.0017
8.8	1.2567	1.2581	0.0014
8.9	1.2535	1.2546	0.0011
9	1.2502	1.2511	0.0009
9.1	1.2469	1.2474	0.0005
9.2	1.2435	1.2437	0.0002
9.3	1.24	1.2399	0.0001
9.4	1.2365	1.2361	0.0004
9.5	1.2328	1.2321	0.0007
9.6	1.2292	1.2281	0.0011
9.7	1.2254	1.224	0.0014
9.8	1.2215	1.2198	0.0017
9.9	1.2176	1.2154	0.0022
10	1.2136	1.2111	0.0025

Appendix B

Azimuth and ellipticity calculation

Despite their apparently obvious character, the equations (4.1.5), (4.1.6) require a certain amount of calculus. As a matter of fact, in order to deduce this expressions we have at first to consider the projections E_x and E_y of the $\vec{E}$ for z=0, by introducing specific phases ϕ_x and ϕ_y:

$$E_x = E_{0x}\cos(\omega t + \phi_x)$$
$$E_y = E_{0y}\cos(\omega t + \phi_y)$$

Now we project this relations on the system (η, ξ) according to the Figure B.1:

$$E_\xi = E_x\cos(\psi) + E_y\sin\psi$$
$$E_\eta = -E_x\sin\psi + E_y\cos\psi$$

Meanwhile, in the rotated system (η, ξ) we have

$$E_\xi = a\cos(\omega t + \phi)$$
$$E_\eta = \pm b\cos(\omega t + \phi)$$

where a is the semi-minor axis of the ellipse, b is the semi-major axis and ϕ is a common phase. From this two last groups of identities we deduce

$$a\cos(\omega t + \phi) = E_x\cos(\psi) + E_y\sin\psi$$

$$\pm b\cos(\omega t + \phi) = -E_x\sin\psi + E_y\cos\psi$$

Substituting now the expression of the field along x and y

$$a\cos(\omega t + \phi) = E_{0x}\cos(\omega t + \phi_x)\cos\psi + E_{0y}\cos(\omega t + \phi_y)\sin\psi$$
$$\pm b\cos(\omega t + \phi) = -E_{0x}\cos(\omega t + \phi_x)\sin\psi + E_{0y}\cos(\omega t + \phi_y)\cos\psi$$

Using trigonometric identities we may develop the first relation and then, by grouping the terms in $\cos(\omega t)$ and $\sin(\omega t)$ we may find

$$a\cos(\phi) = E_{0x}\cos\phi_x\cos\psi + E_{0y}\cos(\phi_y)\sin\psi \tag{B.0.1}$$
$$a\sin(\phi) = E_{0y}\sin\phi_x\cos\psi + E_{0y}\sin(\phi_y)\sin\psi \tag{B.0.2}$$

And in the same way:

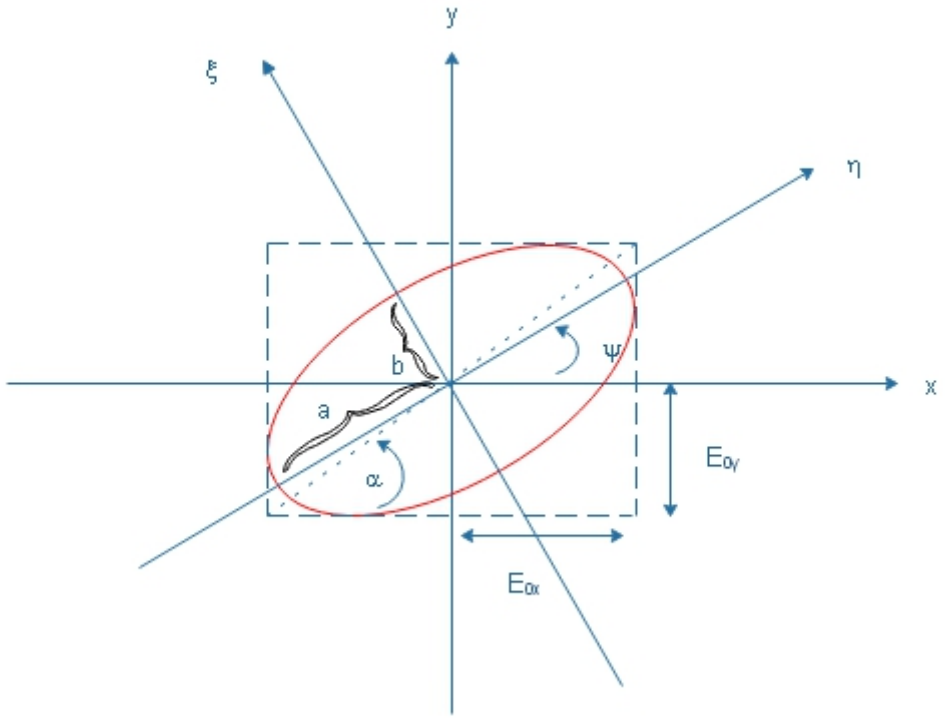


Figure B.1: Principal angles defining the polarization ellipse. A second reference frame (η, ξ) oriented after the axis of the ellipse is also represented

$$\mp b \sin(\phi) = -E_{0x} \cos\phi_x \sin\psi + E_{0y} \cos(\phi_y) \cos\psi \tag{B.0.3}$$

$$\pm b \cos(\phi) = -E_{0x} \sin\phi_x \sin\psi + E_{0y} \sin(\phi_y) \cos\psi \tag{B.0.4}$$

Dividing (B.0.1) by (B.0.4) and (B.0.2) by (B.0.3) we find directly:

$$\frac{E_{0x} \cos\phi_x \cos\psi + E_{0y} \cos(\phi_y) \sin\psi}{-E_{0x} \sin\phi_x \sin\psi + E_{0y} \sin(\phi_y) \cos\psi} = \frac{E_{0y} \sin\phi_x \cos\psi + E_{0y} \sin(\phi_y) \sin\psi}{E_{0x} \cos\phi_x \sin\psi - E_{0y} \cos(\phi_y) \cos\psi}$$

Working on this last relation and making again use of trigonometric identities we will easily find that:

$$(E_{0x}^2 - E_{0y}^2) \sin(2\psi) = 2E_{0x}E_{0y} \cos(2\psi) \cos(\phi_y - \phi_x)$$

and then

$$\tan(2\psi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 - E_{0y}^2} \cos\varepsilon \tag{B.0.5}$$

where $\varepsilon = \phi_y - \phi_x$. Playing again with the relations (B.0.1) - (B.0.4) we may find easily from $(B.0.1) \cdot (B.0.4) - (B.0.2) \cdot (B.0.3)$ that

$$\pm ab = E_{0x}E_{0y}[\sin\phi_y \cos\phi_x - \sin\phi_x \cos\phi_y]$$

which is equivalent to

$$\pm ab = E_{0x}E_{0y} \sin\varepsilon$$

Also, from the same set of equations we may infer that

$$a^2 + b^2 = E_{0x}^2 + E_{0y}^2$$

Taking now the ellipticity as $\tan\chi = \pm b/a$ we may find that:

$$\sin(2\chi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 + E_{0y}^2}\sin\varepsilon \tag{B.0.6}$$

Appendix C

Stokes parameters, ellipticity and azimuth

From the definition of the Stokes parameters as it was stated in the the section 4.3 we saw that:

$$\begin{array}{lcl} I = & \langle I \rangle & = \langle E_{0x}^2 \rangle + \langle E_{0y}^2 \rangle \\ Q = & \langle I_0 \rangle - \langle I_{90} \rangle & = \langle E_{0x}^2 \rangle - \langle E_{0y}^2 \rangle \\ U = & \langle I_{45} \rangle - \langle I_{-45} \rangle & = \langle 2E_{0x}E_{0y}\cos\varepsilon \rangle \\ V = & \langle I_{RCP} \rangle - \langle I_{LCP} \rangle & = \langle 2E_{0x}E_{0y}\sin\varepsilon \rangle \end{array} \tag{C.0.1}$$

