

Imperial College of Science, Technology and Medicine
The Blackett Laboratory
Photonics Group

High numerical aperture Müller matrix polarimetry and applications to multiplexed optical data storage

Carlos Alberto Macías Romero

Thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy of Imperial College
of Science, Technology and Medicine.

March 17, 2010

Abstract

Optical data storage has reached a point where it is no longer possible to increase the storage capacity of an optical disk by conventional means. The work presented here, however, proves experimentally the validity of a novel solution to increase the capacity of an optical disk. Polarisation is used to store information multiplexed into shape-birefringent pits. A high numerical aperture scanning confocal microscope capable of measuring the Müller matrix of a sample is built to measure and characterise the polarisation properties of such pits. A number of results found during the development of the apparatus are given. Novel contributions are shown in the areas of imaging, the theory of coherence and polarisation, polarisation imaging, linear algebra, polarisation confocal microscopy, and optical data storage. It is shown that the storage capacity of any existent non-volumetric optical technology can be increased by 7.16 times, at least.

To my family

Acknowledgements

I have climbed mountains above 5000 metres from the sea level. I have climbed cliffs 300 metres high. I have walked tens of kilometres for days. I have camped in adverse conditions. However, writing this thesis has been the greatest challenge I have faced in my life. I honestly felt that I was going to die or to go crazy while trying. I never felt so defeated in my life. I do not think I could have finished writing it without the love, care and understanding of Laura Romero my mom, Sitara Romero my aunt, Gloria Flores my grandma, Priscille Varillon my girlfriend, Carlos Macías my dad, and Peter Török my supervisor. I truly thank them from my soul.

If I look back I can see perfectly the chain of events that took me to do this PhD. It all started with Javier Piñataro, the lab technician that recommended me to Professor Ojeda–Castañeda who put me into contact with Professor Dainty and led me to meet Peter Török. I thank them all for being part of the beginning and particularly Peter for being there all the way. I also want to thank my mates from university Osvaldo, Mary and David for teaching me Physics.

During the process of the PhD I met a large amount of people that contributed somehow to my own development, both in a personal and a professional way. The list is large but I would like to mention those who remained close. Martin, Simon and James from the workshop for materialising my capricious ideas. Peter Munro, Arthur van de Nes and Carl Paterson for sharing their knowledge. My flatmates/officemates/partymates and friends Christoforos, Alexandros (now I understand you dude), Pável, Alex (Cachorro), Matt (George), Pierre, Gung–Hsuan, Ren and Mike.

Finally I want to acknowledge the people in México that contributed to my formation as a scientist and a human being. My chemistry teacher Araceli. My brother Juan and my sisters Anamarina, Barbara, Tony and Victoria. My stepmother Bar-

bara Mouriño. My friends Alvaro, Nabor, Samuel, Luis Enrique, Aranda, Miguel, Daniela, Jose Luis, Pilar and Ricardo. I apologise if I forget anyone.

This research was funded by the Consejo Nacional de Ciencia y Tecnología México (CONACyT) and by the Engineering and Physical Sciences Research Council (EPSRC).

Contents

Abstract	3
Acknowledgements	7
Contents	9
List of figures	13
Introduction	17
1 Imaging	21
1.1 Heuristic description of a radiating field	21
1.2 Ellipse of polarisation	23
1.3 High NA imaging	27
1.4 Scanning confocal imaging	31
1.4.1 Spatial and spatial frequency resolution	32
1.4.2 The longitudinal component of a dipole	39
1.5 Conclusion	46
2 Polarisation and coherence	47
2.1 Fluctuations of light	47
2.2 Degree of polarisation and degree of coherence	50
2.2.1 Degree of temporal and spatial coherence	51
2.2.2 Degree of temporal polarisation	56
2.2.3 Degree of spatial polarisation	60
2.3 Conclusions	64

3	Polarimetry	67
3.1	Müller–Stokes formalism	67
3.1.1	Stokes vector	67
3.1.2	Müller matrix	70
3.2	Müller–Stokes polarimetry	74
3.2.1	Matrix conditioning	76
3.3	Conclusion	82
4	Polarisation imaging	83
4.1	Confocal micropolarimetry	83
4.1.1	Pinhole size dependance	86
4.2	Imaging techniques	91
4.2.1	Stokes imaging	91
4.2.2	Müller imaging	92
4.3	Illumination engineering	97
4.4	Conclusions	99
5	El experimento	101
5.1	Experimental layout	101
5.2	Scanning	103
5.2.1	Pre-shaping	103
5.3	LabVIEW software	105
5.4	Imaging optics	106
5.4.1	Point spread function of the system	107
5.5	Polarisation optics	109
5.5.1	Broadband polarimeter	109
5.5.2	Pockel’s cells	112
5.6	Conclusion	113
6	Calibration	115
6.1	Unveiling the ECM	115
6.1.1	The Heart of the Method	116
6.1.2	Solving a particular case of the Sylvester equation	117

6.2	Calibration of the system	120
6.2.1	Two-step calibration	121
6.2.2	Double pass calibration	122
6.2.3	Three-step calibration	124
6.3	Pinhole polarisation dependance	126
6.4	Conclusion	127
7	MODS	129
	Conclusion	141
	Bibliography	145
A	Jones calculus	153
B	ECM proofs	157

List of Figures

1.1	Polarisation ellipse.	26
1.2	High NA 4-f system	27
1.3	Focusing in a high NA system	29
1.4	Collection of light in a high NA system	30
1.5	Complete confocal imaging system (not to scale)	31
1.6	Spatial frequency response of a high NA imaging system to a dipole emitting horizontally polarised light. Top: HTF obtained from the total field(left), from the x component(centre) and from y (right). Bottom: the respective cross sections, and a comparison with the scalar case(left) .	34
1.7	Spatial frequency response of a high NA imaging system to a dipole emitting circularly polarised light. Top: HTF obtained from the total field(left), from the x component(centre) and from y (right). Bottom: the respective cross sections, and a comparison with the scalar case(left).	35
1.8	Impulse response of a high NA imaging system to a horizontally polarised dipole. Top: DSF obtained from the magnitude of the field(left), from the x component(centre) and from the y component(right). Bottom: the respective cross sections, and a comparison with the scalar case(left). . . .	36
1.9	Coherent two point resolution	37
1.10	Resultant irradiance from the superposition of two Bessel functions with opposite phase. The plot shows cross sections of the observation plane while the distance between these functions is changed.	38
1.11	Field distributions for E_x (top) and E_y (bottom) in the back focal plane of a collecting lens ($NA = 0.966$) arising from orthogonal electric dipoles orientated along the x , y and z coordinate axes respectively (left to right).	40

1.12	Proposed optical setup for determination of the full three-dimensional orientation of an electric dipole.	41
1.13	a) Proposed phase mask b) Rotational misalignment c) Translational misalignment.	42
1.14	a) Extinction ratios as a function of mask misalignment (bottom-left axes) and spectral bandwidth (top-right axes). b) SNR for finite detector sizes. .	43
2.1	Young interferometer	51
2.2	Michelson interferometer	53
3.1	Ellipse of polarisation and the Poincaré sphere.	70
3.2	Block diagram of a polarimeter.	75
4.1	Experimental setup	84
4.2	Polarisation distribution at the back focal plane of a lens with $NA = 1$. . .	85
4.3	Polarisation distribution at the confocal plane.	85
4.4	Extinction ratio for different pinhole sizes resulting from the image of a dipole	87
4.5	Extinction ratio (colour scale) for different polarisation illuminations and for a pinhole of one Airy unit in size.	87
4.6	Degree of purity due to an excited dipole for different pinhole size	88
4.7	Elliptical purity (left) and azimuthal purity (right) for different polarisation illuminations and a pinhole of one Airy unit in diameter.	89
4.8	Polarisation distribution at the pinhole produced by two dipoles.	89
4.9	Elliptical and azimuthal degree of purity when two dipoles are scanned along the x axis and $y = 0$ with different pinholes sizes	90
4.10	Stokes parameters extracted from the Stokes image of two dipoles	91
4.11	Stokes parameters taken from the Stokes image when two dipoles are scanned along the x axis with different pinholes sizes	92
4.12	Müller matrix of two fixed dipoles placed one Airy unit apart.	93
4.13	Distribution of the Lu-Chipman parameters obtained from the Müller matrix due to two fixed dipoles placed one Airy unit apart.	94
4.14	Retardance (yellow) and diattenuator (green) vectors calculated from the Müller matrix of two fixed dipoles.	95

4.15	Müller matrix due to a free dipole.	96
4.16	Lu–Chipman parameters calculated from the Müller matrix due to a free dipole.	97
4.17	Venn diagrams representing equation (4.4).	98
5.1	Experimental setup	102
5.2	Top: Desired waveform and response of the stage. Centre: Shaped waveform. Bottom: comparison between the shaped waveform and the desired waveform	104
5.3	Zernike coefficients of the aberrations present in the objective lens (left), and the corresponding Zernike polynomials (right). H stands for horizontal and V for vertical.	106
5.4	Confocal signal from a low reflecting surface	107
5.5	Confocal signal from a low reflecting surface	108
5.6	Schematic of the broadband polarimeter	109
5.7	The polarimeter	111
5.8	Angle (left) and wavelength (right) dependence of the condition number	111
6.1	Block diagram of Müller matrix polarimetry	116
6.2	Simplified schematic of the confocal polarisation microscope.	120
6.3	Block diagram representing the path of the light in the confocal polarisation microscope built for this research	122
6.4	Parameters from the polar decomposition of the confocal signal produced by a mirror for different pinhole diameters	126
7.1	Confocal image of the tracks in a DVD	130
7.2	MODS Pits	131
7.3	Stokes parameters measured when illuminating the engraved pits with linearly polarised light	132
7.4	Stokes parameters measured when illuminating the engraved pits with circularly polarised light	133

7.5	Linescans of the polarisation images obtained when illuminating the pits with linearly (left column) and circularly (right column) polarised illumination. Each row corresponds to a rows in figure (7.2) and each marker to a pit.	134
7.6	Müller matrix of the MODS pits	135
7.7	Magnitude of the retardance (top left), depolarisation (bottom right), polarisance (bottom left), and diattenuation (top right) induced by the MODS Pits to the state of polarisation of the illumination	136
7.8	Poincaré sphere representation of the polarisance vector (left), diattenuator vector (centre), and fast axis (right) corresponding to each pixel of the image	136
7.9	Stokes parameters measured when illuminating the pits with unpolarised light	137
7.10	Left: Linescans obtained from the image representing the azimuth angle of the polarisance vector. Right: Linescans obtained from the image representing the azimuth angle of the diattenuation vector	138
7.11	Storage capacity as a function of the tracking error when illuminating with circularly polarised light (right), and with unpolarised light (left).	139
7.12	Storage capacity as a function of the windowing error obtained when illuminating with circularly polarised light(blue), and with unpolarised light (red). The vertical axis represents the storage increase, and the horizontal axis denotes half of the length of the window	139

Introduction

Storing data is probably the most important factor for a culture and their knowledge to survive throughout history. Take for example the Egyptian culture whose legacy is largely intact as opposed to the Mayan culture whose records were destroyed during the Spanish conquest. It does not matter if the data is stored in stone, paper, optical disks, magnetic disks or solid state circuits, there is a need for the human race to conserve and transmit their knowledge.

A major breakthrough in the history of data storage occurred during the 1950's. In this decade the compact disk (CD) and the hard drive were invented. It was not until the 80's, however, when the boom of both technologies took place. By then the hard drive could store up to 10 Mb. The CD on the other hand was already storing 650Mb. This led to the thought that light could be the ultimate resource for data storage.

The overwhelming success of the CD was due to the availability of lasers and better objective lenses. The key for its success, however, was the relatively low cost and size compared to the hard drive. Nowadays the story is different. The optical disk industry has reached a point where it is impractical to increase storage capacity by conventional means, i.e. increasing the numerical aperture (NA) of the objective lens and decreasing the wavelength of light. Further decreasing of the wavelength would require the use of CaF lenses which could increase the production costs. The only possible way to increase the NA further is by means of employing an immersion technique[1] which can lead to difficult technical challenges. Assuming the technical difficulties can be solved, it is the price of an optical disk that makes the technology available to most of the world. The cost of a Gb of optical storage is currently 5 pence as opposed to the 2 per Gb paid for magnetic storage. Even if the production costs are maintained at its current level, it might not be possible for optical storage

to survive the technological race against magnetic or solid state storage. That is only if the current storage increase method is not discontinued. A change of paradigm is needed for optical data storage to survive. Only then an affordable solution to store data can exist.

The paradigm proposed to promote the survival of the optical disk is the use of the polarisation of light to multiplex and store information. With this approach it is possible to significantly increase the capacity while maintaining the same storage area. The theory behind polarisation encoding method was developed during and after the SLAM project[2, 3], and a patent has been filed[4]. Experimental work however is needed to corroborate the theory.

The objective of this research is to prove experimentally if polarisation can be used to store information in a high numerical aperture regime and to what extent. The areas explored to reach this objective are polarimetry[5] and scanning confocal microscopy[6].

The work presented in this thesis results from the development process of a polarisation sensitive confocal microscope and the craving for understanding of polarisation and its interaction with matter. The work is both experimental and theoretical. Within the theoretical work it is possible to find sections that follow rigorous procedures and others using intuitive thoughts. Novel contributions given in this thesis are in the areas of imaging, the theory of coherence and polarisation, polarisation imaging, linear algebra, polarisation confocal microscopy, and optical data storage.

Chapter 1 provides the insight necessary to understand light as an electromagnetic field and scanning confocal imaging. The concepts of EM field, and polarisation are explained including the basic concepts regarding imaging from a vectorial point of view. Vectorial ray tracing is introduced, and confocal microscopy is then explained. As part of the contribution of this work a difference between spatial and spectral resolution in confocal imaging is given.

The objective of chapter 2 is to introduce a formulation to understand light from a stochastic point of view. The chapter includes an overview of the time variations in a scalar and vectorial fields. A solution to the ongoing issue of coherence versus polarisation is also proposed here. The chapter is divided into two parts. The first part introduces the electric coherency matrix, the mutual coherency matrix and the

polarisation matrix. The second part describes the concepts of temporal and spatial degree of polarisation, depolarisation, and introduces a novel metric to account for spatial variations in the polarisation.

Chapter 3 reviews the notation used in this work for the measurement of polarisation and introduces the theory of the measurement itself. The first section introduces the Stokes vector, the Poincaré sphere, and the Müller matrix. The second section focuses on division of amplitude polarimetry (DOAP). The latter section also includes an overview of the condition number of a matrix, and a technique to improve it in the context of DOAP.

Chapter 4 is devoted to the combination of high numerical aperture confocal microscopy and polarimetry. It discusses the advantages and constraints of the resulting technique. The role of the confocal detector is reviewed, and the imaging parameters that can be used with this apparatus are introduced. The polarisation point spread function of the system is outlined, and an illumination engineering method is developed to exploit geometrical and physical properties of objects.

Chapter 5 presents the layout of the experimental system. Individual parts of the system are addressed here in three main sections: scanning, imaging, and polarisation optics. The objective lens used in the system is described and characterised by means of interferometry. A measurement of the axial and transverse PSF of the system is shown and analysed. A description of the polarisation state generator (PSG) and the polarisation state analyser (PSA) is given, and the device used for scanning is described together with its calibration.

The aim in chapter 6 is to explain the calibration of a confocal polarisation microscope, in particular the polarisation sensitive part of the apparatus. Three calibration procedures which derive from the Eigenvalue Calibration Method (ECM) are considered here. These are a two-step, a three-step, and the double pass calibration. A measurement of the polarisation properties of the confocal polarisation microscope are shown for different pinhole sizes.

The research culminates in chapter 7 with the answers to the questions that motivated the objective of building a state-of-the-art confocal polarisation microscope. It explains the background and the principle behind multiplexed optical data storage (MODS). Müller matrix images of the prototype and the method to retrieve the

data are given.

The conclusion regarding MODS is that a polarisation sensitive disk could store 7.1 times more information than any of the existing non volumetric optical technologies.

Chapter 1

Imaging

The objective of this chapter is to provide the insight necessary to understand light as an electromagnetic field and scanning confocal imaging. The concepts of field, polarisation, and ellipse of polarisation are explained. Basic concepts and definitions regarding imaging from a vectorial point of view are given. Vectorial ray tracing is introduced to aid the description of high numerical aperture optical system. Scanning confocal microscopy is then explained and a comparison with conventional microscopy is given. The novel contribution of this chapter is the difference between spatial and spectral resolution in confocal imaging. Also as a novel contribution at the end of the chapter a method to retrieve the longitudinal component of the field is given.

Most of the concepts introduced in this chapter serve as a basis for a more in-depth understanding. Subsequent chapters address more in detail concepts such as scanning confocal imaging and polarisation.

1.1 Heuristic description of a radiating field

To understand light as an electromagnetic wave take for example an oscillating dipole[7] in vacuum. The oscillations produced by the particle correspond to a wave radiating in space. Characterised by the strength and direction of the dipole, these waves can be represented by two vector valued functions $\vec{B}$ and $\vec{E}$ termed the magnetic and the electric vector, respectively. These functions are also known as field quantities since both can be defined throughout space. A dipole is composed

of a negative and a positive electric pole, which together create an electric potential difference. This means that when the poles are moved apart or close to each other a free electric current density $\vec{J}$ is created. The electric current in turn induces a circulation of the magnetic field, i.e. $\nabla \times \vec{B}$, according to

$$\nabla \times \vec{B} = \vec{J} + \epsilon_0 \frac{\partial \vec{E}}{\partial t}, \quad (1.1)$$

where ϵ_0 denotes the vacuum permeability¹. If now the movement of the poles is changing in time, i.e. $\partial \vec{E} / \partial t \neq 0$, the magnetic field starts to vibrate inducing an oscillating electric field described by

$$\nabla \times \vec{E} = -\mu_0 \frac{\partial \vec{B}}{\partial t}, \quad (1.2)$$

where μ_0 represents the vacuum permittivity². The vibrations of the electric field create in turn a vibrating magnetic field that oscillates in time. The oscillation of the magnetic field induces an oscillation of the electric field, and the process repeats itself creating an electromagnetic wave radiating in $(\mathbf{r}; t) = (x, y, z; t)$. Any point $P(\mathbf{r}, t)$ in space away from the source has in consequence an associated electric and a corresponding magnetic field changing in time or changing in space for a specific time.

The direction in which the energy of the electromagnetic field propagates is known as energy flow. This flow can be described in a dynamic field by the Poynting vector[8] defined as

$$\vec{s} = \vec{E} \times \vec{B}. \quad (1.3)$$

Here $\times$ denotes cross-product and c is the speed of light. A particular case where the energy flow coincides with the direction of propagation of a transversal field is in free space. It is also in this case where it is convenient to introduce the concepts of polarisation and the ellipse of polarisation.

¹ $\epsilon_0 = 8.8542 \times 10^{-7} \text{ F/m}$

² $\mu_0 = 4\pi \times 10^{-7} \text{ H/m}$

1.2 Ellipse of polarisation

Polarisation is a property of light on which the work presented in this thesis is based. This property can be defined as the trajectory drawn in space by the end point of the electric vector field. To describe polarisation however, it is convenient to introduce a pictorial representation known as the ellipse of polarisation. In this representation the oscillations of the field are projected to a plane perpendicular to the direction of propagation.

The ellipse of polarisation is introduced next using the solutions of the vectorial wave equations. The electric and magnetic wave equations are first introduced using the assumptions of free space propagation. The equivalence of the solutions of these equations to describe polarisation in free space is given, and then the ellipse of polarisation is defined.

The vectorial wave equations are derived using equations (1.1), (1.2),

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}, \quad (1.4)$$

and

$$\nabla \cdot \vec{B} = 0 \quad (1.5)$$

when both the total charge density ρ and $\vec{J}$ equal zero. The resulting equations are

$$\nabla^2 \vec{E} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0, \quad (1.6)$$

$$\nabla^2 \vec{B} - \frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2} = 0. \quad (1.7)$$

The most fundamental solutions to equations (1.6) and (1.7) are transverse plane waves whose x, y, z components are in general

$$E_i = a_i \cos[\varphi_i(\mathbf{r}) - \omega t], \quad (1.8a)$$

$$B_i = b_i \cos[\psi_i(\mathbf{r}) - \omega t]. \quad (1.8b)$$

Each component represents the i th projection of the vibrations of the field in the

respective $i = x, y, z$ axis. The real numbers a_i and b_i represent the amplitude of each component. The frequency of the vibrations of the light known as angular frequency is denoted by ω . The light is referred to as monochromatic when it contains a single frequency. When a small range of frequencies is present the light is termed quasimonochromatic. In this case ω is given by the centre or mean average frequency and denoted instead by $\bar{\omega}$. The light is termed polychromatic when it contains a large range of frequencies. In this thesis however, the treatment is assumed to be with quasimonochromatic light. The spatially dependant function $\varphi_i(\mathbf{r})$ represents the initial conditions of the field that define the geometry and direction of its wavefront³. This function can be written as

$$\varphi_i(\mathbf{r}) = k \vec{n}(\mathbf{r}) \cdot \mathbf{r} + \delta_i(\mathbf{r}) \quad (1.9)$$

for $i = x, y, z$, where $k = 2\pi/\lambda$ is the number of oscillations per period, or equivalently the number of wavelengths λ in a period. This number k is known as the wavenumber, and $\vec{k} = k \vec{n}(\mathbf{r})$ is termed the wavevector. Here $\vec{n}(\mathbf{r})$ is a vector parallel to the direction of propagation, i.e. normal to the wavefront, and $\delta_i(\mathbf{r})$ defines the relative phase between the components. Note that position vectors are shown in lowercase and boldface, and vector functions are denoted by an arrow on top of the corresponding symbol. This is to distinguish vector functions from vectors.

In the case of free space propagation, using equation (1.8a) is equivalent to using equation (1.8b) when describing polarisation since $\vec{n}(\mathbf{r}) \cdot \vec{a} = 0$ and $\vec{n}(\mathbf{r}) \cdot \vec{b} = 0$. This means that both $\vec{E}$ and $\vec{B}$ are perpendicular to the direction of propagation $\vec{n}$. It follows in addition from equations (1.1) and (1.2) that $\vec{E}$ and $\vec{B}$ are orthogonal to each other given that

$$\vec{B} = \vec{n}(\mathbf{r}) \times \vec{E}. \quad (1.10)$$

Consequently both vectors are suitable for defining the polarisation of the field in free space. The electric field however is the one commonly measured and hence used.

Note that equation (1.9) represents the relation between the scalar and the vectorial nature of the light, i.e. between wavefront and polarisation. According to this

³Surface given by the collection of all points with position vector $\mathbf{r}$ oscillating at equal phase[9]

equation the vibrations of the electric field in free space are always tangential to the wavefront. To then describe uniquely the polarisation of a field using the ellipse of polarisation, the vibrations must be projected always to a plane perpendicular to the direction of propagation. A simple case is for example an homogeneously polarised plane wave. In this case $\vec{k} \cdot \mathbf{r}$ is constant, and it is only necessary to know the direction of propagation from one point in the field to extrapolate the projection.

Consider a polarised field formed of a single plane wave to introduce the ellipse of polarisation. Equation (1.8a) is rewritten for convenience in complex notation as

$$E_i = A_i e^{i\delta_i} \quad (1.11)$$

where $A_i = a_i e^{i(\vec{k} \cdot \mathbf{r} - \omega t)}$ for $i = x, y, z$. The ellipse of polarisation can then be defined using the ellipsometric parameters ϵ and θ given by [10]

$$\sin(2\epsilon) = \frac{2A_x A_y \sin(\varphi_y - \varphi_x)}{A_x^2 + A_y^2}, \quad (1.12)$$

where $-\pi/4 \leq \epsilon \leq \pi/4$, and

$$\tan(2\theta) = \frac{2A_x A_y \cos(\varphi_y - \varphi_x)}{A_x^2 - A_y^2} \quad (1.13)$$

with $-\pi/2 \leq \theta \leq \pi/2$. An arbitrary representation of the ellipse of polarisation is given in figure (1.1). As presented in this figure θ represents the azimuth angle of the polarisation. That is to say the angle between the major axis of the ellipse and a reference axis which in this case is the horizontal axis. The ellipticity of the ellipse is given by ϵ . When $\epsilon = 0$ the light is linearly polarised, and when $\epsilon = \pi/4$ it is circularly polarised. In the case of the intermediate values the light is said to be elliptically polarised.