Based on the expression for the ellipticity (χ)

$$\sin(2\chi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 + E_{0y}^2}\sin\varepsilon$$

we infer that

$$V = 2E_{0x}E_{0y}\sin\varepsilon$$

and then

$$V = (E_{0x}^2 + E_{0y}^2)\sin(2\chi) \tag{C.0.2}$$

Using now the fact that for a completely polarized wave we need to have

$$Q^2 + U^2 + V^2 = I^2 \tag{C.0.3}$$

and substituting with (C.0.2) and (C.0.1) we obtain:

$$(E_{0x}^2 - E_{0y}^2)^2 + (2E_{0x}E_{0y})^2\cos^2\varepsilon + (E_{0x}^2 + E_{0y}^2)^2\sin^2(2\chi) = (E_{0x}^2 + E_{0y}^2)^2$$

$$\Leftrightarrow \quad (E_{0x}^2 - E_{0y}^2)^2 + (2E_{0x}E_{0y})^2\cos^2\varepsilon + (E_{0x}^2 + E_{0y}^2)^2 - (E_{0x}^2 + E_{0y}^2)^2\cos^2(2\chi) = (E_{0x}^2 + E_{0y}^2)^2$$

$$\Leftrightarrow \quad (E_{0x}^2 - E_{0y}^2)^2 = (E_{0x}^2 + E_{0y}^2)^2\cos^2(2\chi) - (2E_{0x}E_{0y})^2\cos^2\varepsilon$$

However, in the Chapter 4 we saw that the azimuth angle is given by

$$\tan(2\psi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 - E_{0y}^2}\cos\varepsilon$$

Replacing now $2E_{0x}E_{0y}\cos\varepsilon$ by the corresponding term:

$$(E_{0x}^2 - E_{0y}^2)^2 = (E_{0x}^2 + E_{0y}^2)^2\cos^2(2\chi) - (E_{0x}^2 - E_{0y}^2)^2\tan^2(2\psi)$$

$$\begin{aligned} \Leftrightarrow \quad & (E_{0x}^2 - E_{0y}^2)^2(1 + \tan^2(2\psi)) = (E_{0x}^2 + E_{0y}^2)^2 \cos^2(2\chi) \\ \Rightarrow \quad & Q = (E_{0x}^2 - E_{0y}^2) = (E_{0x}^2 + E_{0y}^2)\cos(2\chi)\cos(2\psi) \end{aligned} \tag{C.0.4}$$

And again because

$$\tan(2\psi) = \frac{2E_{0x}E_{0y}}{E_{0x}^2 - E_{0y}^2}\cos\varepsilon$$

we find directly

$$U = 2E_{0x}E_{0y}\cos\varepsilon = (E_{0x}^2 + E_{0y}^2)\cos^2(2\chi)\sin(2\psi) \tag{C.0.5}$$

As a conclusion we may write now

$$\begin{aligned} I &= E_{0x}^2 + E_{0y}^2 \\ Q &= (E_{0x}^2 + E_{0y}^2)\cos(2\chi)\cos(2\psi) \\ U &= (E_{0x}^2 + E_{0y}^2)\cos(2\chi)\sin(2\psi) \\ V &= (E_{0x}^2 + E_{0y}^2)\sin(2\chi) \end{aligned} \tag{C.0.6}$$

Appendix D

Mueller matrix of the spectropolarimeter

The birefringent bloc is characterized by the phase differences $\Delta\phi_1$ and $\Delta\phi_2$, while the analyzer by the angle θ.

$$\frac{1}{2}\begin{bmatrix} 1 & \cos(2\theta)\cos(\Delta\phi_2) & \cos(\Delta\phi_1)\sin(2\theta) - \sin(\Delta\phi_1)\cos(2\theta)\sin(\Delta\phi_2) & \cos(\Delta\phi_1)\cos(2\theta)\sin(\Delta\phi_2) + \sin(\Delta\phi_1)\sin(2\theta) \\ \cos(2\theta) & \cos^2(2\theta)\cos(\Delta\phi_2) & \cos(\Delta\phi_1)\cos(2\theta)\sin(2\theta) - \sin(\Delta\phi_1)\cos^2(2\theta)\sin(\Delta\phi_2) & \cos(\Delta\phi_1)\sin(\Delta\phi_2)\cos^2(2\theta) + \sin(\Delta\phi_1)\sin(2\theta)\cos(2\theta) \\ \sin(2\theta) & \cos(2\theta)\cos(\Delta\phi_2)\sin(2\theta) & \cos(\Delta\phi_1)\sin^2(2\theta) - \sin(\Delta\phi_1)\cos(2\theta)\sin(2\theta)\sin(\Delta\phi_2) & \sin(\Delta\phi_1)\sin^2(2\theta) + \cos(\Delta\phi_1)\text{os}(2\theta)\sin(\Delta\phi_2)\sin(2\theta) \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{D.0.1}$$

Appendix E

List of symbols and abbreviations

LHP	horizontal polarization
LVP	vertical polarization
L+45P	polarization at $+45^{\circ}$
L-45P	polarization at -45°
RCP	right circular polarization
LCP	left circular polarization

Appendix F

Appendix: Useful Mueller matrices

- Linear polarizer making an angle θ with the x axis:

$$M(\theta) = \frac{1}{2}\begin{bmatrix} 1 & \cos(2\theta) & \sin(2\theta) & 0 \\ \cos(2\theta) & \cos^2(2\theta) & \sin(2\theta)\cos(2\theta) & 0 \\ \sin(2\theta) & \sin(2\theta)\cos(2\theta) & \sin^2(2\theta) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Horizontal polarizer

$$M(0) = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- 45° polarizer

$$M(\tfrac{\pi}{4}) = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- $\frac{\pi}{2}$ polarizer

$$M(\tfrac{\pi}{2}) = \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Rotator (of angle θ)

$$R(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Wave plate (ϕ)

$$M(\phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{bmatrix}$$

- Quarter wave plate (QWP)

$$M(\tfrac{\pi}{2}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

- Half wave plate (HWP)

$$M(\pi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

- Rotated wave plate (rotation θ, phase ϕ)

$$M(\theta,\phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \cos\phi \sin^2 2\theta & (\cos\phi - 1)\sin 2\theta \cos 2\theta & \sin\phi \sin 2\theta \\ 0 & (\cos\phi - 1)\sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos\phi \cos^2 2\theta & \sin\phi \cos 2\theta \\ 0 & -\sin\phi \sin 2\theta & -\sin\phi \cos 2\theta & \cos\phi \end{bmatrix}$$

Bibliography

[1] Birefringence at an Isotropic-Uniaxial Interface: Waves, Rays, and Fresnel Coefficients. http://demonstrations.wolfram.com/BirefringenceAtAnIsotropicUniaxialInterfaceWavesRaysAndFresn/. Accessed: 2018-11-21.

[2] Edmund Optics, introduction to Polarization. https://www.edmundoptics.com. Accessed: 2018-12-12.

[3] Eksmaoptics. http://eksmaoptics.com. Accessed: 2018-12-20.

[4] Entry on L-S Coupling of angular momenta. https://chem.libretexts.org. Accessed: 2018-12-12.

[5] Magnesium floride. https://www.korth.de/index.php/162/items/21.html. Accessed: 2018-10-01.

[6] Optical constants of magnesium fluoride. https://refractiveindex.info. Accessed: 2018-09-26.

[7] Wikipedia, entry on Angular momentum operator. https://en.wikipedia.org/wiki/Angular_momentum_operator. Accessed: 2019-1-2.

[8] Wikipedia, entry on Circular Polarization. https://en.wikipedia.org/wiki/Circular_polarization. Accessed: 2018-12-12.

[9] E. Akkermans, P. E. Wolf, and R. Maynard. Coherent backscattering of light by disordered media: analysis of the peak line shape. *Physical review letters*, 56, N. 14, 1986.

[10] J. Bailey, L. Kedziora-Chudczer, and K. Bott. Polarized radiative transfer in planetary atmospheres and the polarization of exoplanets. *arXiv: astro-ph.EP/1807.04443v1*, 2018.

[11] V. Chandrasekharan and H. Damany. Birefringence of sapphire and magnesium fluoride in the vacuum ultraviolet, and retardation plates. *Applied Optics*, 7, 1968.

[12] Edward Collett. *Field Guide to Polarization*. SPIE-The International Society for Optical Engineering, 2005.

[13] Thomas M. Cotter, Michael E. Thomas, and William J.Tropf. *Magnesium Fluoride* (MgF_2). Handbook of optical constants of solids II, 1991.

[14] A. Duncanson and W. H. Stevenson, R. Some properties of Magnesium Fluoride crystallized from the melt. *Proc. Phys.Soc.*, 72, 1958.

[15] W. Bruce Hapke, Robert M. Nelson, and Smythe William D. The opposition effect of the moon: The contribution of the coherent backscatter. *Science*, 260, N. 5107, 1993.

[16] Eugene Hecht. *Optics*. Pearson Education, 2017.