Using the ellipse of polarisation to describe the polarisation state of a complicated field can be misleading. Particularly in the case of a 3D field or when the field contains a transversally inhomogeneous polarisation distribution. It is important to understand that the oscillations are occurring in a point of the electric field and not necessarily in an area. The plane wave representation is only to aid the understanding. The concept of the ellipse of polarisation can be generalised to any given

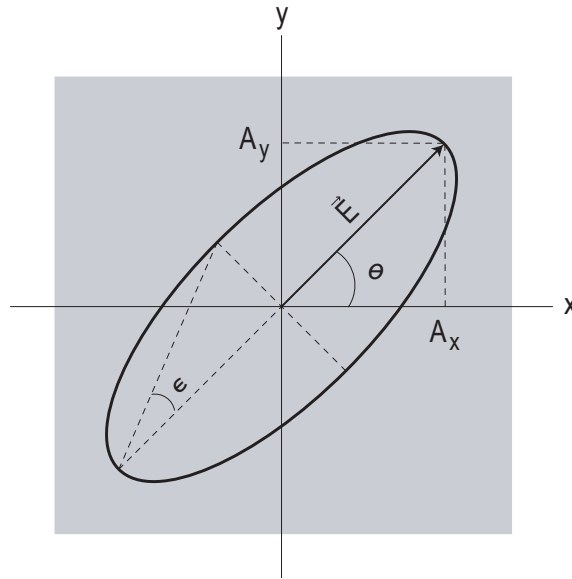


Figure 1.1: Polarisation ellipse.

wavefront using for example the notion of angular spectrum decomposition⁴[11]. A general expression however of the electric field can be simply written as

$$\vec{E} = \begin{pmatrix} A_x e^{i\delta_x} \\ A_y e^{i\delta_y} \\ A_z e^{i\delta_z} \end{pmatrix}, \quad (1.14)$$

where any spatial dependance is implicit, and $A_i = a_i(\mathbf{r}) \exp^{i\varphi_i(\mathbf{r}) - i\omega t}$ for $i = x, y, z$.

For example a widely used field throughout this thesis is the far field produced by an electric dipole given by[9]

$$\vec{E}^d(\mathbf{r}) = \frac{1}{r^3} \mathbf{r} \times (\mathbf{r} \times \vec{p}), \quad (1.15)$$

where $\vec{p} = (p_x, p_y, p_z)$ is a vector with the moment of the dipole as components. Such a field is useful in this thesis because of the applications in imaging systems of high numerical aperture.

⁴The components of a field are decomposed in a sum of weighted plane waves with a given polarisation state

1.3 High NA imaging

The objective of having introduced light as an EM field is to understand high numerical aperture (NA) systems. In these systems where the light is focused or collected at large angles, it is necessary to consider the vectorial nature of the light[9]. The explicit definition of NA is given by

$$NA = n \sin \alpha, \quad (1.16)$$

where α is the limiting convergence semiangle of the light, and n is the refractive index in which the lens is embedded.

Scalar theory can be used to describe systems of low NA, but when the NA is large it is necessary to take into account the polarisation of the light. In high NA systems the transmittance of the lenses depends on the state of polarisation of the light. Rigorous electromagnetic theory is commonly used to model for example the propagation of a field through a high NA imaging system. In this research however, it is more convenient to use a method termed vectorial ray tracing. This is because there is no need to calculate the local interactions within an object with the light. The method is introduced here with the aim of providing the tools to understand later basic high numerical aperture (NA) imaging.

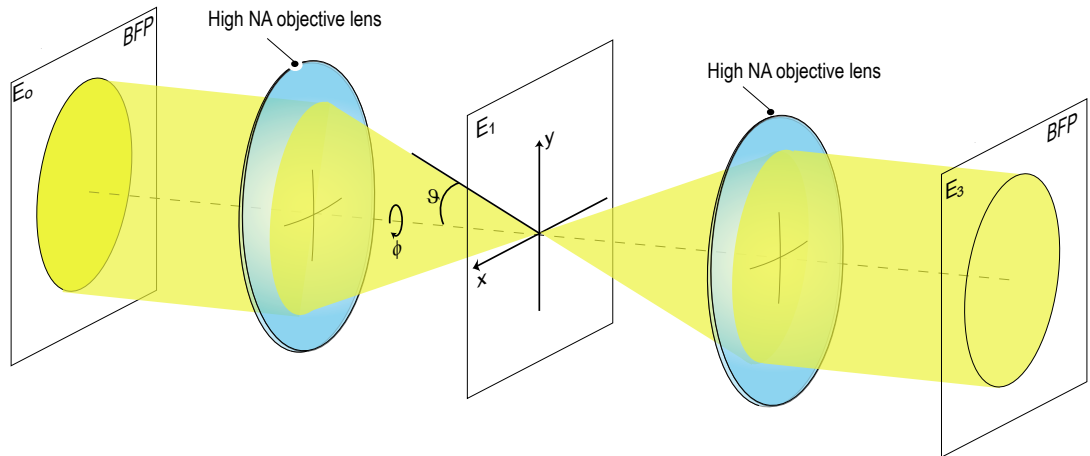


Figure 1.2: High NA 4-f system

Vectorial ray tracing is a method based on coordinate transformations and the use of Fresnel coefficients. To introduce this method consider the 4-f system depicted

in figure (1.2) by parts. In the first part a collimated and polarised beam is focused by a high NA lens. In the second part the diverging field is collected also by a high NA lens.

To introduce the first part of vectorial ray tracing let $\vec{E}_0$ be a collimated and polarised beam incident onto a lens. The field at the focal plane is given by[12]

$$\vec{E}_1 = \mathbf{R}(-\phi)\mathbf{F}_t\mathbf{L}(\vartheta)\mathbf{R}(\phi)\vec{E}_0, \quad (1.17)$$

where

$$\mathbf{R}(\phi) = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.18)$$

$$\mathbf{L}(\vartheta) = \begin{pmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{pmatrix}, \quad (1.19)$$

and,

$$\mathbf{F}_t = \begin{pmatrix} t_{\perp} & 0 & 0 \\ 0 & t_{\parallel} & 0 \\ 0 & 0 & t_{\perp} \end{pmatrix}. \quad (1.20)$$

As depicted in figure (1.3) ϕ and ϑ define a spherical coordinate system⁵. The Fresnel transmission coefficients t of the surface are given as usual by[13]

$$t_{\perp} = \frac{2 \sin \vartheta_t \cos \vartheta_i}{\sin(\vartheta_i + \vartheta_t)} \quad (1.21a)$$

$$t_{\parallel} = \frac{2 \sin \vartheta_t \cos \vartheta_i}{\sin(\vartheta_i + \vartheta_t) \cos(\vartheta_i - \vartheta_t)}, \quad (1.21b)$$

where $\parallel$ refers to the parallel and $\perp$ to the perpendicular component of the field in respect to the surface in which the field is incident. Here ϑ_i is the angle of incidence and ϑ_t is the angle of transmission. Both angles are given with respect to the normal of the surface of incidence. It is possible to relate these angles using Snell's

⁵Normally the convention is to use θ instead of ϑ in this case. The change is due to the convention of using θ to denote a polarisation parameter later introduced and used extensively throughout this work.

law $n_i \sin \vartheta_i = n_t \sin \vartheta_t$. The refractive index of the medium of incidence is denoted by n_i , and n_t is the refractive index of the medium of transmission.

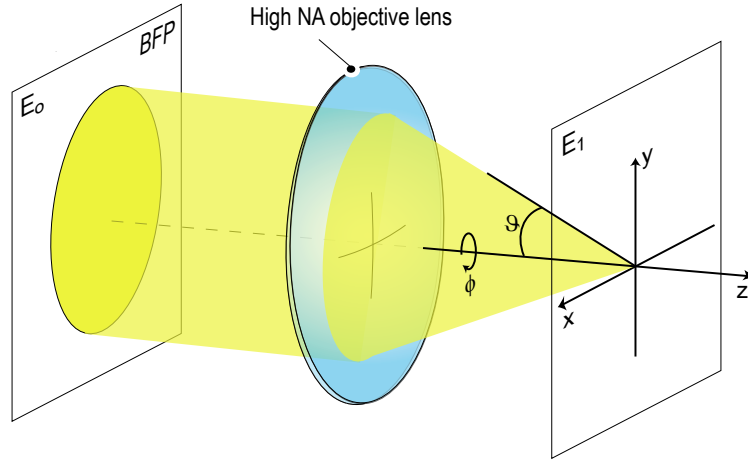


Figure 1.3: Focusing in a high NA system

The matrices in equations (1.18), (1.19), and (1.20) are known as the generalised Jones matrices. The matrix $\mathbf{R}(\phi)$ in equation (1.17) represents a rotation of the coordinate system of the field with the z axis as pivot. The rotation aligns the frame of reference of the vibrations of electric field to the meridional planes⁶ of the system. If the input beam is for example vertically polarised, after the rotation the beam appears to be radially polarised. It is then possible to apply the transformation $\mathbf{L}(\vartheta)$ induced by the lens and the Fresnel coefficients $\mathbf{F}$. In this case $\mathbf{L}(\vartheta)$ focuses the light. The initial coordinate system is then restored by applying $\mathbf{R}(-\phi)$.

The field in the focus interacts with a given object which re-emits a possibly modified field. The interaction can be described for example by the Green's tensor method, where the object is modelled as a three dimensional collection of dipoles with a given refractive index. Roughly speaking, in this method the field produced by each dipole due to the incident field and the neighbouring dipoles is calculated at a given surface, the outcome is a new electromagnetic field $\vec{E}_2$. A detailed description of the Green's tensor method falls beyond the scope of this thesis, but a good review can be found in [14]. The complex amplitude, state of polarisation and state of coherence of the re-emitted field thus depend on the geometrical configuration of the object and refractive indices.

⁶Any plane along the optical axis which contains a ray

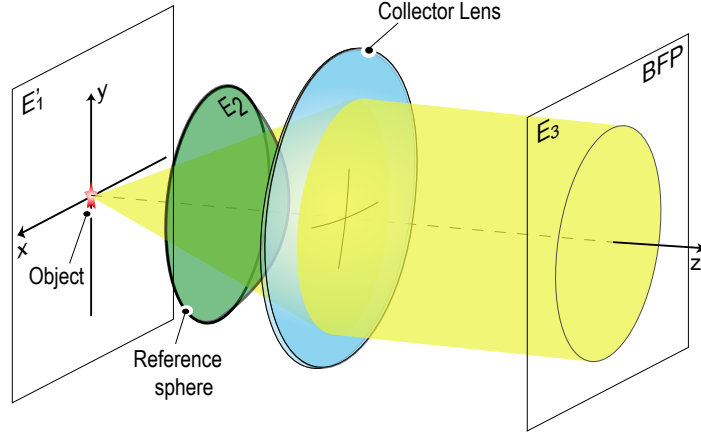


Figure 1.4: Collection of light in a high NA system

In the second part of vectorial ray tracing the re-emitted field $\vec{E}_2$ is collected by an objective lens. The field at the back focal plane (BFP) of the lens is given by[15]

$$\vec{E}_3 = \mathbf{R}(-\phi)\mathbf{L}(-\vartheta)\mathbf{F}_t\mathbf{R}(\phi)\vec{E}_2. \quad (1.22)$$

Note that the collector lens and the high NA lens are assumed to be identical.

To summarise, a polarised plane wave is focused upon an object using a high NA lens, the scattered light from the object is then collected and collimated by an identical high NA lens. The output field from the imaging system can thus be given by

$$\vec{E}_3 = \mathbf{R}(-\phi)\mathbf{L}(-\vartheta)\mathbf{F}_t\mathbf{R}(\phi)\mathbf{N}\mathbf{R}(-\phi)\mathbf{L}(\vartheta)\mathbf{F}_t\mathbf{R}(\phi)\vec{E}_0, \quad (1.23)$$

where $\vec{E}_0$ is the field input to the system, and $\mathbf{N}$ is a matrix describing an object assumed to be non depolarising.

At this point different detection methods can be used such as interferometry[16], polarisation imaging[17, 18] or scanning confocal imaging[19] among others. This research is based in fact on the combination of polarisation imaging and scanning confocal imaging. The latter is introduced next and chapter 4 deals with polarisation imaging.

1.4 Scanning confocal imaging

Confocal imaging[6, 15, 20] is a tomographic technique⁷ where an image is constructed point by point by scanning. As shown in figure (1.5) the particular characteristic of a scanning confocal microscope is that the light is emitted by a point source and detected by a point detector. In this system the plane of the source and the detection plane are conjugate planes to the focal plane of the high NA lens.

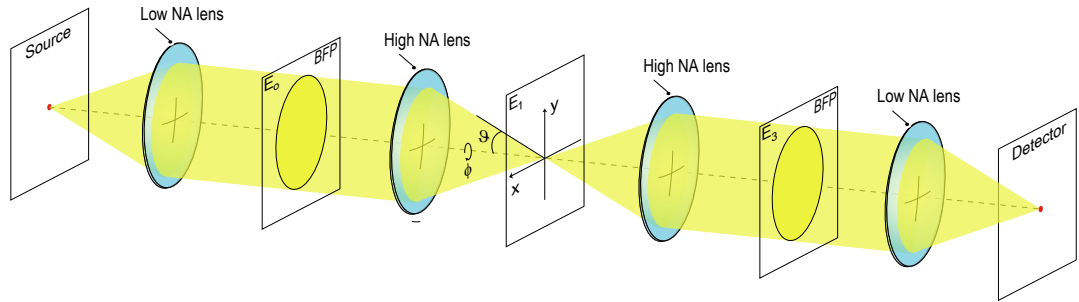


Figure 1.5: Complete confocal imaging system (not to scale)

The device used for detection in confocal imaging can be chosen depending on the experimental requirements. It is possible to use for example a photodetector after a pinhole or after an optical fibre. Both devices act as a three dimensional spatial filter, where the portion of the light measured depends on the size of the detector. In the case of a small pinhole, the difference between these devices is their polarisation maintaining properties. The pinhole acts as an antenna coupling and emitting particular polarisation states. The optical fibre also couples particular polarisation states, but in general the states are not maintained after delivery. Even if the fibre is calibrated, the temperature and stress dependence complicates the usage in polarisation sensitive applications. For the same reason the detection device chosen in this research is a pinhole followed by a photodetector.

In the type of confocal microscope introduced here (point detector) the detection is a coherent process[12] and hence sensitive to the phase. In contrast with the microscopes using an extended detector, the detector integrates incoherently the field. During coherent detection different parts of the field are added coherently by

⁷Imaging by sectioning

the lens and then integrated by the photodetector. As a result the device takes into account the interference effects happening at the detection plane. Mathematically the coherent detection is expressed as

$$I_{confocal} = \left| \int_0^{2\pi} \int_0^R \vec{E}_4(r', \phi') S_D r' dr' d\phi' \right|^2, \quad (1.24)$$

and the detection using an extended detector is given by

$$I_{conventional} = \int_0^{2\pi} \int_0^R \left| \vec{E}_4(r', \phi') \right|^2 S_D r' dr' d\phi', \quad (1.25)$$

where r and ϕ are the dimensions of the detector, S_D represents the sensitivity of the detector and $\vec{E}_4(r', \phi')$ represents the field at the detection plane. Note how in the first case the field is first integrated and then the phase information disappears. The phase plays an important role as opposed to the second case where the phase is not used. The interactions at the detector are in consequence of great importance when using a phase sensitive confocal microscope. As seen next the concepts of spatial frequency and spatial resolution are for example no longer equivalent. The contribution of the longitudinal component to the field also disappears during confocal detection.

1.4.1 Spatial and spatial frequency resolution

Resolution refers here to the ability of an optical system to distinguish in an image different features of an object. Within resolution it is possible to define spatial resolution and spatial frequency resolution. There is no need to distinguish between these two types of resolution in wide-field microscopy. In scanning confocal microscopy, however, this distinction is one of the main properties of this imaging technique.

In scanning confocal imaging the frequency content in an image depends on the NA as in wide-field imaging. The capability however of distinguishing different features within the object changes. In scanning confocal imaging the spatial resolution depends on the geometrical properties of the object and on optical properties such as birefringence. This is due to the fact that scanning confocal microscopes

are coherent imaging systems[11]. Take for example two objects spaced one Airy unit. These objects can be resolved if an optical path difference corresponding to a phase of π exists among them. In fact the objects can still be resolved if the distance between them is less than one Airy unit. On the other hand, the features in the objects corresponding to high spatial frequencies are lost due the low-pass filtering nature of the lens. This relation is shown next using the transfer function of a confocal imaging system and the Rayleigh criterion.

Spatial frequency resolution refers here to the ability of a system to image the spatial frequency content of an object[21]. A measure of this ability is given for example by the Amplitude Transfer Function (ATF). This is a function that describes the frequency response of a low NA diffraction-limited coherent system. The analogous function more suitable for high NA systems is known as the 3D Coherent Transfer function (CTF)[22]. Both ATF and CTF describe the frequency response of the system using a single function. It is known however that the irradiance distribution in the focal plane of a high NA lens depends on the polarisation of the light, e.g. the geometry of the distribution formed when focusing circularly polarised light is different from focusing linearly polarised light[9]. This means that the spatial frequency resolution should depend on the polarisation state of the illumination. The High NA Transfer Function (HTF) introduced in this work is derived analogously to the ATF.

The ATF of a scanning confocal imaging system can be obtained by convolving the pupil function⁸ of the objective lens and the pupil function of the collector lens[6]. For simplicity assume however that the pupil function of both lenses have the same function $A(\mathbf{r})$. If $A(\mathbf{r})$ is known, the ATF can then be described by

$$\text{ATF}(\rho) = \left| \int A(\mathbf{r} - \rho) A(\mathbf{r}) d\mathbf{r} \right|. \quad (1.26)$$

In high NA imaging however, the effective pupil function depends on the polarisation state used for illumination. To then find the HTF of the confocal system, it is thus necessary to first define a polarisation state, find the pupil function of the lens, and then obtain the HTF with the autoconvolution of the resulting pupil function. Note

⁸Function that describes the transmittance of a lens

however that due to the vectorial nature of the problem each component of the field has a pupil function. The function $A(\mathbf{r})$ is thus written as $\vec{A}(\mathbf{r}) = [A_x(\mathbf{r}), A_y(\mathbf{r})]$

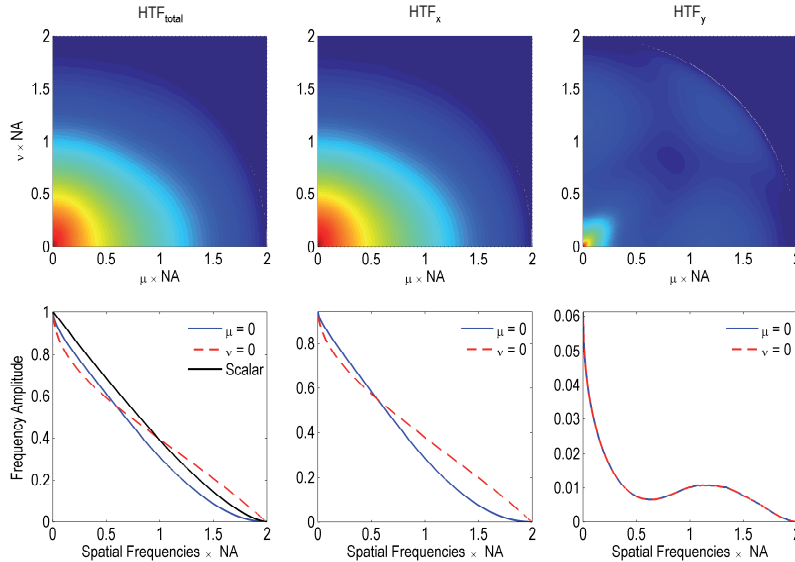


Figure 1.6: Spatial frequency response of a high NA imaging system to a dipole emitting horizontally polarised light. Top: HTF obtained from the total field(left), from the x component(centre) and from y (right). Bottom: the respective cross sections, and a comparison with the scalar case(left)

In this work the collimated far field produced by an electric dipole is considered to be the pupil function of the lens. To define $\vec{A}(\mathbf{r})$, the field is first found in the far field using equation (1.15), and then equation (1.22) is used to describe the collimation property of the lens. The total HTF then follows from the autoconvolution of $|A_x(\mathbf{r})| + |A_y(\mathbf{r})|$. The HTF of the vertical component is given by the autoconvolution of $|A_y(\mathbf{r})|$, and the HTF of the horizontal component is obtained with the autoconvolution of $|A_x(\mathbf{r})|$.

Figure (1.6) shows the HTF of a scanning confocal microscope when using vertically polarised illumination. The top row in this figure shows the total HTF of the field in the left, the centre plot depicts the HTF of the horizontal component, and the HTF of the vertical component is shown in the right hand side plot. The bottom row depicts the horizontal and vertical profiles of the respective HTFs, and a comparison between the total HTF and the scalar result. It can be seen that even though the cutoff frequency is the same in both components, the high frequencies are attenuated more in the vertical direction. In the case however of using circularly

polarised illumination the attenuation is symmetric as shown in figure (1.7).

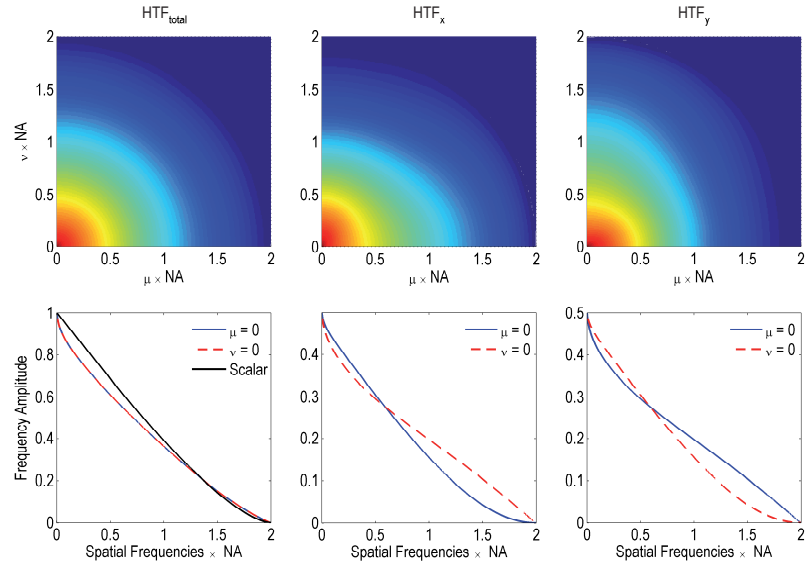


Figure 1.7: Spatial frequency response of a high NA imaging system to a dipole emitting circularly polarised light. Top: HTF obtained from the total field(left), from the x component(center) and from y (right). Bottom: the respective cross sections, and a comparison with the scalar case(left).

The Fourier transform of the HTF is termed as the Dipole Spread Function (DSF) of the system. Analogously to the Point Spread Function (PSF), this function represents the response of the system to a dipole field. Figure (1.8) shows for example the DSF of the confocal system when using vertically polarised illumination. The top row in this figure is obtained by Fourier transforming the equivalent columns from figure (1.6). The bottom row depicts the horizontal and vertical profiles of the respective DSFs.

In systems that use incoherent detection the spatial resolution follows from the PSF of the system. In analogy with low NA systems, this is because spatial resolution is usually defined by the Rayleigh criterion[21] which in turn is related to the PSF. The criterion states that two points can be distinguished if they are placed at a distance greater or equal than

$$\sigma = \frac{1.22\lambda}{\text{NA}}, \quad (1.27)$$

where the coefficient 1.22 is obtained from the first zero of a Bessel function, i.e. the

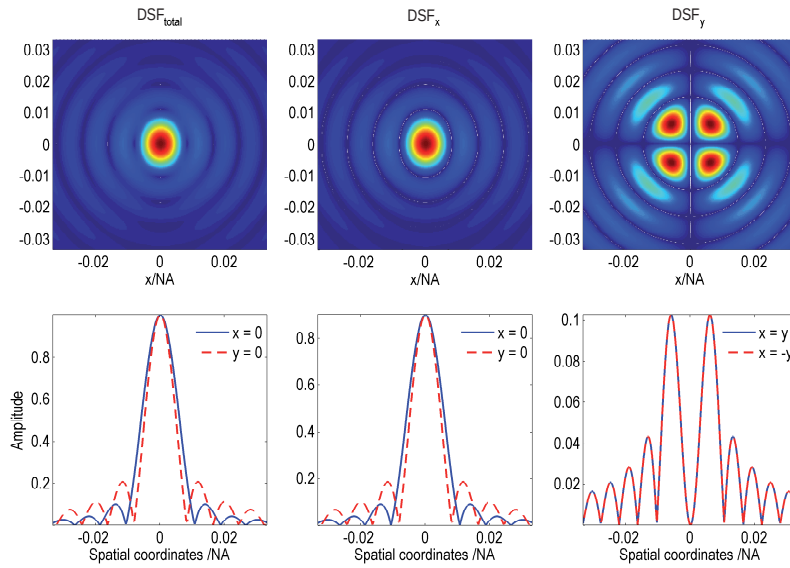


Figure 1.8: Impulse response of a high NA imaging system to a horizontally polarised dipole. Top: DSF obtained from the magnitude of the field(left), from the x component(center) and from the y component(right). Bottom: the respective cross sections, and a comparison with the scalar case(left).

PSF of a lens.

In the coherent detection the spatial resolution depends on the phase difference between the objects to be resolved. To show this dependance consider two dipoles located at $\mathbf{r} = -\mathbf{d}$ and $\mathbf{r} = \mathbf{d}$. The resulting irradiance from the coherent addition at the image plane can be written as

$$I = \left| N\vec{E} \star \delta(\mathbf{r} - \mathbf{d}) + \vec{E} \star \delta(\mathbf{r} + \mathbf{d}) \right|^2, \quad (1.28)$$

where the field $\vec{E}$ of both dipoles is assumed to be the same, and $N = e^{i\phi_N}$ represents a possible phase difference. The term $\vec{E} \star \delta(\mathbf{r} \pm \mathbf{d})$ represents the convolution of $\vec{E}$ with a Dirac delta[23] centred at $\pm \mathbf{d} = (x_d, y_d, z_d)$. Expanding equation (1.28) yields

$$I = \overbrace{\left| \vec{E} \star \delta(\mathbf{r} - \mathbf{d}) \right|^2 + \left| \vec{E} \star \delta(\mathbf{r} + \mathbf{d}) \right|^2}^{\text{Incoherent}} + \underbrace{N\vec{E} \star \delta(\mathbf{r} - \mathbf{d}) \left[\vec{E} \star \delta(\mathbf{r} + \mathbf{d}) \right]^+ + \vec{E} \star \delta(\mathbf{r} + \mathbf{d}) \left[N\vec{E} \star \delta(\mathbf{r} - \mathbf{d}) \right]^+}_{\text{Interference}} \quad (1.29)$$

where $^+$ denotes transpose conjugate. Note that this expression can be divided in two parts. Two of the terms represent the incoherent addition of the fields and the other two represent an interference effect. Also note that a coherent system is equivalent to an incoherent one when

$$N\vec{E} \star \delta(\mathbf{r} - \mathbf{d}) \left[\vec{E} \star \delta(\mathbf{r} + \mathbf{d}) \right]^+ + \vec{E} \star \delta(\mathbf{r} + \mathbf{d}) \left[N\vec{E} \star \delta(\mathbf{r} - \mathbf{d}) \right]^+ = 0, \quad (1.30a)$$

or similarly[24] when

$$2\vec{E}^+(\mathbf{r} - \mathbf{d})\vec{E}(\mathbf{r} + \mathbf{d}) \cos(\phi_N) = 0. \quad (1.31)$$

The phase shift thus needed for a coherent system to produce an image as an incoherent one is

$$\phi_N = \frac{\pi}{2}. \quad (1.32)$$

Figure (1.9) shows three cases computed using equation (1.28). The left plot in this figure represents the case when the dipoles are in phase, the centre plot is created with the $\pi/2$ phase shift, and depicts the Rayleigh criteria, and the right hand side plot represents the case when the dipoles have opposite phase, i.e. π .

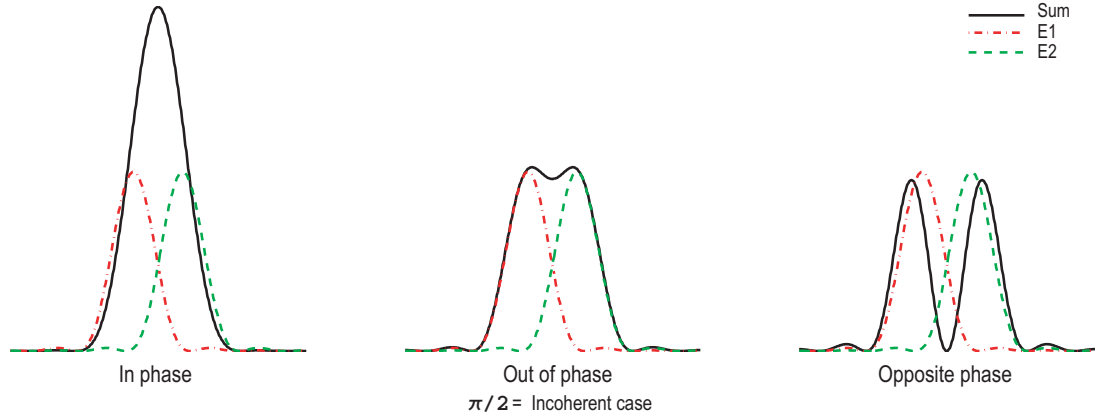


Figure 1.9: Coherent two point resolution

When comparing incoherent and coherent detection, it is possible to say that the latter has lower resolution when the phase difference lies in the range

$$0 < \phi_N < \frac{\pi}{2}. \quad (1.33)$$

To investigate the rest of the range it is convenient to replace $\vec{E}$ by a scalar function $U(\mathbf{r}')$. Take for example[21]

$$U(\mathbf{r}') = \frac{J_1(|\mathbf{r}'|)}{|\mathbf{r}'|}, \quad (1.34)$$

where $J_1(|\mathbf{r}'|)$ is the Bessel function of first order and first kind and $\mathbf{r}' = (x, y, 0)$. Substituting equations (1.32) and (1.34) in (1.28) yields

$$I = \left| \frac{J_1(|\mathbf{d} + \mathbf{r}'|)}{|\mathbf{d} + \mathbf{r}'|} - \frac{J_1(|\mathbf{d} - \mathbf{r}'|)}{|\mathbf{d} - \mathbf{r}'|} \right|^2. \quad (1.35)$$

Note that equation (1.35) and equation (1.28) are equivalent when in the latter $\phi_N = \pi$. Also note from equation (1.35) that the destructive interference occurs at $\mathbf{r} = 0$ regardless of the distance $\mathbf{d}$ between the sources. As shown in figure (1.10) however, the peak value of I certainly decreases when $\mathbf{d}$ is decreased. This figure shows the measured irradiances along the centre of the distribution (right axis) in function of the distance between the sources (left axis). In theory it is possible to differentiate between both distributions despite the value of $\mathbf{d}$. In practice the noise in the system limits the spatial resolution of the system.

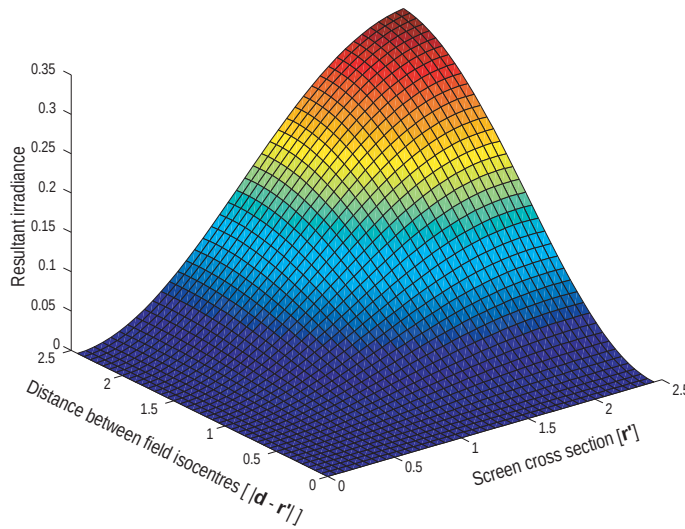


Figure 1.10: Resultant irradiance from the superposition of two Bessel functions with opposite phase. The plot shows cross sections of the observation plane while the distance between these functions is changed.

It is possible to conclude that the spatial resolution of a scanning confocal microscope depends on the optical properties of the object to be imaged. The resolution

in the spatial frequency domain depends on both the optical system and on the state of polarisation emitted by the object to be images. A different consequence but still feature of coherent detection is the cancellation of the longitudinal component of the field. The next section expands more on the subject and proposes a method to retrieve the lost component.

1.4.2 The longitudinal component of a dipole

The work in this section was published in [25]. The ideas and results were generated jointly with Matthew R. Foreman.

Start by considering a fluorescent molecule as an electric dipole emitter with moment $\mathbf{p} = (p_x, p_y, p_z)$. A high numerical aperture (NA) lens is used to collect and collimate the far-field radiation pattern. The lens is assumed to be ideal and immersed in a medium of the same refractive index as that containing the dipole. This constitutes a special case, however a fuller treatment in which the dipole field propagates through dielectric interfaces (e.g. a coverglass-immersion fluid interface) does not alter the symmetry of the problem [26] and hence the following discussion is unchanged.

Vectorial ray tracing can be used to find the electric field $\mathbf{E}_{bf}$ in the back focal plane of the collector lens by using generalised Jones matrices. This yields [27]

$$\vec{E}_{bf} = \frac{1}{2\sqrt{\cos \vartheta}} \begin{pmatrix} p_x a_1^x - p_y a^y - 2p_z a_1^z \\ p_y a_2^x - p_x a^y - 2p_z a_2^z \\ 0 \end{pmatrix} \quad (1.36)$$

where

$$\begin{aligned} a_1^x &= (1 + \cos \vartheta) - (1 - \cos \vartheta) \cos 2\phi \\ a_2^x &= (1 + \cos \vartheta) + (1 - \cos \vartheta) \cos 2\phi \\ a^y &= (1 - \cos \vartheta) \sin 2\phi \\ a_1^z &= 2 \sin \vartheta \cos \phi \\ a_2^z &= 2 \sin \vartheta \sin \phi \end{aligned} \quad (1.37)$$

The $\sqrt{\cos\vartheta}$ factor is required to ensure energy conservation [28].

Equation (1.36) shows that it is possible to consider the field in the back focal plane as having contributions from three independent electric dipoles aligned with the coordinate axes. The field distributions associated with each of these dipoles are shown in figure (1.11).

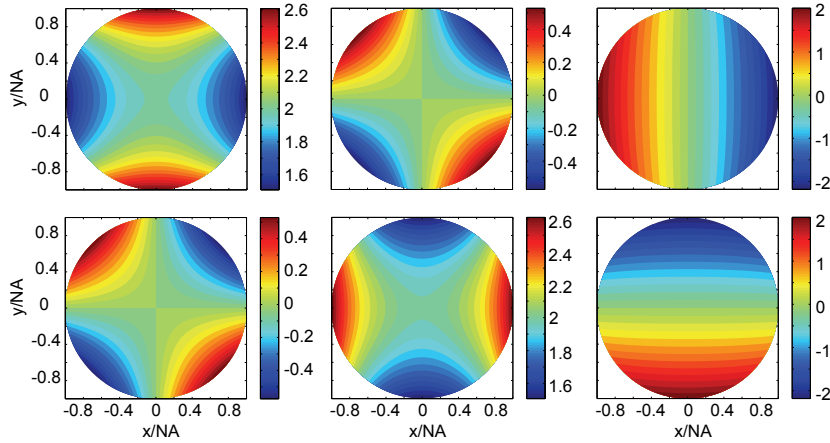


Figure 1.11: Field distributions for E_x (top) and E_y (bottom) in the back focal plane of a collecting lens ($\text{NA} = 0.966$) arising from orthogonal electric dipoles orientated along the x , y and z coordinate axes respectively (left to right).

Refocusing of the collected beam by a second lens corresponds to a coherent integration over the back focal plane. The symmetry of the distributions in figure (1.11) immediately shows that points in the integration plane will either cancel pairwise or superpose constructively. The resulting field at the focus thus takes the form $\vec{E}_f \propto (p_x, p_y, 0)$ where the z component is negligible since the light is refocused with a low NA lens. This can also be shown rigorously by integration of equation (1.36). Refocusing by a high NA lens, as described by the Debye-Wolf diffraction integral [29], gives a similar result whereby $\vec{E}_f \propto (K_1 p_x, K_1 p_y, K_2 p_z)$ at the focus where K_1 and K_2 are constants and $K_2/K_1 \ll 1$ [27].

This result encapsulates the difficulty in determining the p_z component of a dipole, however also suggests a means by which it can be overcome. A significant p_z dependence can be introduced into the “image” of the dipole by breaking the symmetry of the distributions in the back focal plane of the collector lens. This could be done, for example, by apodisation or phase modification of the beam. The latter is considered since it does not inherently reduce the optical throughput of

the system, which is a major issue in single molecule experiments [30]. The optical setup shown in figure (1.12) is proposed as a means to detect the full 3D orientation of an electric dipole in real time.

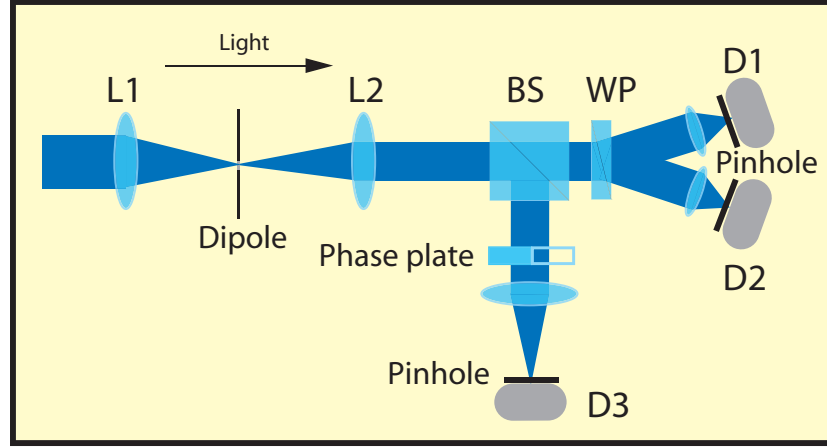


Figure 1.12: Proposed optical setup for determination of the full three-dimensional orientation of an electric dipole.

Light collected from the dipole is incident into a beam splitter from which one of the output beams is further passed through a Wollaston prism (WP) which splits the field into its constituent x and y components. The other portion of light output from the beam splitter is passed through a phase mask which imposes a π phase delay to the beam in the 1st and 4th quadrants with respect to the beam in the 2nd and 3rd quadrants so as to break the symmetry as discussed above. The mask shown in figure (1.13) has an amplitude transmittance function of

$$T(x, y) = \begin{cases} -1 & \text{for } x \geq 0 \\ +1 & \text{for } x < 0 \end{cases} \quad (1.38)$$

In its simplest form this could be implemented using a glass block of appropriate thickness placed across one half of the beam.

Finally the field in each arm is refocused onto point detectors (such as avalanche photodiodes) D_1 , D_2 and D_3 , so that detection is field sensitive yielding measurements of $C_x|p_x|^2$, $C_y|p_y|^2$ and $C_z|p_z|^2$ respectively. It is reasonable to assume that $C_x = C_y$ since the two detection arms are identical albeit for a 90° rotation in the state of polarisation. Note that introduction of an additional arm with a polariser at 45° to the x axis allows the angular ambiguity as to which quadrant the dipole

lies in to be removed, although the light incident on each detector is reduced. Furthermore the magnitude of the dipole components for a general dipole will depend on the excitation field which then has consequences for the SNR in each detector. The illumination must hence be matched to the specific detection needs.

For orientational measurements it is only necessary to work with the ratios between the detector signals. The value of the constant $C = C_x/C_z$ must be found by means of a calibration scan of a free dipole, which can be physically realised using point scatterers [31]. Note that C is dependent on the splitting ratio of the beam splitter and can hence be controlled to some extent, but does cause a trade-off of SNR between different detector arms. Since the scatterer re-radiates as a dipole with an effective moment $\mathbf{p}_{eff} \propto \vec{E}_{ill}$ the three detector signals will theoretically map the focused field distribution, which can be calculated exactly using the Debye-Wolf integral. It is then a simple matter to determine the constant of proportionality. This method hence also provides a means to measure the longitudinal component of an electric field distribution.

Of importance in any practical implementation of the proposed detection scheme are the alignment tolerances of the phase mask. Both rotational and translational misalignments of the mask are considered as depicted in figure (1.13) since any misalignment can be treated as a combination of both. A rotational misalignment by

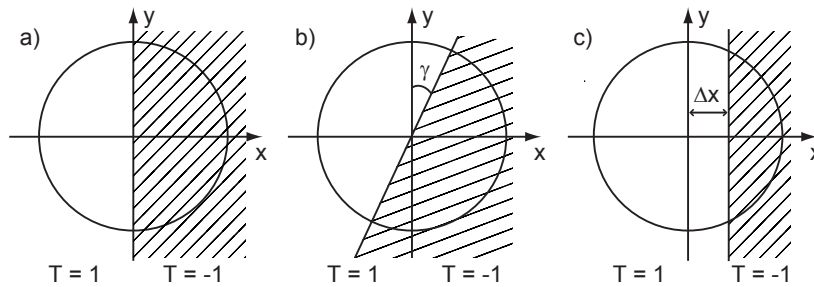


Figure 1.13: a) Proposed phase mask b) Rotational misalignment c) Translational misalignment.

an angle γ modifies the transmittance function, however assuming a symmetric pupil (which is normally the case) integration over the field does not yield any dependence on γ . The detector signals are thus insensitive to pure rotational misalignments, a property which can again be attributed to the symmetrical nature of the back focal plane field distributions.

Translational misalignments of the mask do however have a detrimental effect on the detector signals. A horizontal shift of Δx modifies the transmittance function to

$$T(x, y) = \begin{cases} -1 & \text{for } x \geq \Delta x \\ +1 & \text{for } x < \Delta x \end{cases} \quad (1.39)$$

which in general causes a mixing of the p_x and p_y signals into D_3 . Figure (1.14) shows the degree to which this affects the system, as found by numerical simulations and as parameterised by the extinction ratios between the measured signal in D_3 for a pure x and pure y dipole (denoted D_3^x and D_3^y respectively) relative to that for a pure z dipole (D_3^z).

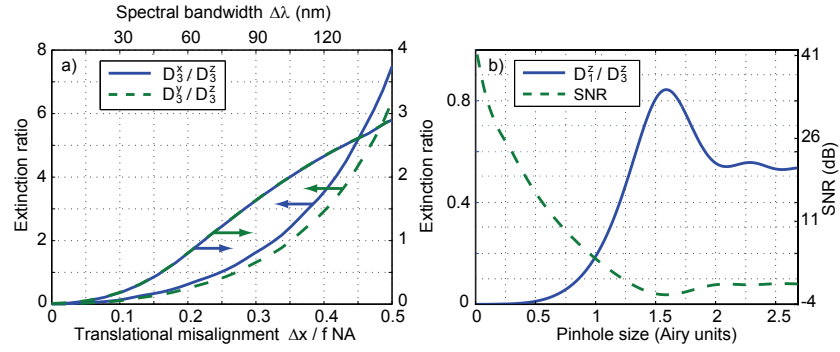


Figure 1.14: a) Extinction ratios as a function of mask misalignment (bottom-left axes) and spectral bandwidth (top-right axes). b) SNR for finite detector sizes.

Hitherto discussion has been limited to monochromatic light of wavelength λ . This is however an unrealistic assumption since fluorescent molecules emit a spectrum $S(\lambda)$ of wavelengths of bandwidth $\Delta\lambda$. A non-zero bandwidth will have consequences with regard to the glass slab used to impose the π phase delay since it can only be designed to operate perfectly at a single wavelength. The signal recorded in each detector can be calculated using

$$D_\nu^{tot} = \int_{\lambda_{min}}^{\lambda_{min} + \Delta\lambda} S(\lambda) D_\nu(\lambda) d\lambda \quad \nu = 1, 2, 3 \quad (1.40)$$

where $D_\nu(\lambda)$ is the detector reading for light of wavelength λ in the presence of a phase plate designed for operation at λ_0 and λ_{min} is the lower bound on the spectrum.

Results from numerical simulations assuming a Lorentzian spectrum profile cen-

tered on $\lambda_0 = 450\text{nm}$ are shown in figure (1.14a). Other spectral profiles give similar results. Only the extinction ratios for D_3 are shown since no variation is seen in D_1 and D_2 . Note that the p_x and p_y signals mix into the D_3 signal and become double that of the p_z measurement for a bandwidth of $\sim 100\text{nm}$. Although typical fluorophores have bandwidths of approximately 30nm this mixing effect can limit the performance of the setup. Use of broadband waveplates to implement the phase mask can however greatly improve the situation.

As a final consideration the previous assertion of point detectors is relaxed. In general D_1 and D_2 operate as confocal microscopes with finite sized pinholes illuminated with x and y polarised light respectively, an analysis of which can be found in [12]. Thus here only measuring a pure p_z dipole is considered. It is found that the SNR arising in this case is worse than for measuring pure p_x and p_y dipoles and is hence the limiting case.

Finite sized detectors record the integrated intensity of the focused light distribution over their extent, however when imaging a dipole the off-axis intensity is dependent on all three dipole components [27]. For a dipole with moment $\mathbf{p} = (0, 0, 1)$ parasitic signals are thus introduced into D_1 and D_2 . Treating these signals as noise the SNR can be defined as $10 \log_{10} (D_3^z / 2D_1^z)$, the behaviour of which is shown in figure (1.14b) along with that of the extinction ratio $D_1^z / D_3^z = D_2^z / D_3^z$. Typical values of the SNR in single molecule experiments are $\sim 15\text{dB}$ [30]. As such figure (1.14b) shows that the effect of a finite detector size becomes more dominant once the radius of the detector is approximately half the size of the Airy disc. Detectors smaller than this are thus preferable.

An alternative to retrieve the p_z component of a dipole is to use a radial polarisation converter and a Wollaston prism instead of the step-phase plate. Introducing an optical path difference of π across the field before focusing produces a purely p_z contribution along the optical axis. If now a field with dominant longitudinal component can be produced using either a radial polarisation converter [32] or a dipole with moment placed along the optical axis, both cases could generate a p_z dependence on the field. Analogous to how a retarder transforms elliptically polarised light into linear, the possibility of using a radial polarisation converter to transform the longitudinal component to transversal is investigated here.

The retrieval of the p_z component can be shown using vectorial ray tracing. The field produced by a dipole is collimated, passed through a radial polarisation converter and then analysed using a Wollaston prism. The collimated field of a dipole can be found using equation (1.22) with $\vec{E}_2 = \mathbf{r} \times (\mathbf{r} \times \vec{p})$, and the radial polarisation converter can be represented using the matrix $\mathbf{R}(\phi)$ given in equation (1.18). Multiplying the matrix with the collimated field yields

$$\vec{E}_3 = \frac{1}{\sqrt{\cos \vartheta}} \begin{pmatrix} p_z \sin \vartheta - \cos \vartheta (p_x \cos \phi + p_y \sin \phi) \\ p_x \sin \phi - p_y \cos \phi \\ 0 \end{pmatrix}. \quad (1.41)$$

To then analyse the E_x and E_y components the field is passed through a Wollaston prism. Two beams emerge from the prism and each beam is measured by a photodetector. Given that the irradiance in each detector is given by

$$I_x = \frac{\pi^2}{4} (p_x^2 + p_y^2 + 2p_z^2), \quad (1.42a)$$

$$I_y = \frac{\pi^2}{2} (p_x^2 + p_y^2), \quad (1.42b)$$

the $|p_z|$ component of the dipole can thus be found by

$$|p_z| = \frac{1}{\pi} \sqrt{2I_x - I_y}. \quad (1.43)$$

Note that the method has several advantages. The light efficiency of the method is in theory 100%. The losses that can be found in practice are the back reflections of the optics. Also note that the measurement yields the magnitude of p_z independently of the ellipticity present in the beam.