[17] Serge Huard. *Polarisation de la lumière*. Masson, 1993.

[18] Andrew P. Ingersol. Polarization measurements of Mars and Mercury: Rayleigh scattering in the Martian atmosphere. *The Astrophysical Journal*, 1971.

[19] R. Ishikawa and all. Strategy for realizing high-precision VUV Spectro-Polarimeter. *Solar Phys 289*, 2014.

[20] James C. Kemp. Circular polarization of planets. In T. Gehrels, editor, *Planets, Stars and Nebulae studied with photopolarimetry*, chapter 6, pages 607–616. The University of Arizona Press, 1974.

[21] M.F. Kessler and all. The infrared space observatory (ISO). *Space Science Reviews, 61:45-60*, 1992.

[22] David S. Kliger, James W. Lewis, and Randall Cora E. *Polarized light in optics and spectroscopy*. Academic Press, Inc., 1990.

[23] Ludmilla Kolokolova, James Hough, Anny Chantal, and Levasseur Regourd. *Polarimetry of stars and planetary systems*. Cambridge University Press, 2015.

[24] S. V. Lagerros, J. and all. Thermal physics of asteroids II: Polarization of the thermal microwave emission from asteroids. *Astron. Astrophys.*, 315(3),625-632, 1996.

[25] S. V. Lagerros, J. and all. ISOPHOT polarization measurements of the Asteroids (6) Hebe and (9) Metis at 25 μm. *Icarus*, 142,454-463, 1999.

[26] Jean-Louis Leroy. *La polarisation de la lumière et l'observation astronomique*. Edition des archives contemporaines, 2001.

[27] H. Li, H. Refractive index of alkaline earth halides and its wavelength and temperature derivatives. *Journal of Physical and Chemical Reference Data, 9*, 1980.

[28] Pierre Magain. Traitement des données expérimentales. University Lecture, 2018.

[29] Alexander Morozhenko and Vid'machenko Anatoliy. Polarimetry and physics of solar system bodies. In *Photopolarimetry in Remote Sensing*, chapter 6, pages 369–384. Springer Science, 2005.

[30] Coralie Neiner and Jean-Claude Bouret. Pollux: European Study of a UV Spectropolarimeter. *The LUVOIR Interim Report*, 2018.

[31] Coralie Neiner and Jean-Claude Bouret. POLLUX:European Study of a UV Spectropolarimeter. In *The large UV Optical Infrared Surveyor: The LUVOIR Interim Report*, chapter 10, pages 246–268. 2018.

[32] Martin Pertenais. *Spectropolarimetrie stellaire UV et visible depuis l'espace*. These Doctorat, 2016.

[33] E. Raouafi, N. Coronal polarization. *arXiv: astro-ph.SR/1104.0598v1*, 2011.

[34] E. Raouafi, N., K. Solanki S., and T. Wiegelmann. Hanle effect diagnostics of the coronal magnetic field. *arXiv:astro-ph.SR/0801.2202v1*, 2008.

[35] L. Rossi and D. M. Stam. Circular polarization signals of cloudy (exo)planets. *arXiv: astro-ph.EP/1805.0868v1*, 2018.

[36] E. Russell, E. and All. Galileo photopolarimeter/radiometer experiment. *Space Science Reviews, 60: 531-563*, 1992.

[37] C. Simon, Maria and V. Gottschalk, Karin. Waves and rays in uniaxial birefringent crystals. *Optik*, 118,457-470, 2007.

[38] Frans Snik and all. Prototyping for the Spectropolarimeter for Planetary EXploration (SPEX): calibration and sky measurements. *SPIE*, 8160, 2015.

[39] Donguk Song and all. Optical alignment of the High-Precision UV Spectro-Polarimeter (CLASP2). *NASA Technical Reports*, 2018.

[40] William Sparks and all. Compact and robust method for full Stokes spectropolarimetry. *Applied Optics*, 51, 2012.

[41] J.O. Stenflo. Solar magnetic fields as revealed by Stokes polarimetry. *arXiv:astro-ph.SR/1309.5454v1*, 2013.

[42] Jaap Tinbergen. *Astronomical polarimetry*. Cambridge University Press, 2005.

[43] John D. van der Laan and ALL. Evolution of circular and linear polarization in scattering environments. *OSA*, 2015.