In summary it was found that by breaking the symmetry of the back focal plane field distributions from an imaged electric dipole it is possible to determine the longitudinal component of its dipole moment. More specifically a scheme whereby one half of the collected beam is subject to a π phase delay has been presented. Potential sources of error for experimental implementations have also been discussed. Future work in this field is to include experimental verification of the principles of operation of the method. Finally note that the discussed technique works equally

well when imaging the electric field of magnetic dipoles.

1.5 Conclusion

The concepts of polarisation and the ellipse of polarisation were given after introducing light an electromagnetic field. Vectorial ray tracing was also explained with the aim of understanding basic high NA imaging, and scanning confocal imaging was introduced together with an explanation of the type of detection and consequences. In the case of coherent detection the spatial and spectral resolution are not equivalent concepts. The resolution in the spatial frequency domain depends on the imaging system and on the state of polarisation emitted by the object. The spatial resolution depends on both the imaging system and on the optical properties of the object. It was also concluded that in confocal imaging the aforementioned resolution concepts can be equivalent if the optical path difference between two objects is $\pi/2$.

An efficient method to measure the longitudinal component of the field was also introduced. The measurement however is left as future work. The next chapter addresses more in detail the concepts of polarisation and coherence of a field. An insight to the difference between degree of polarisation and coherence is given.

Chapter 2

Polarisation and coherence

The objective of this chapter is to introduce a formulation that can describe the polarisation of a beam and its statistics. The chapter includes an overview of the scalar and vectorial variations of a field. In fact, this chapter proposes a solution to the coherence versus polarisation ongoing issue in the literature.

The chapter is divided in two parts. The first part of the chapter introduces a formalism that accounts for fluctuations in the beam, i.e. the electric coherency matrix, the mutual coherency matrix and the polarisation matrix. The second part then focuses closely on quantifying the fluctuations in a field, and introduces a novel metric to account for spatial depolarisation effects.

2.1 Fluctuations of light

Most electromagnetic fields found in practice are neither steady nor spatially homogeneous. Take for example, the type of illumination used in this work, a field emitted by a “Coherent Cube diode laser”¹. The polarisation of this field is fluctuating in time, the beam’s cross section contains a polarisation distribution, the wavefront is highly distorted and the field has a finite bandwidth. This means that the phase and the state of polarisation of the propagating field are changing in time and in space. Depending now on the nature of the variations, the changes are considered to be part of the coherence or the polarisation properties of the beam. More specifically, a beam is partially polarised when the chances of measuring a particular polarisation

¹As in the brand and make

state can be described by a probability distribution. In the extreme case when every polarisation state is equally probable to be present in the beam the light is termed unpolarised. Analogously, a beam is partially coherent when the wavefront is fluctuating with a given probability distribution or, in the extreme case, incoherent when it is not possible to determine an average wavefront. In any case, both coherence and polarisation effects can be considered using a single matrix known as the electric coherency matrix. This matrix is a temporal version of a matrix first introduced in 1930 by Wiener[33] in the frequency domain. Both the electric coherency matrix and Wiener's matrix can describe partially coherent and partially polarised fields, however, only the electric coherency matrix is considered in this research.

As shown by Gil[34], the electric coherency matrix can describe a two or a three dimensional field. For simplicity however, only the derivation of the 2D description is given here; the 3D case can be derived if necessary following the same operations as the 2D case. First consider an arbitrary electric field given by

$$\vec{E}(\mathbf{r}, t) = \begin{pmatrix} A_x(\mathbf{r}, t)e^{i\delta_x(\mathbf{r}, t)} \\ A_y(\mathbf{r}, t)e^{i\delta_y(\mathbf{r}, t)} \end{pmatrix}. \quad (2.1)$$

The field is then evaluated at $(\mathbf{r}_1, t)$ and directly multiplied² by the Hermitian adjoint of the field itself evaluated at $(\mathbf{r}_2, t - \tau)$, where $\mathbf{r}_1$ and $\mathbf{r}_2$ are two arbitrarily chosen points in space and τ is a time delay. The result of this multiplication is then averaged yielding the 2×2 Hermitian matrix known as the *electric coherency matrix*³. The latter can be written explicitly as

$$\mathbf{\Gamma}(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle \vec{E}(\mathbf{r}_1, t) \otimes \vec{E}(\mathbf{r}_2, t - \tau)^\dagger \rangle \quad (2.2a)$$

$$= \begin{pmatrix} \langle E_x(\mathbf{r}_1, t)E_x^*(\mathbf{r}_2, t - \tau) \rangle & \langle E_x(\mathbf{r}_1, t)E_y^*(\mathbf{r}_2, t - \tau) \rangle \\ \langle E_y(\mathbf{r}_1, t)E_x^*(\mathbf{r}_2, t - \tau) \rangle & \langle E_y(\mathbf{r}_1, t)E_y^*(\mathbf{r}_2, t - \tau) \rangle \end{pmatrix}. \quad (2.2b)$$

The symbols † and * denote the Hermitian adjoint and complex conjugate operations respectively, $\otimes$ represents the Kronecker[35, 36, 37] or direct product and the brackets denote average. This average is defined depending on the statistics of the

²Kronecker product

³This matrix is also known in the literature as the *mutual coherency matrix*

fluctuations. If the fluctuations are ergodic and they occur in a time interval much shorter than the integration time of the measurement, the average can be defined as a time average.

Particular attention is now given to two special cases of the electric coherency matrix: the self coherency matrix and the polarisation matrix. The first is used to investigate the relation between coherence and polarisation, and the polarisation matrix is used to describe partially polarised light. Additionally, a notation more suitable for the laboratory is derived later in section 3.1 using the polarisation matrix.

The *self coherency matrix* is nothing more than the electric coherency matrix evaluated at a single point $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$,

$$\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, \tau) = \begin{pmatrix} \langle E_x(\mathbf{r}, t) E_x^*(\mathbf{r}, t - \tau) \rangle & \langle E_x(\mathbf{r}, t) E_y^*(\mathbf{r}, t - \tau) \rangle \\ \langle E_y(\mathbf{r}, t) E_x^*(\mathbf{r}, t - \tau) \rangle & \langle E_y(\mathbf{r}, t) E_y^*(\mathbf{r}, t - \tau) \rangle \end{pmatrix}, \quad (2.3)$$

and it represents the correlation of the field with itself at two instants. This matrix $\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, \tau)$ is then evaluated at $\tau = 0$ yielding

$$\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, 0) = \begin{pmatrix} \langle A_x^2 \rangle & \langle A_x A_y e^{i\Delta_{xy}} \rangle \\ \langle A_y A_x e^{i\Delta_{yx}} \rangle & \langle A_y^2 \rangle \end{pmatrix}, \quad (2.4)$$

where $\Delta_{jk} = \delta_j - \delta_k$ is the relative phase difference between $j, k = x, y$. The matrix in equation (2.4) was originally introduced by Fano in 1954 using the name of the density matrix[38]. A few years later Wolf[39, 40], Parrent and Roman[41], and Barakat[42] started to develop the theory of partial polarisation and called $\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, 0)$ the coherency matrix. To avoid confusion however, $\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, 0)$ is nowadays known as the *polarisation matrix*. The matrix $\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, 0)$ is widely used in the literature to study the theory of partially polarised light[43], and it represents the correlation of the field with itself.

In this work, the polarisation matrix is used to understand the temporal and spatial polarisation variations of a beam. The temporal fluctuations are an extensively studied subject in the literature, however, the spatial variations are not. The latter in fact have triggered an ongoing discussion in the literature regarding the

spatial distribution of polarisation in a beam and its coherence. The next section elaborates on this discussion and proposes a solution by introducing a novel metric.

2.2 Degree of polarisation and degree of coherence

It is possible to find nowadays a disagreement in the literature regarding polarisation and coherence. The scientific community has not yet agreed if coherence is a scalar or a vectorial property of the light. Take for example the case of a monochromatic and spatially inhomogeneous polarised field emitted by a significantly small source. Since a fully coherent field means a perfectly correlated field[21], it is not clear if this example should be considered as partially coherent or not. From a scalar point of view, such field is considered fully coherent since its phase is correlated in time and space. From a vectorial point of view however, the field is considered partially coherent since it has a polarisation distribution. Indeed it is possible to regard the polarisation properties of the beam as a coherence effect, yet in this research it is considered as inconvenient. Given that the scalar properties of a beam are in general independent of the vectorial properties[8], it is more convenient to leave the term coherence for scalar effects and create a new metric to account for spatial variations in the polarisation. Only in this context it is possible to ensure that a metric to solely quantify each one of the contributions to the fluctuations of a beam exists. Given that it is possible to identify four contributions: scalar temporal, scalar spatial, vectorial temporal and vectorial spatial, hence four metrics are necessary to complete the theory of polarisation and coherence. The *degree of temporal coherence* (DOTC)[43] measures the fluctuations of the beam due to the spectral content of the beam itself. The *degree of spatial coherence* (DOSC)[43] quantifies the variations throughout the wavefront of the beam. The temporal fluctuations in the polarisation are being measured by the *degree of polarisation* and the spatial variations are quantified by the *degree of spatial polarisation*.

As one of the contributions of this work the degree of spatial polarisation is introduced here. To avoid confusion the degree of polarisation is renamed here as the *degree of temporal polarisation* and in fact generalised here to any dimensions. Before

expanding further on both of the polarisation metrics, the well known coherence metrics, i.e. the DOTC and the DOSC are introduced next. The aim is to illustrate the implications that the polarisation has on coherence, and hence justify the need for the new metric. Ideally the four previously mentioned metrics will complete the theory of polarisation and coherence.

2.2.1 Degree of temporal and spatial coherence

Degree of coherence (DOC) is commonly defined as the fringe visibility in an interferometric experiment[43]. This metric is commonly used to quantify the degree of correlation within the field. In particular when the DOC is termed spatial, the correlation refers to the transverse direction of the beam. The *degree of spatial coherence* (DOSC) can be measured using for example the two pinhole version of the Young's experiment[16] shown in figure (2.1). Alternatively when the DOC is termed temporal, the correlation is in the longitudinal direction of the beam. In this case the *degree of temporal coherence* (DOTC) can be measured using for example a Michelson's interferometer[44]. As opposed to the DOSC, the DOTC is a spatially independent metric. The DOTC is actually a correlation between the field and a delayed copy of itself.

Both the DOTC and the DOSC are derived next using the Young's experiment and a Michelson's interferometer, respectively. The derivation is first from a scalar point of view and then from a vectorial point of view.

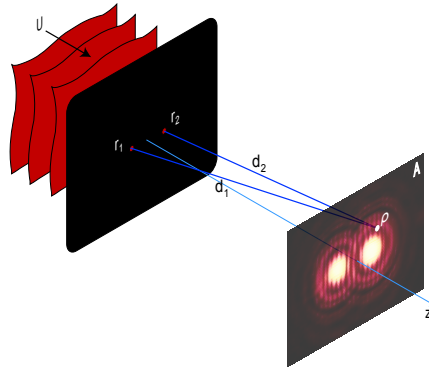


Figure 2.1: Young interferometer

Consider first the scalar derivation of the degree of spatial coherence. Assuming a scalar wave U incident upon the pinholes located at $\mathbf{r}_1$ and $\mathbf{r}_2$, the measured

interference at the screen A of figure (2.1) is then described by

$$I(\boldsymbol{\rho}, t) = |K_1 U(\mathbf{r}_1, t) + K_2 U(\mathbf{r}_2, t - \tau)|^2 \quad (2.5a)$$

$$= |K_1 U(\mathbf{r}_1, t)|^2 + |K_2 U(\mathbf{r}_2, t - \tau)|^2 + 2|K_1 K_2| \text{Re}[U(\mathbf{r}_1, t) U^*(\mathbf{r}_2, t - \tau)]. \quad (2.5b)$$

Here $\boldsymbol{\rho}$ is a coordinate vector defined in screen A. Note that the pattern depends on the optical path differences $\tau = t_2 - t_1$, where the time is related to the distance by $t_1 = d_1/c$ and $t_2 = d_2/c$; here c is the speed of light. The complex constant $K_j \approx -i/(\bar{\lambda} \mathbf{r}_j) p_j$ is known as the propagation factor of each pinhole $j = 1, 2$, where $\bar{\lambda}$ is the mean wavelength and p_j is the area of each pinhole. Additionally, the constants K_j derived in reference [9] from the Huygens-Fresnel principle have the property

$$I_j(\boldsymbol{\rho}) = |K_j|^2 I(\mathbf{r}_j), \quad \text{for } j = 1, 2. \quad (2.6)$$

Assuming now that the fluctuations of the light are ergodic and shorter than the integration time of the measurement, it is possible to define a time average. Given a probability density function $f(U, t)$, a time average is defined as

$$I(\boldsymbol{\rho}) = \langle I(\boldsymbol{\rho}, t) \rangle = \int U(\boldsymbol{\rho}, t) U^*(\boldsymbol{\rho}, t) f(U, t) dt. \quad (2.7)$$

Equation (2.6) is then used to write (2.5) in terms of $\boldsymbol{\rho}$, and the result is time averaged yielding,

$$I(\boldsymbol{\rho}) = I_1(\boldsymbol{\rho}) + I_2(\boldsymbol{\rho}) + 2\sqrt{I_1(\boldsymbol{\rho})I_2(\boldsymbol{\rho})} \text{Re} \left\{ \frac{\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)}{\sqrt{I(\mathbf{r}_1)I(\mathbf{r}_2)}} \right\}, \quad (2.8)$$

where

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle U(\mathbf{r}_1, t) U^*(\mathbf{r}_2, t - \tau) \rangle \quad (2.9)$$

is known as the *mutual coherence function* [39]; this function defines a measure of the correlation between two points in terms of their optical path difference. When the mutual coherence function is normalised by the square root of the intensity at each point, $\sqrt{I(\mathbf{r}_1)I(\mathbf{r}_2)} = \sqrt{\Gamma(\mathbf{r}_1, \mathbf{r}_1, 0)\Gamma(\mathbf{r}_2, \mathbf{r}_2, 0)}$, it yields the *spatial degree of*

coherence⁴

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)}{\sqrt{\Gamma(\mathbf{r}_1, \mathbf{r}_1, 0)\Gamma(\mathbf{r}_2, \mathbf{r}_2, 0)}}. \quad (2.10)$$

Finally, in order to make the relationship between degree of coherence and fringe visibility clear let

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| e^{-i[2\pi\bar{\nu}\tau - \varphi(\mathbf{r}_1, \mathbf{r}_2, \tau)]}, \quad (2.11)$$

where $\bar{\nu}$ is the mean frequency and φ is a phase given by the imaginary part of Γ . The substitution of equation (2.11) into (2.8) yields

$$I(\boldsymbol{\rho}) = I_1(\boldsymbol{\rho}) + I_2(\boldsymbol{\rho}) + 2\sqrt{I_1(\boldsymbol{\rho})I_2(\boldsymbol{\rho})} |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| \cos[2\pi\bar{\nu}\tau - \alpha(\mathbf{r}_1, \mathbf{r}_2, \tau)], \quad (2.12)$$

and since fringe visibility is defined as[44]

$$\mathcal{V} = \frac{I_{\max}(\boldsymbol{\rho}) - I_{\min}(\boldsymbol{\rho})}{I_{\max}(\boldsymbol{\rho}) + I_{\min}(\boldsymbol{\rho})}, \quad (2.13)$$

it follows that

$$\mathcal{V} = \frac{2\sqrt{I_1(\boldsymbol{\rho})I_2(\boldsymbol{\rho})}}{I_1(\boldsymbol{\rho}) + I_2(\boldsymbol{\rho})} |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)|. \quad (2.14)$$

It is possible to conclude from equation (2.14), that the DOSC is proportional to the fringe visibility. In the special case when $I_1(\boldsymbol{\rho}) = I_2(\boldsymbol{\rho})$, then $\mathcal{V} = |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)|$.

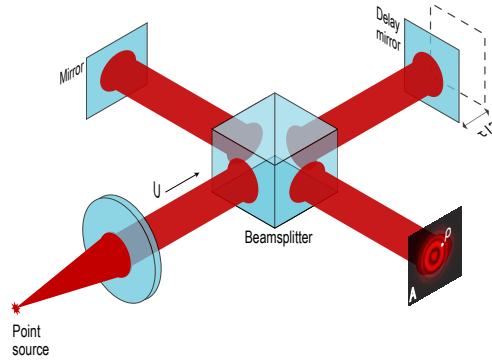


Figure 2.2: Michelson interferometer

Regarding the derivation of the degree of temporal coherence (still from a scalar point of view), the steps to derive the DOSC are strictly followed. In this case

⁴This metric is also known as the *complex degree of mutual coherence*.

however, the interference pattern is formed by the wave $U(\mathbf{r}, t)$ and its time delayed copy $U(\mathbf{r}, t - \tau)$. This interference can be achieved using for example the Michelson interferometer in figure (2.2). After derivation, the *degree of temporal coherence*⁵ is given by

$$\gamma(\mathbf{r}, \mathbf{r}, \tau) = \frac{\Gamma(\mathbf{r}, \mathbf{r}, \tau)}{|\Gamma(\mathbf{r}, \mathbf{r}, 0)|}. \quad (2.15)$$

where

$$\Gamma(\mathbf{r}, \mathbf{r}, \tau) = \langle U(\mathbf{r}, t) U^*(\mathbf{r}, t - \tau) \rangle \quad (2.16)$$

is the *self coherence function*.

It is possible to conclude from equations (2.9) and (2.16), that the visibility of the fringes $\mathcal{V}$ in either experiment depends on how well correlated the wavefronts are. In the case of Young's interferometer, $\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ is a metric of the correlation in the transverse direction of the beam i.e. spatial coherence, while for the Michelson interferometer, $\gamma(\mathbf{r}, \mathbf{r}, \tau)$ is a measure of the correlation in the longitudinal direction of the beam, i.e. temporal coherence. As it was shown by Fresnel[45] however, the fringe visibility also depends on the state of polarisation of each wave; specifically on the dot product between the polarisation states that are being correlated.

To show this polarisation dependance, let $\vec{E}(\mathbf{r}_1, t)$ and $\vec{E}(\mathbf{r}_2, t - \tau)$ be a pair of arbitrary polarised electric fields emergent from the pinholes placed at $\mathbf{r}_1$ and $\mathbf{r}_2$ in figure (2.1). The irradiance pattern produced by the superposition of $\vec{E}(\mathbf{r}_1, t)$ and $\vec{E}(\mathbf{r}_2, t - \tau)$ at the point $\boldsymbol{\rho}$ can be described by

$$I(\boldsymbol{\rho}) = I_1(\boldsymbol{\rho}) + I_2(\boldsymbol{\rho}) + 2\sqrt{I_1(\boldsymbol{\rho})I_2(\boldsymbol{\rho})} \operatorname{Re} \{ \mu(\mathbf{r}_1, \mathbf{r}_2, \tau) \}, \quad (2.17)$$

where $I_i(\boldsymbol{\rho}) = \langle \vec{E}^\dagger(\mathbf{r}_i, t) \cdot \vec{E}(\mathbf{r}_i, t) \rangle$ is the contribution to the irradiance from the i th pinhole at the screen and

$$\mu(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{\operatorname{Tr} [\mathbf{\Gamma}(\mathbf{r}_1, \mathbf{r}_2, \tau)]}{\sqrt{\operatorname{Tr} [\mathbf{\Gamma}(\mathbf{r}_1, \mathbf{r}_1, 0)]} \sqrt{\operatorname{Tr} [\mathbf{\Gamma}(\mathbf{r}_2, \mathbf{r}_2, 0)]}} \quad (2.18)$$

is the vectorial equivalent of the scalar degree of spatial coherence; equation (2.10). Note for future reference that in equation (2.18), the matrix $\mathbf{\Gamma}(\mathbf{r}_1, \mathbf{r}_2, \tau)$ is the electric coherency matrix.

⁵This metric is also known as the *complex degree of self coherence*.

Analogous to the derivation of equation (2.12), equation (2.17) can be written in a more intuitive manner as

$$I(\boldsymbol{\rho}) = I_1(\boldsymbol{\rho}) + I_2(\boldsymbol{\rho}) + 2\sqrt{I_1(\boldsymbol{\rho})I_2(\boldsymbol{\rho})} |\mu(\mathbf{r}_1, \mathbf{r}_2, \tau)| \cos[2\pi\bar{\nu}\tau - \varphi(\mathbf{r}_1, \mathbf{r}_2, \tau)], \quad (2.19)$$

which shows that the fringe visibility $\mathcal{V} \propto |\mu(\mathbf{r}_1, \mathbf{r}_2, \tau)|$, and hence proportional to the vector dot product of the two interfering fields. It then follows that by replacing $\vec{E}(\mathbf{r}_2, t - \tau)$ with $\vec{E}(\mathbf{r}, t - \tau)$ and $\vec{E}(\mathbf{r}_1, t)$ with $\vec{E}(\mathbf{r}, t)$, equation (2.18) can be written as

$$\mu(\mathbf{r}, \mathbf{r}, \tau) = \frac{\text{Tr}[\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, \tau)]}{|\text{Tr}[\mathbf{\Gamma}(\mathbf{r}, \mathbf{r}, 0)]|}, \quad (2.20)$$

and hence the same conclusion applies to the degree of temporal coherence.

From the preceding discussion, it is possible to conclude that the fringe visibility depends on both the scalar and the vectorial variations of a beam. It appears that a beam can only be categorised as fully coherent if it is also homogeneously polarised in time and in space. That is if the fringe visibility is considered as a metric to quantify the degree of correlation. In general however, the phase and frequency content of a beam are independent of the polarisation[46]. It is thus possible to split the concepts of coherence and polarisation. For scalar problems, the DOTC and the DOSC can be estimated using equations (2.15) and (2.10) respectively. For vectorial problems however, the polarisation of the beam should be homogenised⁶ before using equations (2.18) and (2.20) to estimate the DOSC and the DOTC respectively. The term coherence is thus assigned to only scalar effects ensuring that a measure to quantify purely wavefront fluctuations exists. However, the fact of homogenising the polarisation creates a gap in the theory of polarisation and coherence. The spatial variations in the polarisation of the beam are not being considered. To then fill the gap in the theory, it is necessary to create a metric to account for the spatial variations, i.e. the degree of spatial polarisation.

The degrees of temporal and spatial polarisation are explained next. Both concepts are introduced by deriving their mathematical expressions in a novel way.

⁶Make spatially and temporally invariant

2.2.2 Degree of temporal polarisation

First consider the case of the degree of temporal polarisation (DOTP). This metric is equivalent to the widely known degree of polarisation (DOP). The DOTP however, is derived here in a more intuitive and general way. The derivation follows the conceptual definition of the DOP.

Since the *degree of polarisation* is defined as the ratio between the irradiance of the polarised part of a beam and its total irradiance[10], the derivation of the DOTP is designed as a two-step thought experiment. First a spatially inhomogeneous perfect analyser is used to extinguish the polarised part of the beam and the throughput irradiance is measured. The analyser is then removed and the irradiance is again measured. The ratio of these two irradiance measurements yield the DOTP.

To derive the DOTP let $\mathbf{\Gamma}(\mathbf{r}_j, \mathbf{r}_j, 0)$ describe a partially polarised⁷ field formed of j homogeneously polarised patches. Since depolarisation is a linear effect [34] $\mathbf{\Gamma}_j = \mathbf{\Gamma}(\mathbf{r}_j, \mathbf{r}_j, 0)$ may be expressed as the linear combination of an unpolarised field $\mathbf{\Gamma}^u$ and a fully polarised one $\mathbf{\Gamma}_j^p$ as

$$\mathbf{\Gamma}_j = \mathbf{\Gamma}_j^p + \mathbf{\Gamma}_j^u. \quad (2.21)$$

The argument is dropped for convenience. Here $\mathbf{\Gamma}_j^u = \alpha \mathbf{I}$, where α is a complex number and $\mathbf{I}$ is the identity matrix. Given that the matrix $\mathbf{\Gamma}_j^p$ is Hermitian symmetric, it has at least one non zero eigenvalue[47].

The field is now propagated through a non depolarising optical element described by a Jones matrix $\mathbf{G}_j$ (see appendix A). Equation (A.3) is Kronecker-multiplied by its Hermitian adjoint and the result is time averaged[10],

$$\langle \vec{E}_{out} \otimes \vec{E}_{out}^\dagger \rangle = \langle \mathbf{G}_j \vec{E}_{in} \otimes \vec{E}_{in}^\dagger \mathbf{G}_j^\dagger \rangle, \quad (2.22)$$

yielding

$$\mathbf{\Gamma}_j^{out} = \mathbf{G}_j \mathbf{\Gamma}_j \mathbf{G}_j^\dagger, \quad (2.23)$$

where $\mathbf{G}_j$ is an arbitrary Jones matrix. Equation (2.23) describes the propagation of

⁷It is meant by partially polarised a field that has possibly inhomogeneous polarisation structure in the transverse spatial dimensions and, at any point along the transverse spatial extent of the beam, a polarisation that has random time-evolution.

partially polarised light through non-depolarising media, and the trace of the output matrix represents the throughput irradiance. Equation (2.21) is then substituted into (2.23), and the trace is taken of both sides of the resulting expression giving

$$\text{Tr} [\mathbf{G}_j \mathbf{\Gamma}_j \mathbf{G}_j^\dagger] = \text{Tr} [\mathbf{G}_j \mathbf{\Gamma}_j^p \mathbf{G}_j^\dagger] + \text{Tr} [\mathbf{G}_j \mathbf{\Gamma}_j^u \mathbf{G}_j^\dagger]. \quad (2.24)$$

Note that the additive property⁸ of the trace was used here to split the terms on the right hand side. Once the irradiance expression is derived, equation (2.24), the next step is to choose $\mathbf{G}_j$ so the polarised part of the beam is completely extinguished, as discussed before. To accomplish this the eigenvectors $\tilde{\mathbf{X}}_j$ in the eigensystem of $\mathbf{G}_j$

$$\mathbf{G}_j = \tilde{\mathbf{X}}_j \tilde{\mathbf{\Phi}}_j \tilde{\mathbf{X}}_j^\dagger \quad (2.25)$$

are chosen to be the same as in the eigensystem of $\mathbf{\Gamma}_j^p$

$$\mathbf{\Gamma}_j^p = \tilde{\mathbf{X}}_j \mathbf{\Lambda}_j^p \tilde{\mathbf{X}}_j^\dagger, \quad (2.26)$$

and the eigenvalues are chosen to be related by

$$\tilde{\mathbf{\Phi}}_j = \sqrt{\tilde{\lambda}_j^p \mathbf{I} - \tilde{\mathbf{\Lambda}}_j^p}, \quad (2.27)$$

where $\tilde{\lambda}_j^p = \lambda_j^p / \text{Tr} [\mathbf{\Gamma}_j^p]$ with λ_j^p the non-zero eigenvalue of $\mathbf{\Gamma}_j^p$ and $\tilde{\mathbf{\Lambda}}_j^p$ is a diagonal matrix containing $\tilde{\lambda}_j^p$ as the only non-zero element⁹. The matrix $\tilde{\mathbf{\Phi}}_j$ is also diagonal and contains the eigenvalues of $\mathbf{G}_j$, and $\mathbf{I}$ is the identity matrix.

Equations (2.25-2.27) are then substituted in (2.24), and the resulting expression is rearranged using the multiplicative property¹⁰ of the trace. The result of this is

⁸ $\text{Tr} [\mathbf{A} + \mathbf{B}] = \text{Tr} [\mathbf{A}] + \text{Tr} [\mathbf{B}]$

⁹Since $\tilde{\mathbf{G}}$ is by definition dimensionless, it is necessary for $\tilde{\mathbf{\Phi}}_j$ and $\tilde{\mathbf{X}}_j$, to also be dimensionless.

¹⁰ $\text{Tr} [\mathbf{AB}] = \text{Tr} [\mathbf{BA}]$

given by

$$\text{Tr} [\mathbf{G}_j \mathbf{\Gamma}_j \mathbf{G}_j^\dagger] = \text{Tr} [\mathbf{\Gamma}_j^p \mathbf{G}_j^\dagger \mathbf{G}_j] + \text{Tr} [\mathbf{\Gamma}_j^u \mathbf{G}_j^\dagger \mathbf{G}_j] \quad (2.28a)$$

$$= \text{Tr} [\mathbf{\Gamma}_j^p \tilde{\mathbf{X}}_j |\tilde{\mathbf{\Phi}}_j|^2 \tilde{\mathbf{X}}_j^\dagger] + \text{Tr} [\mathbf{\Gamma}_j^u \tilde{\mathbf{X}}_j |\tilde{\mathbf{\Phi}}_j|^2 \tilde{\mathbf{X}}_j^\dagger] \quad (2.28b)$$

$$= \text{Tr} [\mathbf{\Lambda}_j^p (\tilde{\lambda}_j^p \mathbf{I} - \mathbf{\Lambda}_j^p)] + \alpha \text{Tr} [\mathbf{I} (\tilde{\lambda}_j^p \mathbf{I} - \mathbf{\Lambda}_j^p)] \quad (2.28c)$$

$$= \alpha \tilde{\lambda}_j^p (\text{Tr} [\mathbf{I}] - 1), \quad (2.28d)$$

or equivalently

$$\text{Tr} [\mathbf{G}_j \mathbf{\Gamma}_j \mathbf{G}_j^\dagger] = \frac{\tilde{\lambda}_j^p \text{Tr} [\mathbf{\Gamma}_j^u]}{q} \quad (2.29)$$

where

$$q = \frac{\text{Tr} [\mathbf{I}]}{\text{Tr} [\mathbf{I}] - 1}. \quad (2.30)$$

Substituting now equation (2.29) into the trace of equation (2.21) yields

$$\text{Tr} [\mathbf{\Gamma}_j^p] = \text{Tr} [\mathbf{\Gamma}_j] - \frac{q}{\tilde{\lambda}_j^p} \text{Tr} [\mathbf{G}_j \mathbf{\Gamma}_j \mathbf{G}_j^\dagger]. \quad (2.31)$$

Equation (2.31) is the irradiance of the polarised part of the field. If $\mathbf{G}_j$ was known this equation would conclude the first stage of the derivation. In general however this is not the case. A closer look at the eigensystems shows how the matrix $\mathbf{G}_j$ can be chosen.

To choose $\mathbf{G}_j$ let

$$\mathbf{\Gamma}_j = \tilde{\mathbf{X}}_j' \mathbf{\Lambda}_j \tilde{\mathbf{X}}_j'^\dagger \quad (2.32)$$

be the eigensystem of $\mathbf{\Gamma}_j$. Since equation (2.21) can be written, using the respective, eigensystems as

$$\tilde{\mathbf{X}}_j' \mathbf{\Lambda}_j \tilde{\mathbf{X}}_j'^\dagger = \tilde{\mathbf{X}}_j \mathbf{\Lambda}_j^p \tilde{\mathbf{X}}_j^\dagger + \alpha \mathbf{I} \quad (2.33a)$$

$$= \tilde{\mathbf{X}}_j (\mathbf{\Lambda}_j^p + \alpha \mathbf{I}) \tilde{\mathbf{X}}_j^\dagger, \quad (2.33b)$$

the result implies that the eigenvectors of the polarised part are the same as the eigenvectors of the total field. Equations (2.25) and (2.32) are first substituted in the second term of the right hand side of equation (2.31) and then setting $\tilde{\mathbf{X}}_j' = \tilde{\mathbf{X}}_j$

yields

$$\text{Tr} [\mathbf{\Gamma}_j^p] = \text{Tr} [\mathbf{\Gamma}_j] - \frac{q}{\tilde{\lambda}_j^p} \text{Tr} [\mathbf{\Lambda}_j(\tilde{\lambda}_j^p \mathbf{I} - \mathbf{\Lambda}_j^p)]. \quad (2.34)$$

Note that $(\tilde{\lambda}_j^p \mathbf{I} - \mathbf{\Lambda}_j^p)$ is a diagonal matrix of ones and a zero. It then follows that the term $\text{Tr} [\mathbf{\Lambda}_j(\tilde{\lambda}_j^p \mathbf{I} - \mathbf{\Lambda}_j^p)]$ yields the sum of the diagonal elements of $\mathbf{\Lambda}_j$ but without one element. Equation (2.34) can be written in consequence as

$$\text{Tr} [\mathbf{\Gamma}_j^p] = \text{Tr} [\mathbf{\Gamma}_j] - q(\text{Tr} [\mathbf{\Lambda}_j] - \lambda_j), \quad (2.35)$$

where λ_j is one of the eigenvalues of $\mathbf{\Gamma}_j$. In order to find which eigenvalue is λ_j , equation (2.33) is taken one step further. That is

$$\mathbf{\Lambda}_j = \mathbf{\Lambda}_j^p + \alpha \mathbf{I}. \quad (2.36)$$

This equation implies that the largest eigenvalue of $\mathbf{\Gamma}_j$ is always $\lambda_j^p + \alpha$. Equations (2.32) and (2.36) are then substituted in equation (2.35) yielding

$$\text{Tr} [\mathbf{\Gamma}_j^p] = \text{Tr} [\mathbf{\Gamma}_j] - q(\text{Tr} [\mathbf{\Lambda}_j] - \lambda_j) \quad (2.37a)$$

$$\lambda_j^p = \text{Tr} [\mathbf{\Lambda}_j^p] + \alpha \text{Tr} [\mathbf{I}] - q \text{Tr} [\mathbf{\Lambda}_j^p] - q \alpha \text{Tr} [\mathbf{I}] + q \lambda_j \quad (2.37b)$$

$$0 = q \lambda_j - q \lambda_j^p - (q - 1) \alpha \text{Tr} [\mathbf{I}] \quad (2.37c)$$

$$\lambda_j = \lambda_j^p + \alpha. \quad (2.37d)$$

It is now possible to conclude that λ_j is always the largest eigenvalue of $\mathbf{\Gamma}_j$.

Since all the required variables are now known, the first stage of the derivation concludes with equation (2.35), namely the irradiance measurement of the polarised part. The next step is to write the conventional definition of the *degree of polarisation* explicitly as

$$\mathcal{P}^t = \frac{\text{Tr} [\mathbf{\Gamma}_j^p]}{\text{Tr} [\mathbf{\Gamma}_j]}. \quad (2.38)$$

It then follows that the substitution of equation (2.35) in (2.38) yields the final expression of the *degree of temporal polarisation*,

$$\mathcal{P}_j^t = 1 - \frac{q(\text{Tr} [\mathbf{\Gamma}_j] - \lambda_j)}{\text{Tr} [\mathbf{\Gamma}_j]}, \quad q = \frac{\text{Tr} [\mathbf{I}]}{\text{Tr} [\mathbf{I}] - 1}, \quad (2.39)$$

It can be shown by expanding the previous equation in terms of the eigenvalues of $\mathbf{\Gamma}_j$ that

$$0 \leq \mathcal{P}_j^t \leq 1. \quad (2.40)$$

In the case when $\mathcal{P}_j^t = 1$ the beam is fully temporally polarised ($\alpha = 0$), and when $\mathcal{P}_j^t = 0$ the beam is temporally unpolarised ($\lambda_j^p = 0$).

Note that equation (2.39) can be applied to a field of any dimensions as long as the field is described by a polarisation matrix of rank two. Also note that since the trace of a matrix is invariant to rotation [47], $\mathcal{P}_j^t$ is invariant to co-ordinate rotation. $\mathcal{P}_j^t$ can be integrated over a detected area to obtain a single number.

2.2.3 Degree of spatial polarisation

The degree of spatial polarisation (DOSP) is derived from a slightly different procedure than the derivation of the DOTP. In this case the polarised part of the field is examined locally a spatially homogeneous analyser. The throughput irradiance is then divided by the total intensity of the polarised part yielding the DOSP. The result of the derivation is a general and spatially dependant expression to describe the DOSP; general in the sense that it includes temporal effects. If assuming no temporal effects, the general expression can then be simplified leading to a spatial-only expression that is function of a particular case of the DOSC (see section 2.2.1). The general and the spatial-only expressions of the DOSP are derived thoroughly next.

To derive the DOSP consider the polarisation matrix $\mathbf{\Gamma}(\mathbf{r}_j, \mathbf{r}_j, 0) = \mathbf{\Gamma}_j$ of a field having j homogeneously polarised patches, where again

$$\mathbf{\Gamma}_j = \mathbf{\Gamma}_j^p + \mathbf{\Gamma}_j^u. \quad (2.41)$$

The j th patch of the polarised part of the field is then examined using the spatially homogeneous analyser $\mathbf{G}_i$. The polarised part of the field is given by

$$\mathbf{\Gamma}_j^p = \tilde{\mathbf{X}}_j \mathbf{\Lambda}_j^p \tilde{\mathbf{X}}_j^\dagger \quad (2.42)$$

and the analyser is defined as

$$\mathbf{G}_i = \tilde{\mathbf{X}}_i \mathbf{\Phi}_i \tilde{\mathbf{X}}_i^\dagger, \quad (2.43)$$

where $\mathbf{X}_i$ are the eigenvectors from the i th patch in the polarised part of the field and $\mathbf{\Phi}_i$ is a diagonal matrix containing the eigenvalues $\mathbf{G}_i$. Given now that the resulting irradiance from propagating $\mathbf{\Gamma}_j^p$ through $\mathbf{G}_i$ is given by

$$I_{i,j}^p = \text{Tr} \left[\mathbf{G}_i \mathbf{\Gamma}_j^p \mathbf{G}_i^\dagger \right], \quad (2.44)$$

it then follows that to extinguish $\mathbf{\Gamma}_j^p$ when $i = j$, the eigenvalues of $\mathbf{G}_i$ must be defined equivalently to equation (2.27) as

$$\tilde{\mathbf{\Phi}}_i = \sqrt{\tilde{\lambda}_i^p \mathbf{I} - \tilde{\mathbf{\Lambda}}_i^p}. \quad (2.45)$$

The next step is to write the polarimetric measurement, i.e. equation (2.44), in terms of known quantities. First equations (2.43) and (2.45) are substituted in (2.44) giving

$$I_{i,j}^p = \text{Tr} \left[\mathbf{\Gamma}_j^p \tilde{\mathbf{X}}_i (\tilde{\lambda}_i^p \mathbf{I} - \tilde{\mathbf{\Lambda}}_i^p) \tilde{\mathbf{X}}_i^\dagger \right] \quad (2.46a)$$

$$= \tilde{\lambda}_i^p \text{Tr} [\mathbf{\Gamma}_j^p] - \text{Tr} \left[\mathbf{\Gamma}_j^p \tilde{\mathbf{X}}_i \tilde{\mathbf{\Lambda}}_i^p \tilde{\mathbf{X}}_i^\dagger \right] \quad (2.46b)$$

$$= \text{Tr} [\tilde{\mathbf{\Gamma}}_i^p] \text{Tr} [\mathbf{\Gamma}_j^p] - \text{Tr} [\mathbf{\Gamma}_j^p \tilde{\mathbf{\Gamma}}_i^p]. \quad (2.46c)$$

The definition of $\mathbf{\Gamma}_j^u = \alpha_j \mathbf{I}$ and equation (2.41) are then used to expand the previous equation yielding

$$\begin{aligned} I_{i,j}^p &= \text{Tr} [\mathbf{\Gamma}_i] \text{Tr} [\mathbf{\Gamma}_j] - \text{Tr} [\mathbf{\Gamma}_i \mathbf{\Gamma}_j] - \alpha_j (d-1) \text{Tr} [\mathbf{\Gamma}_i] - \\ &\quad - \alpha_i (d-1) \text{Tr} [\mathbf{\Gamma}_j] + (d^2 - d) \alpha_i \alpha_j \end{aligned} \quad (2.47a)$$

$$\begin{aligned} &= \{ \text{Tr} [\mathbf{\Gamma}_i] - \alpha_i (d-1) \} \{ \text{Tr} [\mathbf{\Gamma}_j] - \alpha_j (d-1) \} + \\ &\quad + \alpha_i \alpha_j (d-1) - \text{Tr} [\mathbf{\Gamma}_i \mathbf{\Gamma}_j], \end{aligned} \quad (2.47b)$$

where $d = \text{Tr} [\mathbf{I}]$. Equation (2.47b) can be simplified using the identities

$$\text{Tr} [\mathbf{\Gamma}_j] - \lambda_j = \alpha_j(d - 1), \quad (2.48)$$

$$\text{Tr} [\mathbf{\Lambda}_i \mathbf{\Lambda}_j] = \lambda_i \lambda_j + \alpha_i \alpha_j (d - 1). \quad (2.49)$$

The equivalence of these can be proved with equation (2.36). Substituting then equation (2.48) in (2.47b) and then using (2.49) finally yields

$$I_{i,j}^p = \text{Tr} [\mathbf{\Lambda}_i \mathbf{\Lambda}_j] - \text{Tr} [\mathbf{\Gamma}_i \mathbf{\Gamma}_j]. \quad (2.50)$$

Equation (2.50) represents the polarimetric measurement as a function of the general fields $\mathbf{\Gamma}_i$, $\mathbf{\Gamma}_j$ and their corresponding eigenvalues $\mathbf{\Lambda}_i$ and $\mathbf{\Lambda}_j$.

Note that equation (2.50) reveals a new way of measuring partially polarised fields. Since equation (2.50) and equation (2.44) are equivalent, equation (2.50) also represents the irradiance of a field's polarised part resulting from a polarimetric measurement. The relevance of equation (2.50) becomes apparent in the next chapter when introducing an experimental founded formulism.

The next step is to normalise equation (2.50) to obtain the expression of the degree of spatial polarisation. It is important however to note that despite the general form of equation (2.50), the measurement does not depend on the unpolarised part¹¹ and hence the normalisation is not trivial. In this case it is necessary to divide the measurement by the irradiance of the polarised part of the field at each i, j patch. It is then possible to write the *degree of spatial polarisation* as

$$\mathcal{P}_{i,j}^s = \frac{\text{Tr} [\mathbf{\Lambda}_i \mathbf{\Lambda}_j] - \text{Tr} [\mathbf{\Gamma}_i \mathbf{\Gamma}_j]}{\text{Tr} [\mathbf{\Gamma}_i^p] \text{Tr} [\mathbf{\Gamma}_j^p]}. \quad (2.51)$$

The substitution of equation (2.35) in the previous equation yields

$$\mathcal{P}_{i,j}^s = \frac{\text{Tr} [\mathbf{\Lambda}_i \mathbf{\Lambda}_j] - \text{Tr} [\mathbf{\Gamma}_i \mathbf{\Gamma}_j]}{(\text{Tr} [\mathbf{\Gamma}_i](1 - q) + q\lambda_i) (\text{Tr} [\mathbf{\Gamma}_j](1 - q) + q\lambda_j)}, \quad (2.52)$$

¹¹This is true given that equations (2.44) and (2.50) are equivalent. The apparent temporal dependance disappears in the subtraction of the right hand side. This can also be proved by substituting equation (2.41) in (2.50)

or instead substituting equation (2.39) yields

$$\mathcal{P}_{i,j}^s = \frac{\text{Tr}[\mathbf{\Lambda}_i \mathbf{\Lambda}_j] - \text{Tr}[\mathbf{\Gamma}_i \mathbf{\Gamma}_j]}{\mathcal{P}_i^t \mathcal{P}_j^t \text{Tr}[\mathbf{\Gamma}_i] \text{Tr}[\mathbf{\Gamma}_j]}, \quad (2.53)$$

where $\mathcal{P}_i^t$ and $\mathcal{P}_j^t$ are the degree of temporal polarisation for each point i, j .

It is then possible to conclude that equation (2.53) is a metric that quantifies the similarity between the polarisation states present in an arbitrarily polarised field. Equation (2.53) even takes into account the contributions of temporal effects. Given that a temporally unpolarised field is also spatially unpolarised, equation (2.53) is undefined when $\mathcal{P}^t = 0$.

Note that for a fully temporally polarised beam ($\mathcal{P}_i^t = \mathcal{P}_j^t = 1$), equation (2.53) is equivalent to

$$\mathcal{P}_{i,j}^s = 1 - \frac{\text{Tr}[\mathbf{\Gamma}_{i,j}]^2}{\text{Tr}[\mathbf{\Gamma}_i] \text{Tr}[\mathbf{\Gamma}_j]}. \quad (2.54)$$

In conventional notation, the *degree of spatial polarisation* of a fully temporally polarised beam is then given by

$$\mathcal{P}^s(\mathbf{r}_i, \mathbf{r}_j, 0) = 1 - \frac{\text{Tr}[\mathbf{\Gamma}(\mathbf{r}_i, \mathbf{r}_j, 0)]^2}{\text{Tr}[\mathbf{\Gamma}(\mathbf{r}_i, \mathbf{r}_i, 0)] \text{Tr}[\mathbf{\Gamma}(\mathbf{r}_j, \mathbf{r}_j, 0)]}, \quad (2.55)$$

or since only spatial effects are being considered

$$\text{Tr}[\mathbf{\Gamma}(\mathbf{r}_i, \mathbf{r}_j, \tau)] = |\vec{E}^\dagger(\mathbf{r}_i, t) \cdot \vec{E}(\mathbf{r}_j, t - \tau)| \quad (2.56)$$

for $t = 0$, and hence equation (2.55) can be rewritten as

$$\mathcal{P}^s(\mathbf{r}_i, \mathbf{r}_j, 0) = 1 - \frac{|\vec{E}^\dagger(\mathbf{r}_i, t) \cdot \vec{E}(\mathbf{r}_j, t)|^2}{|\vec{E}(\mathbf{r}_i, t)|^2 |\vec{E}(\mathbf{r}_j, t)|^2} \quad (2.57a)$$

$$= 1 - \cos^2 \theta_{i,j}, \quad (2.57b)$$

where $\theta_{i,j}$ is the angle between the polarisation states i, j . It thus follows that

$$0 \leq \mathcal{P}^s \leq 1. \quad (2.58)$$

As opposed to the DOTP, here $\mathcal{P}^s = 0$ when the beam is fully spatially polarised, and $\mathcal{P}^s = 1$ when it is spatially unpolarised.

As a concluding remark, note the similarity between equation (2.55) and the degree of spatial coherence given by equation (2.18). Squaring equation (2.18) and substituting it in (2.55) for $\tau = 0$ yields

$$\mathcal{P}^s(\mathbf{r}_i, \mathbf{r}_j, 0) = 1 - \mu^2(\mathbf{r}_i, \mathbf{r}_j, 0). \quad (2.59)$$

The previous equation then suggests that the DOSP can also be measured using a two pinhole interferometer.

In summary, the DOTC, the DOSC, the DOTP and the DOSP were derived here with the aim of unifying the theory of polarisation and coherence. The DOTC and the DOSC estimate the temporal and spatial scalar contributions to the fluctuations of a beam, and the DOTP and the DOSP estimate the temporal and spatial polarisation contributions. Only in this context it is possible to solely quantify the fluctuations of a beam due to its frequency content (DOTC) and phase variations (DOSC). The temporal and spatial variations of the polarisation are then quantified using the DOTP and the DOSP, respectively.

2.3 Conclusions

This chapter has introduced the electric coherency matrix, the self coherency matrix and the polarisation matrix. After introducing these matrices, particular attention was paid to the degree of coherence and the degree of polarisation of a beam. A possible solution to conclude the ongoing discussion in the literature regarding polarisation and coherence was given. It was first proposed that the term “coherence” should refer only to the scalar properties of a beam; any spatial variations of the polarisation across the beam should be homogenised before measuring coherence. Only in this context is possible to guarantee that a measure exists to solely quantify temporal and spatial wavefront fluctuations. A novel metric to account for the spatial variations of the polarisation was introduced. This metric was termed the degree of spatial polarisation, and the quantity commonly known as the degree of polarisation was termed the degree of temporal polarisation. Additionally, the mathematical expression of the latter was derived in a novel way. It was shown that the resulting formula of the DOTP is independent of the dimensions of the field and

co-ordinate rotations.

The formulae given in this chapter account for a good theoretical introduction to polarimetry. In fact, these formulae are used to derive a formalism more suitable for the laboratory in the next chapter. A formalism capable of describing the interactions of partially polarised light and depolarising objects is given next.

Chapter 3

Polarimetry

This chapter reviews the notation used in this work for the measurement of polarisation and introduces the theory of the measurement itself. The first section introduces the Stokes vector, the Poincaré sphere and the Müller matrix. The second section focuses on a technique to measure polarisation known as division of amplitude polarimetry (DOAP). The second section also includes an overview of the condition number of a matrix, and a technique to improve it in the context of DOAP.

3.1 Müller–Stokes formalism

Müller and Stokes have developed a formalism that can account for depolarisation effects in both an object and a beam. The formalism is based on a 4×4 matrix termed the Müller matrix, and a 4×1 vector called the Stokes vector. The Müller matrix represents the polarisation changes induced by an object, and the Stokes vector describes the state of polarisation of the light. Both the Stokes vector and the Müller matrix are introduced next in that order.

3.1.1 Stokes vector

As shown by Fano[38], the elements of a Stokes vector $\vec{S}$ can be found using the polarisation matrix $\mathbf{\Gamma}$ introduced in section 2.1. If $\mathbf{\Gamma}$ is decomposed as a linear

combination of Hermitian trace-orthogonal matrices

$$\mathbf{\Gamma} = \frac{1}{2} \sum_{j=0}^3 s_j \boldsymbol{\sigma}_j, \quad (3.1)$$

then the $j = 0, 1, 2, 3$ elements of $\vec{S}$ are given by the coefficients s_j and $\boldsymbol{\sigma}_j$ represents the Pauli matrices and the identity matrix, denoted as

$$\begin{aligned} \boldsymbol{\sigma}_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ \boldsymbol{\sigma}_2 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \end{aligned} \quad (3.2)$$

If now the coefficients are written as $s_j = \text{Tr}[\mathbf{\Gamma} \boldsymbol{\sigma}_j]$, it then follows that

$$\vec{S} = \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} \gamma_{11} + \gamma_{22} \\ \gamma_{11} - \gamma_{22} \\ \gamma_{12} + \gamma_{21} \\ -i(\gamma_{12} - \gamma_{21}) \end{pmatrix} = \begin{pmatrix} \langle A_x^2 + A_y^2 \rangle \\ \langle A_x^2 - A_y^2 \rangle \\ \langle A_x A_y \cos(\Delta_{yx}) \rangle \\ \langle A_x A_y \sin(\Delta_{yx}) \rangle \end{pmatrix}, \quad (3.3)$$

where γ_{ij} are the $i, j = 1, 2$ elements of $\mathbf{\Gamma}$ (see section 3.1). Equation (3.3) represents the Stokes vector as it was once written by Jones[48], but originally derived by Stokes[49] in 1852 in a different yet equivalent notation. By writing the electric field as[10]

$$A_x e^{i\delta_x} = I_0 (\cos \theta \cos \epsilon - i \sin \theta \sin \epsilon), \quad (3.4a)$$

$$A_y e^{i\delta_y} = I_0 (\sin \theta \cos \epsilon + i \cos \theta \sin \epsilon), \quad (3.4b)$$

the Stokes vector derived by Stokes himself is given by

$$\begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = I_0 \begin{pmatrix} 1 \\ \mathcal{P}^t \cos(2\theta) \cos(2\epsilon) \\ \mathcal{P}^t \sin(2\theta) \cos(2\epsilon) \\ \mathcal{P}^t \sin(2\epsilon) \end{pmatrix}, \quad (3.5)$$

where $\mathcal{P}^t$ is the degree of temporal polarisation introduced in section 3.2.2 and is equivalent to

$$\mathcal{P}^t = \frac{\sqrt{s_1^2 + s_2^2 + s_3^2}}{s_0}. \quad (3.6)$$

The symbols θ and ϵ are the ellipsometric parameters shown in figure (3.1), where ϵ describes the ellipticity of the polarisation ellipse and θ is the angle of the major axis of the ellipse to a reference axis. Both quantities play an important role throughout this work.

The Stokes vector can be equivalently derived using the polarisation matrix and a set of irradiance measurements. The element s_0 is given by the total irradiance of the field. The rest of the elements are found by quantifying the irradiance contribution of 6 different polarisation states to the field $\mathbf{\Gamma}$. The element s_1 is found using a horizontal (h) and a linear vertical (v) analyser. The element s_2 is found using a linear analyser placed at $\pi/4$ and at $-\pi/4$ with respect to the horizontal axis. Finally s_3 is found using a right (r) and a left (l) circular analyser. It then follows from section 3.2.4, that Equation (2.50)¹ can be used to write the Stokes vector as

$$\begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} I \\ I_x - I_y \\ I_{-\pi/4} - I_{\pi/4} \\ I_l - I_r \end{pmatrix} = \begin{pmatrix} \text{Tr}[\mathbf{\Gamma}] \\ \text{Tr}[\mathbf{\Gamma}_v \mathbf{\Gamma}] - \text{Tr}[\mathbf{\Gamma}_h \mathbf{\Gamma}] \\ \text{Tr}[\mathbf{\Gamma}_{-\pi/4} \mathbf{\Gamma}] - \text{Tr}[\mathbf{\Gamma}_{\pi/4} \mathbf{\Gamma}] \\ \text{Tr}[\mathbf{\Gamma}_l \mathbf{\Gamma}] - \text{Tr}[\mathbf{\Gamma}_r \mathbf{\Gamma}] \end{pmatrix}, \quad (3.7)$$

The i th matrix $\mathbf{\Gamma}_i$ is constructed using the eigenvectors of the Jones matrix representing the horizontal, vertical, $\pi/4$, $-\pi/4$, left and right circular analysers, respectively.

Another tool of great use in this work is a sphere of unit radius known as the Poincaré sphere[50]. This sphere, shown in figure (3.1), is used to graphically represent a Stokes vector. If $\vec{S}$ is a state of polarisation normalised by the total irradiance, it follows from equation (3.5) that the s_1, s_2 and s_3 components of $\vec{S}$ can be used to define a point within the sphere representing $\vec{S}$. Equivalently 2θ , 2ϵ and $\mathcal{P}$ can also be used to plot $\vec{S}$ within the sphere. In this case 2θ is the angle along the equator, 2ϵ is the azimuthal angle, and $\mathcal{P}$ is the distance to the centre of the sphere. For example, right and left circular polarisation states ($\epsilon = \pm\pi/4$) lie respectively on

¹ $I_{(i)}^p = \text{Tr}[\mathbf{\Lambda}_{(i)} \mathbf{\Lambda}] - \text{Tr}[\mathbf{\Gamma}_{(i)} \mathbf{\Gamma}].$

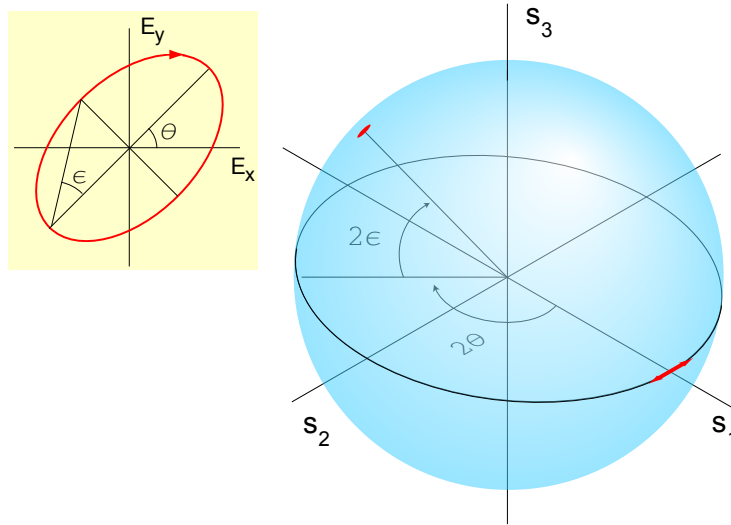


Figure 3.1: Ellipse of polarisation and the Poincaré sphere.

the top and bottom zenith, and all the linear states ($\epsilon = 0$) lie on the equator; the rest of the surface of the sphere represents elliptical states. The origin represents unpolarised light, i.e. $\mathcal{P}^t = 0$, fully polarised states lie on the surface of the sphere, i.e. $\mathcal{P}^t = 1$, and partially polarised states are represented by the inside of the sphere.

Note that the magnitude of the irradiance is lost in the Poincaré representation. Tyo[51] proposed an alternative graphical representation using a conical figure to include the irradiance. Nowadays however, it is more convenient to plot $\vec{S}$ on the Poincaré sphere with a marker² instead of a point and assign a colour scale to it according to the value of s_0 .

3.1.2 Müller matrix

The polarisation interaction between an electromagnetic field and a potentially depolarising object can be characterised by

$$\vec{S}_{out} = \mathbf{M}\vec{S}_{in}, \quad (3.8)$$

where $\mathbf{M}$ is a real positive definitive 4×4 matrix known as a Müller matrix. $\mathbf{M}$ contains all the 2D information of how a polarisation element changes a given incident

²For example a circle or an asterisk

polarisation state $\vec{S}_{in}$ to an output state $\vec{S}_{out}$. A non-depolarising Müller matrix can be written in general as

$$\mathbf{M} = \mathbf{R}(\vartheta) \begin{pmatrix} 1 & -\cos 2\Psi & 0 & 0 \\ -\cos 2\Psi & 1 & 0 & 0 \\ 0 & 0 & \sin 2\Psi \cos \Delta & \sin 2\Psi \sin \Delta \\ 0 & 0 & -\sin 2\Psi \sin \Delta & \sin 2\Psi \cos \Delta \end{pmatrix} \mathbf{R}(-\vartheta), \quad (3.9)$$

where Ψ and Δ are known as the ellipsometric parameters[10] and commonly termed diattenuation and retardance, respectively. The tangent of Ψ gives the ratio of the attenuation between the amplitudes of the s- and p-polarisations, and Δ represents the phase difference induced between the s- and p-polarisations. The matrix $\mathbf{R}(\vartheta)$ is a 4×4 rotation matrix given by

$$\mathbf{R}(\vartheta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\vartheta & \sin 2\vartheta & 0 \\ 0 & -\sin 2\vartheta & \cos 2\vartheta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (3.10)$$

where ϑ is the angle of rotation in respect to a reference axis. Any depolarisation effects can be considered in the formulation by multiplying equation (3.9) with the Müller matrix of a depolariser as shown next.

Depending on the nature of the induced polarisation changes, a polarisation element can be found in the form of a pure retarder, a pure diattenuator, a depolariser or in a combination of these three. A composite Müller matrix can be found by multiplying the individual matrices in the inverse order at which light is passing through. Müller matrices however, are not easy to interpret. Lu and Chipman have developed a Müller matrix interpretation method based on polar decomposition³[52] to solve this problem. The following list summarises the decomposition while introducing the polarisation elements listed above.

Diattenuator. This element shows a preferential transmission or reflection to a given polarisation state. The eigenvector of a diattenuator matrix $\mathbf{M}_D$ corresponding

³Decomposition of a matrix into the product of an unitary matrix and a positive-semidefinite Hermitian matrix

to the largest eigenvalue is known as the diattenuation vector $\vec{D}$. The latter can also be found as

$$\vec{D} = \frac{1}{m_{00}} \begin{pmatrix} m_{01} \\ m_{02} \\ m_{03} \end{pmatrix}, \quad (3.11)$$

where m_{ij} is the ij th element of $\mathbf{M}$ for $i, j = 0 \dots 3$. A diattenuator also has an intrinsic property known as *polarisance*. This is the ability to polarise unpolarised light. The polarisance vector $\vec{P}$ is represented by

$$\vec{P} = \frac{1}{m_{00}} \begin{pmatrix} m_{10} \\ m_{20} \\ m_{30} \end{pmatrix}. \quad (3.12)$$

Both $\vec{P}$ and $\vec{D}$ can be plotted on the Poincaré sphere, and their magnitudes give a measure of the polarisance and the diattenuation present in the element. When $|\vec{D}| = 1$, the diattenuator matrix is given by

$$\mathbf{M}_D = m_{00} \begin{pmatrix} 1 & \vec{D}^T \\ \vec{P} & \vec{P}\vec{D}^T \end{pmatrix}, \quad (3.13)$$

otherwise,

$$\mathbf{M}_D = m_{00} \begin{pmatrix} 1 & \vec{D}^T \\ \vec{P} & \mathbf{0} \end{pmatrix}, \quad (3.14)$$

where $\mathbf{0}$ is a 3×3 matrix of zeroes.

Depolariser. This is an element with the property of modifying the degree of polarisation in a beam. This property is found in inhomogeneous refractive index materials, time changing processes with intervals shorter than the integration time, scatterers, and appears when electronic noise is present in the measurement. The Müller matrix of a depolariser is extracted from an arbitrary Müller matrix $\mathbf{M}$ using the 3×3 submatrix $\mathbf{m}'$ given by

$$\mathbf{M}\mathbf{M}_D^{-1} = \mathbf{M}' = \begin{pmatrix} 1 & \dots \\ \vdots & \mathbf{m}' \end{pmatrix}, \quad (3.15)$$

where the submatrix $\mathbf{m}'$ is then decomposed using singular value decomposition (SVD) as

$$\mathbf{m}' = \sqrt{\lambda_1} \vec{v}_1 \vec{u}_1^T + \sqrt{\lambda_2} \vec{v}_2 \vec{u}_2^T + \sqrt{\lambda_3} \vec{v}_3 \vec{u}_3^T, \quad (3.16)$$

where $\vec{v}_i$ and $\vec{u}_i$ are the $i = 1, 2, 3$ orthogonal vectors resultant from the SVD, and λ_i are the $i = 1, 2, 3$ eigenvalues of $\mathbf{m}'(\mathbf{m}')^T$. The depolariser matrix is given by

$$\mathbf{M}_\Delta = \begin{pmatrix} 1 & \vec{0}^T \\ \vec{P}_\Delta & \mathbf{m}_\Delta \end{pmatrix}, \quad (3.17)$$

where $\vec{0}$ is a 3×1 vector of zeroes and

$$\mathbf{m}_\Delta = \pm(\sqrt{\lambda_1} \vec{v}_1 \vec{v}_1^T + \sqrt{\lambda_2} \vec{v}_2 \vec{v}_2^T + \sqrt{\lambda_3} \vec{v}_3 \vec{v}_3^T). \quad (3.18)$$

The sign is chosen according to the sign of the determinant of $\mathbf{m}'$. In equation (3.17)

$$\vec{P}_\Delta = \frac{\vec{P} - \mathbf{m}\vec{D}}{1 - D^2}, \quad (3.19)$$

where $\mathbf{m}$ is the bottom right 3×3 submatrix of the original Müller matrix $\mathbf{M}$.

Retarder. This element introduces a phase shift among the components. The Müller matrix of a retarder $\mathbf{M}_R$, corresponds to a rotation of the Stokes vector in the Poincaré sphere. The two real eigenvectors of $\mathbf{M}_R$ are known as the fast and slow axis of the optical element. The extraction of a retarder matrix from an arbitrary Müller matrix $\mathbf{M}$ depends on the number of non zero eigenvalues of $\mathbf{m}'$; that is equation (3.16). When all eigenvalues are zero the retarder matrix is the identity matrix. When two of the eigenvalues are zero the retardance is given by

$$R = \cos^{-1} \vec{v} \cdot \vec{u}, \quad (3.20)$$

and the fast axis is calculated from

$$\vec{R} = \frac{\vec{v} \times \vec{u}}{|\vec{v} \times \vec{u}|}, \quad (3.21)$$

where $\vec{v}$ and $\vec{u}$ are the resultant vectors from the SVD of $\mathbf{m}'$ that correspond to the non-zero eigenvalues. In the case when only one eigenvalue is zero the retardance

and the elements a_i of the fast axis are given respectively by

$$R = \cos^{-1}\left(\frac{\text{Tr}[\mathbf{m}_R]}{2} - \frac{1}{2}\right), \quad (3.22)$$

and

$$a_i = \frac{1}{2 \sin(R)} \sum_{j,k=1}^3 \epsilon_{ijk} (\mathbf{m}_R)_{j+1,k+1}, \quad (3.23)$$

where ϵ_{ijk} is the Levi-Civita permutation symbol, and

$$\mathbf{m}_R = \vec{v}_1 \vec{u}_1^T + \vec{u}_2 \vec{v}_2^T + \frac{\vec{v}_1 \times \vec{v}_2}{|\vec{v}_1 \times \vec{v}_2|} \frac{(\vec{u}_1 \times \vec{u}_2)^T}{|\vec{u}_1 \times \vec{u}_2|}. \quad (3.24)$$

The vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{u}_1$ and $\vec{u}_2$ are the resultant vectors from the SVD of $\mathbf{m}'$ that correspond to the non-zero eigenvalues. Finally when all the eigenvalues of $\mathbf{m}'$ are non-zero

$$\mathbf{m}_R = \pm(\vec{v}_1 \vec{u}_1^T + \vec{u}_2 \vec{v}_2^T + \vec{u}_3 \vec{v}_3^T), \quad (3.25)$$

and the sign is chosen the same way as equation (3.18). In this case equations (3.22) and (3.23) are used to find the retardance and the fast axis, respectively.

Having understood what Müller matrices and Stoke vectors are, it is possible to proceed to the introduction of the notation used later in the experiments. The next section presents polarimetry using basic matrix algebra.

3.2 Müller–Stokes polarimetry

Polarimetry refers to the measurement of the polarisation of an electromagnetic field. In particular the measurement is called Stokes polarimetry (SP) or Stokesmetry when one measures a single polarisation state. When measuring the polarisation properties of an object, the measurement is called Müller matrix polarimetry (MMP). Figure (6.1) shows the block diagram of a typical setup capable of performing either of the measurements. The setup is formed by a polarisation state generator (PSG), an object described by $\mathbf{M}$, a polarisation state analyser (PSA) and a detection module $\mathbf{D}$.

Polarimetry can be expressed as a matrix algebra problem[47]. By having control over both the input polarisation states and the output irradiance readings, it is

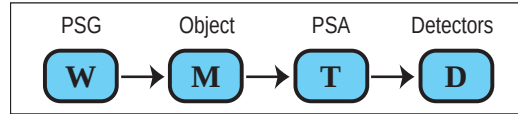


Figure 3.2: Block diagram of a polarimeter.

possible to create an overdetermined system of linear equations. In the case of MMP, since the system of equations has 16 unknowns, at least four polarisation states have to be input to the system and an equivalent set of detected irradiances has to be read out. For SP however, given that the system of equations has 4 unknowns, only one polarisation state is input to the setup and a single set of detected irradiances is read out. The number of unknowns can be further reduced to the number of Stokes elements required, but in this work only complete polarimetry⁴ is considered.

The input is controlled by the polarisation state generator (PSG), and the output by the polarisation state analyser (PSA). If n is the required number of polarisation states, the PSG is denoted by

$$\mathbf{W} = (\vec{S}_1, \dots, \vec{S}_n), \quad (3.26)$$

where $\vec{S}_n$ is the Stokes vector representing the n th polarisation state generated. With regards to the PSA, the chosen method to analyse the polarisation is a technique known as division-of-amplitude polarimetry (DOAP) (see for example [53]). In this technique, the beam to be analysed is split in at least 4 paths⁵, where each path has a different polarisation analyser⁶ followed by a photodetector. The irradiances measured in the photodetectors are then related to a Stokes vector by a non-singular real matrix $\mathbf{T}$, known as the instrument matrix. The matrix $\mathbf{T}$ corresponds in fact to a linear mapping between the polarisation space $\mathbb{P}$ and the irradiance space $\mathbb{I}$. It thus follows that $\mathbf{T}$ can be constructed from the first row of the Müller matrices representing each analyser path. If q is the number of analyser paths, the PSA is then denoted by a $q \times 4$ matrix $\mathbf{T}$.

In short, the PSG generates a number of polarisation states which are possibly modified by a Müller matrix $\mathbf{M}$ and then transformed by $\mathbf{T}$ into irradiances.

⁴The measurement of all the elements of the Stokes vector.

⁵Condition for complete polarimetry

⁶In general, an ellipsometric surface followed by a polariser

Mathematically this is expressed as

$$\mathbf{D} = \mathbf{T}\mathbf{M}\mathbf{W}, \quad (3.27)$$

where $\mathbf{D}$ is a $q \times n$ matrix formed by the q detected irradiances that correspond to the n generated polarisation states. For example in the case of a Stokes vector measurement, $\mathbf{D}$ is a $q \times 1$ matrix vector. When $\mathbf{T}$ and $\mathbf{W}$ are known, the Müller matrix is given by

$$\mathbf{M} = \mathbf{T}^+\mathbf{D}\mathbf{W}^+, \quad (3.28)$$

where $^+$ denotes pseudoinverse⁷ since $\mathbf{T}$ and $\mathbf{W}$ are not in general square matrices. If only $\mathbf{T}$ is known, it is not possible to measure $\mathbf{M}$. It is possible however to find a Stokes vector $\vec{S}_{out}$. Assuming $\vec{S}_{out}$ results from

$$\vec{S}_{out} = \mathbf{M}\vec{S}_{in}, \quad (3.29)$$

equation (3.27) is written as

$$\mathbf{D} = \mathbf{T}\mathbf{M}\vec{S}_{in} \quad (3.30)$$

and hence

$$\vec{S}_{out} = \mathbf{T}^+\mathbf{D}. \quad (3.31)$$

The matter of finding experimentally $\mathbf{T}$ and $\mathbf{W}$ is addressed in chapter 6. It is necessary to first design and optimise theoretically both $\mathbf{T}$ and $\mathbf{W}$. For this purpose the analyser paths and generated polarisation states have to be chosen in such a way that equation (3.27) is well-conditioned. As shown next, the error propagation due to the matrix inversions in equations (3.28) and (3.31) is dependent on the conditioning of the system.

3.2.1 Matrix conditioning

For a system of linear equations to be well-conditioned it must comply with the following requirements[54],

- A solution to the system exists.

⁷Generalisation of the matrix inverse operation[52]

- The solution is unique.
- The solution depends continuously on the data, in some reasonable topology.

However these requirements can be objectively assessed using the condition number κ defined[55, 56] as

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^+\|_2, \quad (3.32)$$

where $\mathbf{A}$ is an arbitrary matrix and in this work $\|\cdot\|_2$ denotes the spectral norm[47], i.e.

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}}, \quad (3.33)$$

$\lambda_{\max}$ being the largest eigenvalue of $\mathbf{A}^\dagger \mathbf{A}$.

There is a method to visualise the conditioning of both $\mathbf{T}$ and $\mathbf{W}$. The rows of $\mathbf{T}$ and the columns of $\mathbf{W}$, are plotted as vectors in a Poincaré sphere each. The end-point of the vectors ideally should fall as far apart as possible from each other, i.e. as orthogonal as possible[57].

In order to investigate now the implications of κ in a polarimetric measurement, consider a Stokes vector measurement,

$$\vec{D} = \mathbf{T}\vec{S}. \quad (3.34)$$

If $\mathbf{T}$ is known, equation (3.34) is a system of linear equations with 4 unknowns. To assess the confidence in the solution of such a system, a perturbation is introduced to the Stokes vector. Letting $\Delta\vec{S}$ be a deviation from $\vec{S}$, equation (3.34) can be written using the Cauchy–Schwarz inequality⁸ as

$$\epsilon(\vec{D}) \leq \kappa(\mathbf{T})\epsilon(\vec{S}), \quad (3.35)$$

where ϵ is the relative error defined as[58, 59]

$$\epsilon(\vec{S}) = \frac{\|\Delta\vec{S} - \vec{S}\|}{\|\vec{S}\|}, \quad (3.36)$$

⁸ $|\langle x, y \rangle| \leq \|x\| \|y\|$

and $\|\cdot\|$ denotes the L^2 -norm. It thus follows from equation (3.35) that the relative error in $\vec{D}$ is upper bounded by the error in $\vec{S}$ magnified κ times. Consequently, the error magnification reduces the confidence of the measurement and also the significant digits that can be trusted. A quantification of the latter is given by [60]

$$\epsilon(\vec{D}) \leq \frac{1}{2}10^{-\xi}. \quad (3.37)$$

where ξ are the trustworthy significant digits. It is possible to say that equation (3.37) is a metric to quantify the resolution of the system to different polarisation states. Given in addition that the signal to noise ratio (SNR) in a system is the inverse of the relative error, equation (3.35) can be written equivalently as

$$\text{SNR}_{\vec{S}} \leq \kappa(\mathbf{T})\text{SNR}_{\vec{D}}. \quad (3.38)$$

For a given $\text{SNR}_{\vec{D}}$, the SNR in $\vec{S}$ is bounded by $\kappa(\mathbf{T})\text{SNR}_{\vec{D}}$.

In summary, a small relative error (high SNR) in the PSG and a PSA with small condition number, results in a measurement with more trustworthy significant digits. As shown by Savenkov[61], in the case of polarimetry and under the L^2 norm definition the theoretical minimum of $\kappa(\mathbf{T})$ is $\sqrt{3}$. It thus follows that

$$\xi \leq \frac{\log \sqrt{3} + \log \epsilon(\vec{S})}{\log 5}. \quad (3.39)$$

Another parameter that could be taken into account is ς . This parameter appears when the Cauchy–Schwarz inequality is converted to an equality as,

$$\|\mathbf{T}^{-1}\vec{D}\| = \varsigma\|\mathbf{T}^{-1}\| \|\vec{D}\|. \quad (3.40)$$

It is then possible to derive an analogous equation to (3.35) as

$$\frac{\|\Delta\vec{S} - \vec{S}\|}{\|\vec{S}\|} \leq \varsigma^{-1} \frac{\|\Delta\vec{D} - \vec{D}\|}{\|\vec{D}\|}. \quad (3.41)$$

If ς is near unity, $\vec{S}$ is insensitive to perturbations regardless of the value of κ .

The SNR and κ are set by the electronics and $\mathbf{T}$, respectively. The condition number however, can still be improved by balancing the columns and rows of the instrument matrix. This improvement can be shown assuming that the instrument matrix can be written as

$$\mathbf{T} = \mathbf{T}_{gain} \mathbf{T}_{pol}, \quad (3.42)$$

where $\mathbf{T}_{gain}$ is a diagonal matrix with the gain of the detectors, and $\mathbf{T}_{pol}$ is an instrument matrix. The substitution of the previous equation in (3.33) yields

$$\|\mathbf{T}\| \cdot \|\mathbf{T}^{-1}\| = \|\mathbf{T}_{gain} \mathbf{T}_{pol}\| \cdot \|(\mathbf{T}_{gain} \mathbf{T}_{pol})^{-1}\|. \quad (3.43)$$

Using the Cauchy–Schwartz inequality, it is then possible to write

$$\|\mathbf{T}\| \cdot \|\mathbf{T}^{-1}\| \leq \|\mathbf{T}_{gain}\| \cdot \|\mathbf{T}_{pol}\| \cdot \|\mathbf{T}_{pol}^{-1}\| \cdot \|\mathbf{T}_{gain}^{-1}\|, \quad (3.44)$$

or equivalently

$$\kappa(\mathbf{T}) \leq \kappa(\mathbf{T}_{gain}) \cdot \kappa(\mathbf{T}_{pol}). \quad (3.45)$$

It thus follows that if $\mathbf{T}_{pol}$ has already been optimised, the condition number of $\mathbf{T}$ then depends entirely on $\mathbf{T}_{gain}$. That means that the condition number can be improved by changing the gain in the detectors. To find now the optimal gain the Van der Sluis equalising theorem[60, 62] is used. First let

$$\mathbf{B}_C = \text{diag}(\|\mathbf{T}(:, j)\|)^{-1} \quad (3.46)$$

and

$$\mathbf{B}_R = \text{diag}(\|\mathbf{T}(i, :)\|)^{-1} \quad (3.47)$$

be diagonal matrices. In equation (3.46), the elements on the diagonal correspond to the norm of the respective row of $\mathbf{T}$, and in equation (3.47) the diagonal elements are given by the norm of the column of $\mathbf{T}$. The Van der Sluis theorem states that there is a pair of diagonal matrices $\mathbf{B}_R$ and $\mathbf{B}_C$ that can lower the condition number

of $\mathbf{T}$. This is written explicitly as

$$\kappa(\mathbf{B}_R \mathbf{T}) \leq \sqrt{n} \cdot \min_{\mathbf{B}} [\kappa(\mathbf{B} \mathbf{T})], \quad (3.48)$$

$$\kappa(\mathbf{T} \mathbf{B}_C) \leq \sqrt{n} \cdot \min_{\mathbf{B}} [\kappa(\mathbf{T} \mathbf{B})], \quad (3.49)$$

where $\min_{\mathbf{B}} [\kappa(\cdot)]$ refers to the smallest condition number achieved by any diagonal matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$.

A couple of examples are now given in order to demonstrate the improvement of κ . Consider the experimental instrument matrix

$$\mathbf{T}_1 = \begin{pmatrix} 0.4825 & -0.4771 & 0.0859 & 0.0851 \\ 0.5175 & 0.4771 & -0.0859 & -0.0880 \\ 0.2531 & -0.0085 & 0.2364 & 0.0446 \\ 0.2543 & -0.0637 & -0.0194 & -0.2493 \end{pmatrix} \quad \kappa(\mathbf{T}_1) = 4.35.$$

Applying Van der Sluis' theorem by steps, first balancing only with $\mathbf{B}_R$ yields

$$\mathbf{B}_R \mathbf{T}_1 = \begin{pmatrix} 0.7001 & -0.6922 & 0.1246 & 0.1235 \\ 0.7242 & 0.6677 & -0.1202 & -0.1232 \\ 0.7246 & -0.0244 & 0.6768 & 0.1276 \\ 0.7019 & -0.1758 & -0.0537 & -0.6881 \end{pmatrix} \quad \kappa(\mathbf{B}_R \mathbf{T}_1) = 3.7919,$$

or only balancing with $\mathbf{B}_C$ gives

$$\mathbf{T}_1 \mathbf{B}_C = \begin{pmatrix} 0.6082 & -0.7039 & 0.3221 & 0.3025 \\ 0.6523 & 0.7039 & -0.3221 & -0.3129 \\ 0.3191 & -0.0126 & 0.8872 & 0.1585 \\ 0.3205 & -0.0940 & -0.0729 & -0.8863 \end{pmatrix} \quad \kappa(\mathbf{T}_1 \mathbf{B}_C) = 2.1855,$$

and finally balancing with both yields

$$\mathbf{B}_R \mathbf{T}_1 \mathbf{B}_C = \begin{pmatrix} 0.8825 & -1.0213 & 0.4674 & 0.4389 \\ 0.9129 & 0.9852 & -0.4508 & -0.4379 \\ 0.9134 & -0.0360 & 2.5397 & 0.4537 \\ 0.8848 & -0.2594 & -0.2013 & -2.4465 \end{pmatrix} \quad \kappa(\mathbf{B}_R \mathbf{T}_1 \mathbf{B}_C) = 2.8659,$$

Clearly balancing with $\mathbf{B}_C$ improves κ by more than half.

As a second example, consider the theoretical instrument matrix of a broadband polarimeter

$$\mathbf{T}_2 = \begin{pmatrix} 1 & -0.5774 & 0.8165 & 0 \\ 1 & -0.5774 & -0.8165 & 0 \\ 1 & 0.5774 & 0 & 0.8165 \\ 1 & 0.5774 & 0 & -0.8165 \end{pmatrix}, \quad \kappa(\mathbf{T}_2) = 1.7321.$$

Repeating the previous procedure, balancing $\mathbf{T}_2$ with $\mathbf{B}_R$ results in

$$\mathbf{B}_R \mathbf{T}_2 = \begin{pmatrix} 0.7071 & -0.4082 & 0.5773 & 0 \\ 0.7071 & -0.4082 & -0.5773 & 0 \\ 0.7071 & 0.4082 & 0 & 0.5773 \\ 0.7071 & 0.4082 & 0 & -0.5773 \end{pmatrix}, \quad \kappa(\mathbf{B}_R \mathbf{T}_2) = 1.7321,$$

or with $\mathbf{B}_C$ gives

$$\mathbf{T}_2 \mathbf{B}_C = \begin{pmatrix} 0.5000 & -0.5000 & 0.7071 & 0 \\ 0.5000 & -0.5000 & -0.7071 & 0 \\ 0.5000 & 0.5000 & 0 & 0.7071 \\ 0.5000 & 0.5000 & 0 & -0.7071 \end{pmatrix}, \quad \kappa(\mathbf{T}_2 \mathbf{B}_C) = 1.0000,$$

and finally with both yields

$$\mathbf{B}_R \mathbf{T}_2 \mathbf{B}_C = \begin{pmatrix} 0.3536 & -0.3536 & 0.5000 & 0 \\ 0.3536 & -0.3536 & -0.5000 & 0 \\ 0.3536 & 0.3536 & 0 & 0.5000 \\ 0.3536 & 0.3536 & 0 & -0.5000 \end{pmatrix}, \quad \kappa(\mathbf{B}_R \mathbf{T}_2 \mathbf{B}_C) = 1.0000.$$

The previous results show how it is possible to lower the theoretical limit of κ for a polarimeter. If this was the case, the relative error of the measurement would not depend on the instrument itself anymore, but only on the SNR of the detection and hence on the smallest element of $\mathbf{B}_C$. This conclusion however, is only true if it is actually possible to implement $\mathbf{B}_C$ and $\mathbf{B}_R$ in the measurement.

Left multiplying $\mathbf{T}$ with $\mathbf{B}_R$ is equivalent in the laboratory to modify the gain in the detectors or to throw away light. Both cases however, can increase the inherent noise of the measurement. It is thus recommended to do a noise analysis before implementing the method. The fact of right multiplying $\mathbf{T}$ with a diagonal matrix was not found here to be possible. Mathematically the matrix $\mathbf{T}$ cannot be multiplied before its inversion due to the matrix order in equation (3.34), and experimentally it was not possible to find a suitable optical element that could perform such balancing.

3.3 Conclusion

This chapter has reviewed the concepts of the Stokes vector, the Poincaré sphere and the Müller matrix. Additionally the Lu–Chipman decomposition method was described while introducing the Müller matrices of a diattenuator, a retarder and a depolariser. To describe the measurement of Stokes vectors and Müller matrices, the chapter introduced a particular case of Müller-Stokes polarimetry, namely division of amplitude polarimetry.

The combination of polarimetry and confocal microscopy results in confocal micropolarimetry, i.e. the measurement of polarisation in a high numerical aperture space. The next chapter introduces and explains the basic properties of this relatively new technology.

Chapter 4

Polarisation imaging

This chapter is devoted to the merger of high NA confocal microscopy and polarimetry. The resulting technique is termed here confocal micropolarimetry. The chapter goes through the advantages and constraints of the system in a rigorous and an intuitive manner. The role of the confocal detector is reviewed and the imaging parameters that can be used with this apparatus are introduced. The theoretical polarisation PSF of the system is shown and an illumination engineering method is developed to exploit geometrical and physical properties of objects.

4.1 Confocal micropolarimetry

Confocal micropolarimetry is a combination of scanning confocal microscopy[6] and polarimetry. It was seen in section 3.2 that polarimetry is a measurement method where a single polarisation state is assigned to a homogenous field, i.e. plane waves. As also seen in section 3.2, confocal microscopy is a technique in which an image is created point by point by scanning. These two techniques combined create a method that images the state of polarisation of the light. Work has been done in this area[63, 64, 65], however, the research reported in this thesis presents an experimental setup capable of taking polarimetric images in three dimensions with nanometre resolution for the first time.

The complete experimental setup is shown in figure (5.1). In this setup the light emitted by the laser passes through a polarisation state generator (PSG) to then be split 50 : 50 and focused onto a pinhole. After passing through the pinhole the

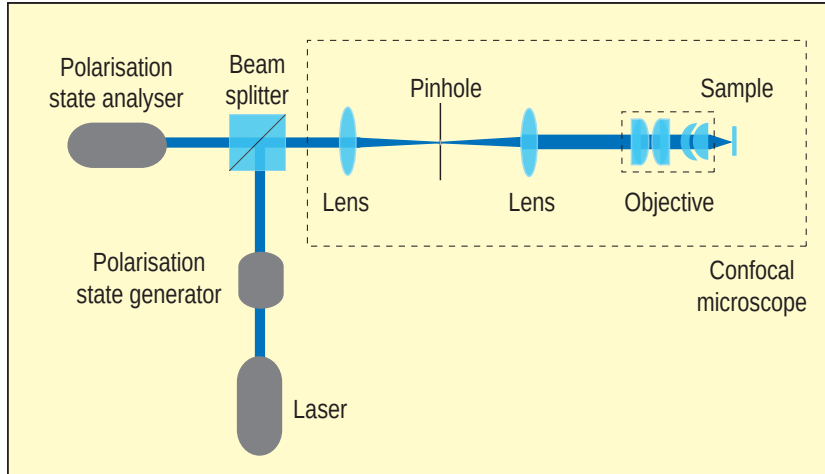


Figure 4.1: Experimental setup

light is collimated, and then focused onto a sample by a high NA lens. In the return path the light is focused again onto the pinhole, the so-called confocal plane, where a portion of the light is filtered out and the rest is collimated to reach a polarisation state analyser (PSA). The light is then split into four analyser paths (see section 3.2) and finally spatially and temporally integrated by the photodetectors. By either stage- or beam-scanning the sample, the recorded irradiances in the PSA are related to the stored x, y, z positions yielding a 3D polarisation image of the sample.

An important characteristic of confocal micropolarimetry is the spatial filtering property of the system. The pinhole filters the unwanted polarisation features in a beam that appear due to refraction. Take for example the case when the rays within a beam are incident at an angle greater than $\approx 10^\circ$ onto an air-glass interface of refractive index 1.5. In this case the polarisation across the beam starts being affected by the Fresnel coefficients of the interface, especially the polarisation of the marginal rays. A ray incident at 10° is equivalent to a lens of 0.17 NA. If the optical elements in the setup are not polarisation corrected¹[66] and the NA is greater than 0.18, one possible way to eliminate the unwanted rays is using confocal detection.

If a system is polarisation corrected[66], the pinhole still plays an important role. When a beam is collimated by a high NA lens the polarisation components of the field are mixed[12] according to the matrices given in section 1.3. This mixing results

¹The Fresnel coefficients of the optical element are ≈ 0

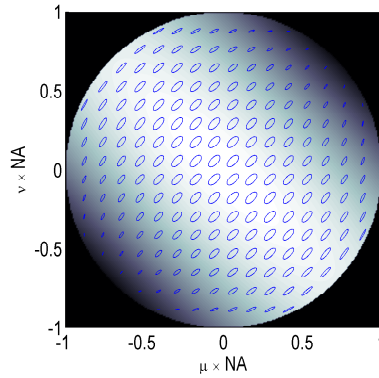


Figure 4.2: Polarisation distribution at the back focal plane of a lens with $NA = 1$.

in a field with a polarisation distribution as shown in figure (4.2). The colour scale depicts irradiance, the ellipses picture the state of polarisation and the magnitude of the major axis of the ellipses corresponds to the magnitude of the field at that ellipse. In this example the figure shows the field at the back focal plane of an objective lens of $NA=1$ resulting from collimating the field of a dipole excited with a polarisation state of $\theta = \pi/4$ and $\epsilon = -\pi/12$. Note how the polarisation across the beam is different from the input apart from the centre of the distribution, and how both the magnitude of the field and the irradiance changes. This occurs because the collimation of a beam is equivalent to an azimuthal projection² of the field.

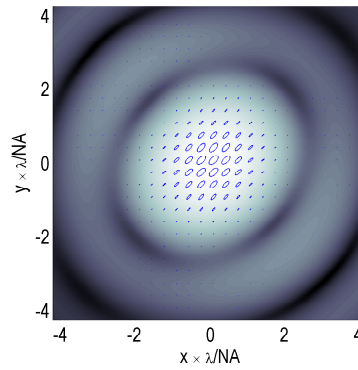


Figure 4.3: Polarisation distribution at the confocal plane.

If the field in figure (4.2) is focused, the resulting field is shown in figure (4.3). In this figure the gray scale represents the irradiance in a logarithmic scale, the ellipses depict again the state of polarisation and the magnitude of the major axis of the

²Projection from 3D to 2D in which the azimuth of all points are shown correctly with respect to the centre.

ellipses is proportional to the magnitude of the field. Note that the focused field also contains a polarisation distribution. Depending on the size of the pinhole, a portion of the distribution passes through it to then be analysed by the PSA. This suggests that the polarisation measurement can be affected by the size of the pinhole.

4.1.1 Pinhole size dependance

In theory it is possible to assume that the pinhole in the imaging system is infinitesimally small and thus analyse the field point by point. In practice however, the pinhole is finite in size. A finite part of the field passes through the pinhole to then be detected incoherently by the PSA. Even though the field at the pinhole can be detected coherently (see section 1.4), the light is analysed incoherently after the pinhole. Consequently the measurement of a fully temporally polarised beam may result as temporally unpolarised if the beam has a spatial polarisation distribution (see section 2.2.3). If for example half of the throughput field is fully vertically polarised and the other half is fully horizontally polarised, the field appears to be temporally unpolarised.

Assuming polarisation corrected optics, the aim is now to characterise the size dependance of the polarisation imaging system on the confocal detector. This is done by defining the extinction ratio and the degree of purity.

Extinction ratio

The extinction ratio is a metric commonly defined as the ratio between the total irradiance of the field and the irradiance resulting from a spatially homogeneous analyser. This metric is a special case of the degree of spatial polarisation since it analyses a single polarisation state (as opposed to all of them, see section 2.2.3). Practically speaking the extinction ratio quantifies the influence of spatial polarisation variations on the degree of temporal polarisation.

Figure (4.4) shows the extinction ratio for different pinhole sizes when imaging an excited free dipole³. The dipole is excited by a polarisation state of azimuth angle $\theta = \pi/4$ and ellipticity $\epsilon = -\pi/12$ (see figure 4.3). It can be seen from figure (4.4) that the extinction ratio improves as the size of the pinhole decreases.

³Dipole whose moment aligns to the incident field

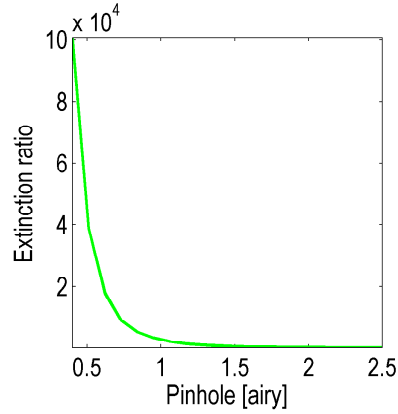


Figure 4.4: Extinction ratio for different pinhole sizes resulting from the image of a dipole

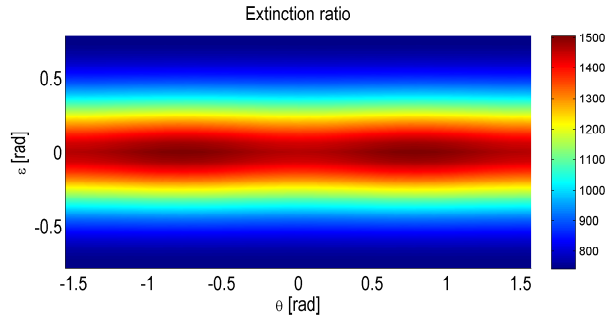


Figure 4.5: Extinction ratio (colour scale) for different polarisation illuminations and for a pinhole of one Airy unit in size.

To further study the system, figure (4.5) shows the extinction ratio for the system when the size of the pinhole is kept constant at one Airy unit (A.U.) and the input polarisation is changed. The ellipticity (ϵ) and the azimuth angle (θ) of the input states are represented by the vertical and horizontal axes, respectively. Note that this figure represents an unwrapped version of the Poincaré sphere.

It can be concluded from figure (4.5) that a pinhole of 1 A.U. has a greater effect on linearly polarised light than to circular. The reason for this is clarified later in section 4.2.2 after showing the Müller matrix corresponding to a dipole.

Degree of purity

The degree of purity is a novel metric introduced here to account for the difference between an expected and a measured polarisation state. This metric quantifies the influence of spatial polarisation variations in the measurement of a Stokes vector.

Since a polarisation state can be defined by the ellipticity (ϵ) and the azimuth angle (θ), two expressions are used to quantify the degree of purity. By using the law of cosines, the azimuthal degree of purity (ADOP) $\Delta\theta$ and the elliptical degree of purity (EDOP) $\Delta\epsilon$ are defined as

$$\Delta\theta = \sqrt{1 - \cos(\theta_{in} - \theta_{out})} \quad (4.1a)$$

$$\Delta\epsilon = \sqrt{1 - \cos(\epsilon_{in} - \epsilon_{out})}, \quad (4.1b)$$

where θ_{in} and ϵ_{in} are the ellipsometric parameters of the input polarisation state of the system, and θ_{out} and ϵ_{out} are those of the measured polarisation state. When both $\Delta\theta$ and $\Delta\epsilon$ are zero, the spatial polarisation variations of the field do not affect the measurement of a polarisation state.

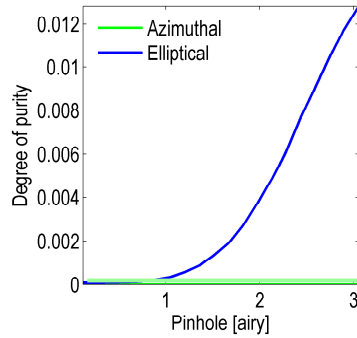


Figure 4.6: Degree of purity due to an excited dipole for different pinhole size

Figure (4.6) shows the ADOP and the EDOP corresponding to the image of a dipole when using pinholes of different sizes. The dipole is excited by a polarisation state of $\theta = \pi/4$ and $\epsilon = -\pi/12$. Interestingly in this particular case, the EDOP (blue) changes slightly while the ADOP (green) stays constant at zero. The greatest value of the EDOP corresponds to 0.5° . This relatively small influence on the measurement appears to be general. Figure (4.7) shows once more the EDOP (left) and the ADOP (right), but for a pinhole of 1 Airy in size and now varying the input polarisation. The vertical axis in both plots represents the elevation from the equator of the Poincaré sphere and both of the horizontal axes represent the equator.

The plots in figure (4.7) show that a confocal imaging system of $NA=1$ and with a pinhole of one A.U. practically maintains the polarisation state of a field.

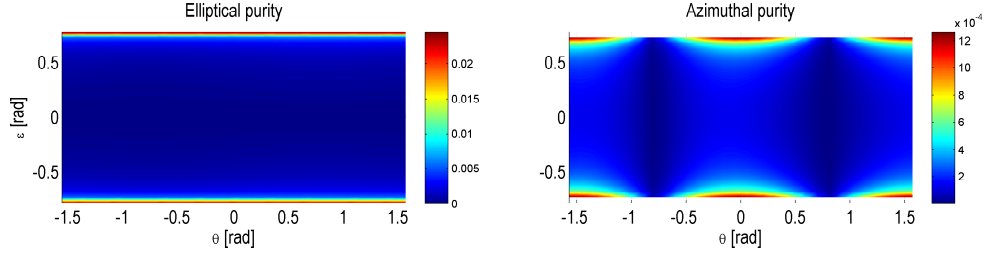


Figure 4.7: Elliptical purity (left) and azimuthal purity (right) for different polarisation illuminations and a pinhole of one Airy unit in diameter.

This conclusion remains true for an optical system using an ideal lens and imaging a dipole. Even though a polarisation distribution exists at the pinhole, see figure (4.3), the irradiance modulation is such that the centre of the distribution has a bigger weight in the integration than the edges where the spatial variations are greater.

The extinction ratio and the degree of purity complement each other. Indeed both metrics quantify the influence of spatial polarisation variations in a polarisation sensitive system, yet on different grounds. The extinction ratio describes the influence of spatial variations on the degree of temporal polarisation and hence it describes the amount of depolarisation present in the field. The degree of purity describes the influence on the remaining two parameters needed to define a Stokes vector, that is to say θ and ϵ .

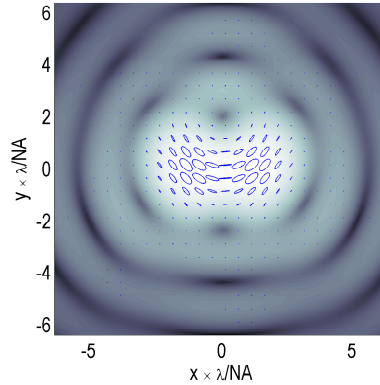


Figure 4.8: Polarisation distribution at the pinhole produced by two dipoles.

For a greater insight into the role of the pinhole consider two dipoles located in the same plane and placed one A.U. apart. It is assumed that one dipole emits a field as if it was excited by a polarisation state of $\theta = \pi/4$ and $\epsilon = -\pi/12$

and the other dipole by $\theta = -\pi/4$ and $\epsilon = \pi/12$. This can be done for example having two molecules fixed at the given azimuth angles, showing the necessary form birefringence⁴ and excited by circularly polarised light. The polarisation distribution produced at the pinhole when imaging these dipoles is shown in figure (4.8). Note that the azimuthal angle varies slowly from $\pi/4$ to $-\pi/4$ along the x axis.

The EDOP and ADOP are calculated next from the image obtained by scanning the dipoles along the x axis at $y = 0$. The EDOP (left) and ADOP (right) are shown in figure (4.9) for different pinholes sizes. In both plots the horizontal axis represents a line scan and the vertical axis represents various pinhole sizes. Note that due to the asymmetry of the polarisation distribution in this example (see figure 4.8), both the EDOP and the ADOP change significantly when increasing the size of the pinhole.

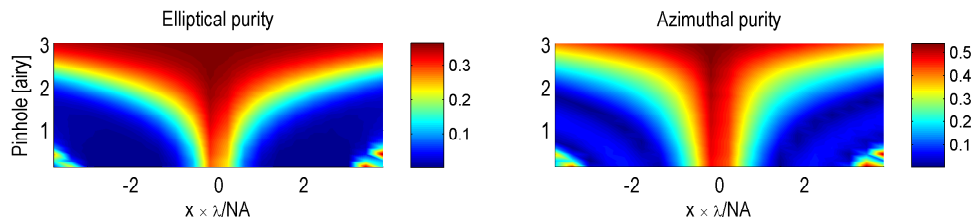


Figure 4.9: Elliptical and azimuthal degree of purity when two dipoles are scanned along the x axis and $y = 0$ with different pinholes sizes

It is possible to say that due to the small changes in both EDOP and ADOP for the single dipole, neither the imaging system nor the pinhole affect significantly a polarisation imaging measurement. The object to be imaged is what underlines the importance of the pinhole in high NA imaging polarimetry.

The next section gives an overview of the imaging parameters used in the rest of this work, namely Stokes and Lu–Chipman parameters. The aim of introducing these parameters is to further characterise the system and for the interpretation of experimental results in the following chapters.

⁴Induced retardance to a field due to the geometrical shape of the object

4.2 Imaging techniques

In this research polarisation imaging is divided in two techniques: Stokes imaging and Müller imaging. The latter images the polarisation properties of an object[67] and Stokes imaging provides information of the state of polarisation of a beam[68]. As it is seen next both techniques yield a number of images with different meaning.

4.2.1 Stokes imaging

In Stokes imaging a single polarisation state is input to the system and a single set of irradiances is being measured (see section 3.2). The set of irradiances is subsequently transformed via the instrument matrix into a Stokes vector. A stoke image can then be formed by plotting the recorded irradiances against the scanning position $\mathbf{r}$. Since a given polarisation state $\vec{S}$ can be fully characterised by the parameters s_1 , $\mathcal{P}^t$, θ and ϵ , the same parameters can be used to interpret the information contained in $\vec{S}$. The depolarised parts of the image are given by $\mathcal{P}^t(\mathbf{r})$, $\theta(\mathbf{r})$ represents the azimuth angle of the polarisation, $\epsilon(\mathbf{r})$ the ellipticity and $s_1(\mathbf{r})$ is the total irradiance of the field. Although the rest of the elements in $\vec{S}$ are not used in this work, for completeness note that s_1 , s_2 and s_3 can be used to represent the amount of light horizontally–vertically, -45° and $+45^\circ$ and left–right circularly polarised, respectively (see section 3.1.1).

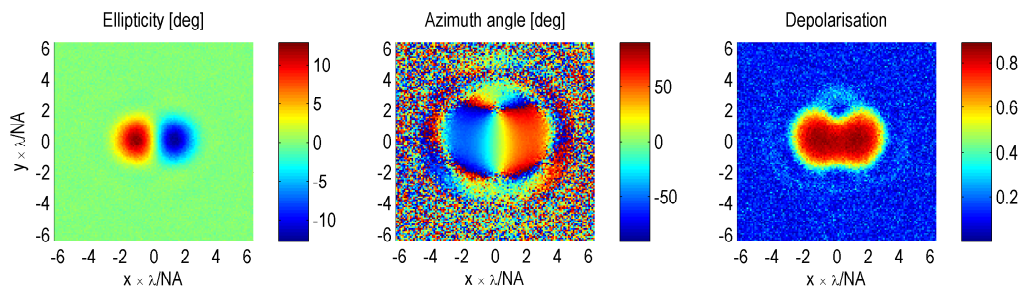


Figure 4.10: Stokes parameters extracted from the Stokes image of two dipoles

To interpret a Stokes image consider the example of two dipoles placed one A.U. apart and with white noise in the detectors⁵. The left dipole is excited by a field having a polarisation state defined by $\theta = -\pi/4$ and $\epsilon = \pi/12$ and the right one

⁵The noise was added to show a non zero plot of $\mathcal{P}^t$

by $\theta = \pi/4$ and $\epsilon = -\pi/12$. The noise was set at 10% of the greatest detected irradiance. The resulting polarisation distribution is the same as in figure (4.8). Figure (4.10) shows the ellipticity (left), azimuth angle (centre) and $\mathcal{P}^t$ (right) from the confocal image of the two dipoles using a point detector. These plots show the influence of detector noise over the measurement and the changes in the ellipticity and the azimuth angle of the polarisation of the beam.

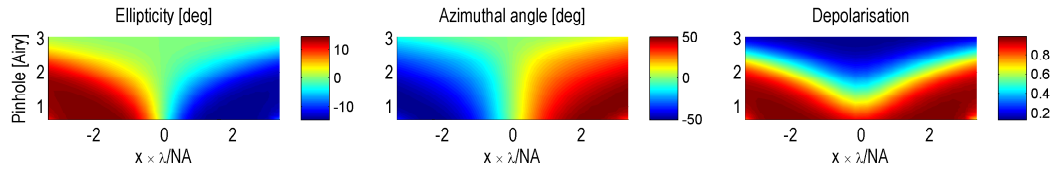


Figure 4.11: Stokes parameters taken from the Stokes image when two dipoles are scanned along the x axis with different pinholes sizes

The influence of the pinhole size on the Stokes parameters is shown in figure (4.11). These plots are calculated from the Stokes images that result from scanning the dipoles along the x axis and changing the size of the pinhole. Note that as the size of the pinhole increases the values of the Stokes parameters change. When the pinhole is 3 A.U. in size, the system cannot resolve the difference between the Stokes parameters at each dipole. Since the plot of depolarisation is calculated from the degree of temporal polarisation, this plot should remain constant as the size of pinhole increases; however, because the pinhole sets the useful area of the detectors (see section 1.4), all polarisation states within this area are added incoherently. The result is a beam that appears to be temporally depolarised and thus may measure as temporal depolarisation (see section 2.2.3 and 4.1.1).

For future work the measured Stokes parameters can be investigated as a function of the distance between the dipoles and the noise in the detectors. This could lead to establishing a relationship between polarisation and the spatial resolution of the dipoles.

4.2.2 Müller imaging

In Müller imaging an object is illuminated by at least four polarisation states. By knowing the illumination states and the instrument matrix it is possible to transform

the detected irradiances into a Müller matrix (see section 3.2). Plotting the scanning position against the measured Müller matrix yields a Müller image.

Müller matrices are sometimes difficult to interpret. One way to gain basic understanding of Müller matrices is by multiplying the matrix by a known Stokes vector. For example, if the matrix is multiplied by a Stokes vector representing unpolarised light, i.e. $(1, 0, 0, 0)^T$, the result is the left column of the Müller matrix. It thus follows that the element m_{11} represents the irradiance distribution at the confocal plane produced by a given object when illuminated with unpolarised light. In the case of imaging a free dipole the element m_{11} represents the unpolarised PSF of the system. Additional information can be obtained by analysing the parameters acquired from the Lu–Chipman decomposition (see section 3.1.2) of the Müller matrix.

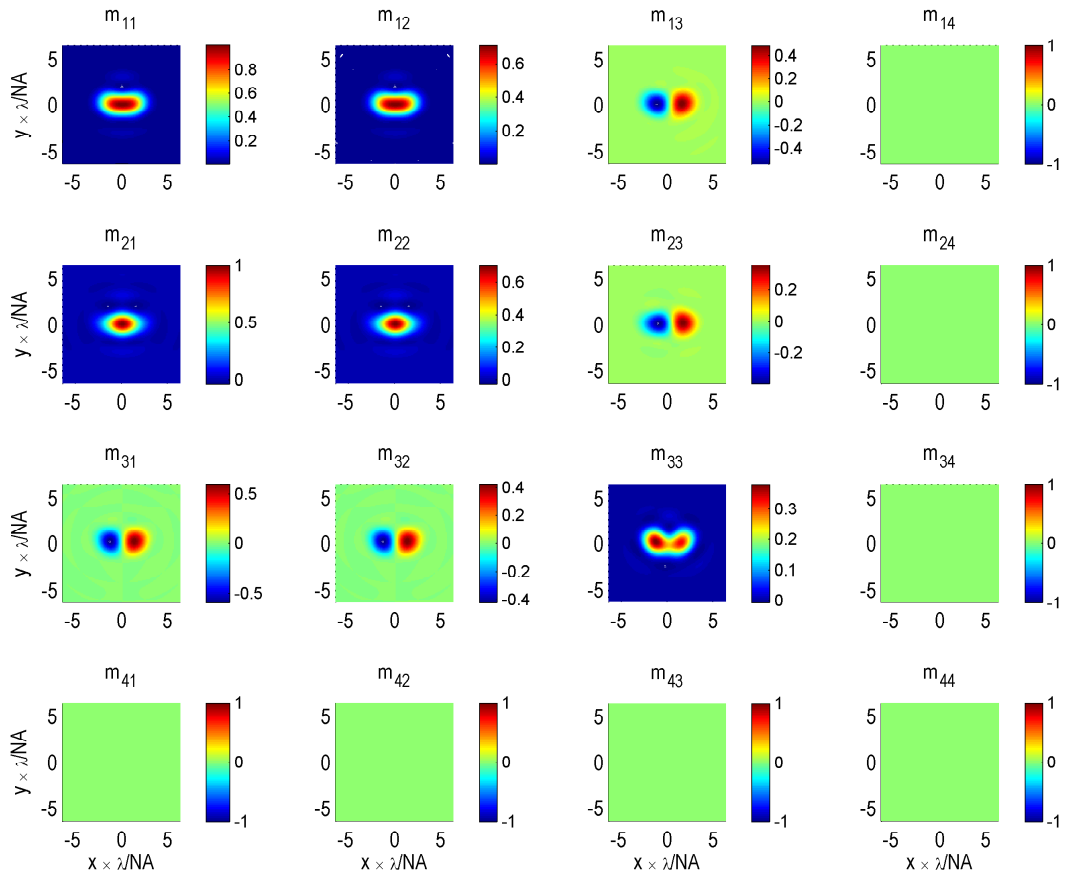


Figure 4.12: Müller matrix of two fixed dipoles placed one Airy unit apart.

The parameters used in this work to interpret a Müller image are the retardance, polarisance, diattenuation and depolarisation resulting from the LC decomposition.

These are the so-called LC parameters. In order to introduce these parameters intuitively consider the two dipole configuration of the previous example, but now without noise⁶. Figure (4.12) shows the Müller matrix obtained when using a point detector to image the dipoles, and the LC parameters are plotted in figure (4.13). To interpret the latter it is important to remember that the LC parameters are obtained from the magnitude of the corresponding vectors normalised by the total irradiance (see section 3.1.2). Because the dipoles are fixed, they fully polarise unpolarised

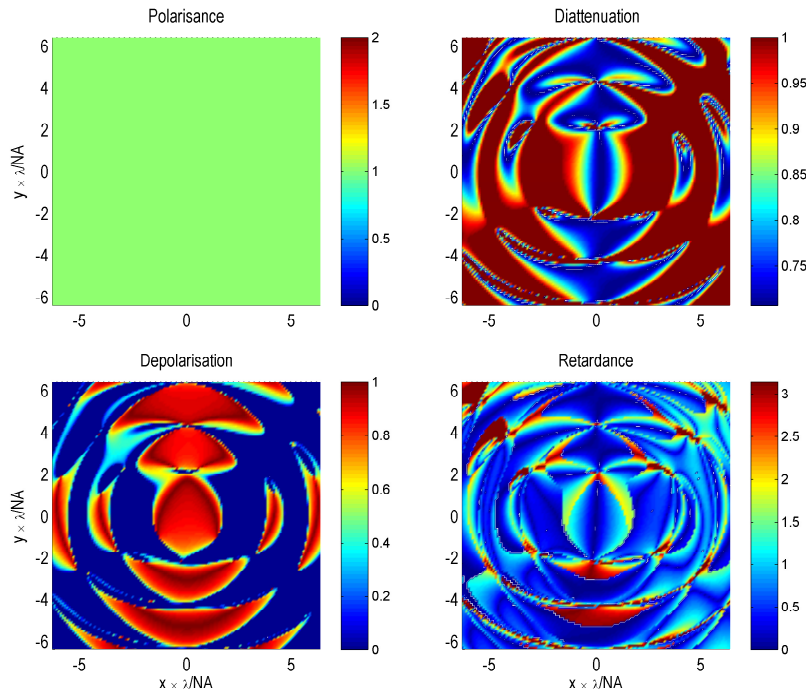


Figure 4.13: Distribution of the Lu–Chipman parameters obtained from the Müller matrix due to two fixed dipoles placed one Airy unit apart.

light. For this reason the polarisance (top left) is unity everywhere. The mixing of the fields emitted by the dipoles results in both an induced retardance (bottom right) and a varying diattenuation (top right). The variation in the depolarisation plot (bottom left) is due to the field distribution of a high NA focused beam. When illuminating with x polarised light, the y and z component are non-zero (see section 1.2). In the portions in figure (4.13) where diattenuation is unity and retardance is non-zero the dipoles behave as an elliptical perfect polariser. Given in addition

⁶Two dipoles are fixed one Airy unit apart in the object plane. The left dipole emits a field as if it was excited by a field having a polarisation state $\theta = -\pi/4$ and $\epsilon = \pi/12$ and the right one by $\theta = \pi/4$ and $\epsilon = -\pi/12$.

that polarisance is unity everywhere, at the points where diattenuation is less than one the dipoles behave as depolariser polariser[69]. This means the dipoles are in this case perfect polarisers but imperfect analysers. Since the emitted field from these portions is fully polarised, it is not possible to specify the polarisation state of the incident field.

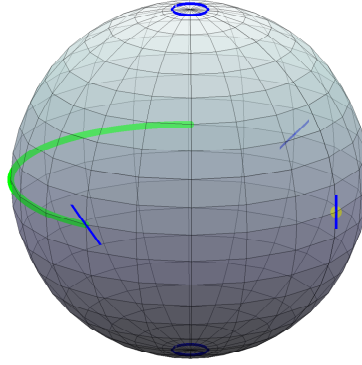


Figure 4.14: Retardance (yellow) and diattenuator (green) vectors calculated from the Müller matrix of two fixed dipoles.

In order to visualise which incident polarisation states are being diattenuated or retarded, the diattenuation and retardance vectors are plotted in the Poincaré sphere in figure (4.14). This figure shows in green the diattenuator axis from each pixel in the image and the fast axis of the retardance is shown in yellow. The blue features are for reference only.

In order to gain an understanding of confocal polarisation imaging systems, it is instructive to study a dipole placed in the focal region of the high NA lens. The dipole is assumed to be induced by the incident electric field and has to rotate before it re-emits. The Müller matrix due to the free dipole found by using a point detector and four different polarisation illuminations is shown in figure (4.15). Using this matrix it is possible to study how the imaging system responds to a given polarisation state[67]. Figure (4.16) shows the LC parameters of the free dipole. The preferential transmission or reflection of the system to a given incident polarisation state is shown by the diattenuation plot (top right). Since the Müller matrix of a free dipole has orthogonal eigenvectors, this plot can be calculated directly from

$$|\vec{D}|_{\text{dipole}} = \frac{|\vec{E}(\theta, \epsilon)|^2 - |\vec{E}(\theta \pm \pi/4, \epsilon)|^2}{|\vec{E}(\theta, \epsilon)|^2 + |\vec{E}(\theta \pm \pi/4, \epsilon)|^2}. \quad (4.2)$$

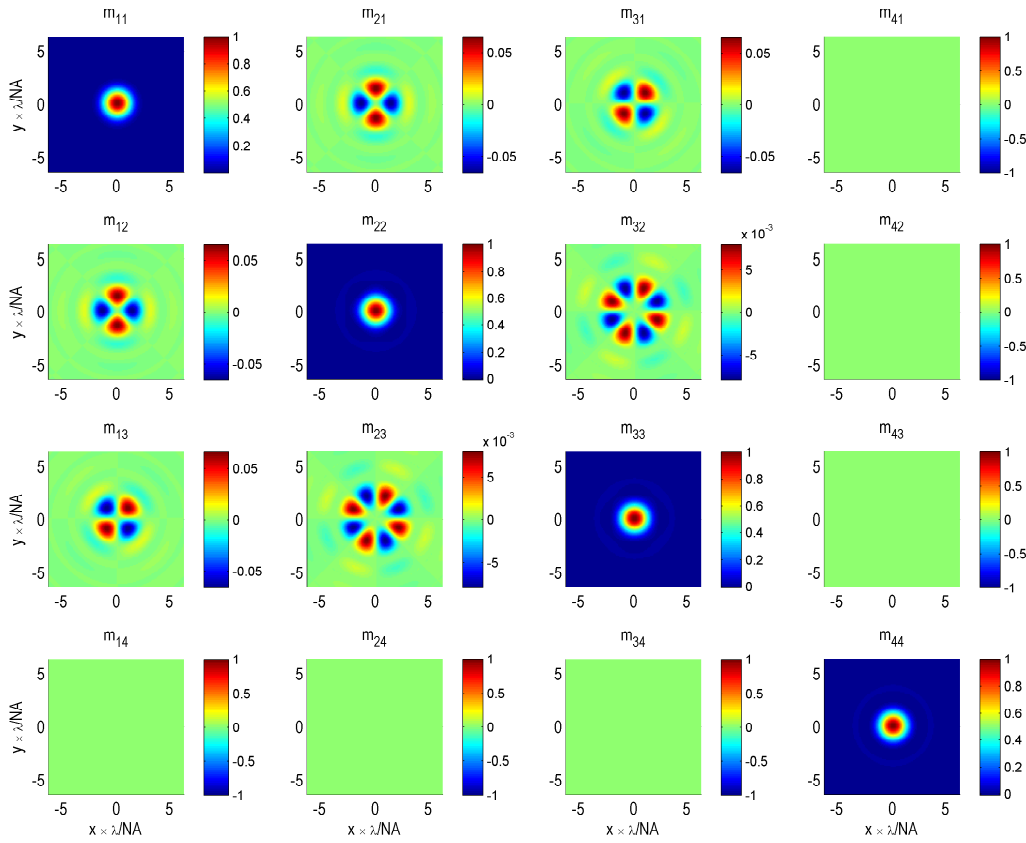


Figure 4.15: Müller matrix due to a free dipole.

The plot of the polarisance (top left) was also found in [70] using the polarisation matrix (see section 2.1). Given that the polarisance and the diattenuation show the same distribution, the optical system is in general a linear analyser–polariser. This is the reason why a preferential influence exists in the extinction ratio to linearly polarised light as shown in section 4.1.1. Note that the centre of both the diattenuation and polarisance is zero, and hence at this point the imaging system has no effect on the polarisation. The depolarisation distribution (bottom left) reveals that the system does not introduce temporal polarisation effects. The retardance plot (bottom right) shows the π phase shift corresponding to the minima of the Airy pattern.

The use of Stokes or Müller imaging depends on the application and the experimental resources. The Lu–Chipman parameters represent the polarisation response of an object to a plane wave illumination. In order to consider different types of illumination such as spherical waves, it is necessary to rely on Stokes imaging en-

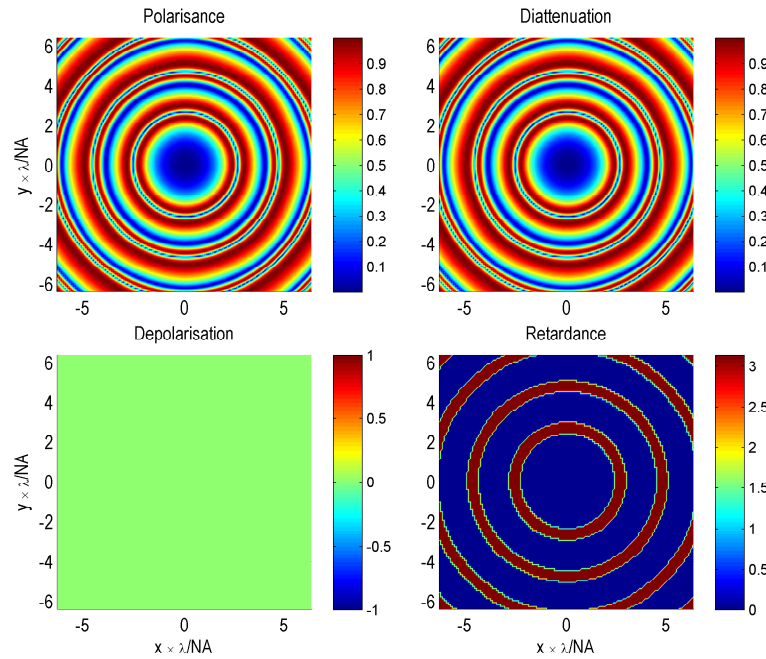


Figure 4.16: Lu–Chipman parameters calculated from the Müller matrix due to a free dipole.

tirely. However the two techniques combined can yield better imaging results such as higher spatial resolution, contrast and even filtering. The next section explains a method to choose the polarisation state of the illumination that can improve the imaging.

4.3 Illumination engineering

Due to the polarisation dependance of both physical and geometrical properties, a given object responds differently to different polarisation illuminations. Take for example a dipole fixed horizontally in respect to the reference axis. This dipole can only couple and re-emit the horizontal component of the illumination. The aim of this section is to introduce a method to exploit this polarisation dependance by engineering the polarisation state of the illumination. In principle the method can be expanded to consider a structured illumination. In this work however a single polarisation state is considered for simplicity. The method consists of first finding the polarisation properties of a given object and using these properties to then define a metric in the context of set theory.

To find the polarisation properties of an object the Müller matrix of the object itself is measured and decomposed using the Lu–Chipman method. The decomposition yields the retarder and diattenuator Müller matrices of each pixel in the image, from which the respective eigenpolarisations can be extracted. Let $\vec{V}_j^p$ and $\vec{U}_j^p$ be the eigenpolarisations that correspond to the property $p = \text{retarder}, \text{diattenuator}$ of the pixel j . These eigenpolarisations can then be used to define a family of sets as

$$\mathbb{S}_j^p = \{\vec{S}(\theta, \epsilon) : \vec{S}^T(\theta, \epsilon) \cdot \vec{V}_j^p = q_j(\theta, \epsilon) \text{ and } \vec{S}^T(\theta, \epsilon) \cdot \vec{U}_j^p = q_j(\theta, \epsilon)\}, \quad (4.3)$$

where $\vec{S}(\theta, \epsilon)$ represents the polarisation state of the illumination and $q_j(\theta, \epsilon)$ is a metric that can be chosen depending on the relation between eigenpolarisations and eigentransmittances⁷. For example setting the value of an eigentransmittance to zero, means that the corresponding eigenpolarisation is not transmitted at all. The state $\vec{S}(\theta, \epsilon)$ can thus be found by specifying $q_j(\theta, \epsilon)$. A particular case of interest to the optical data storage application in this work is when $q(\theta, \epsilon) = 0$. In this case the dot product between the polarisation state of the illumination $\vec{S}$ and the eigenpolarisations of the object is zero. This means that the family $\mathbb{S}_j^p$ contains all the $\vec{S}$ that suffer the greatest change after interacting with the j th portion of the object.

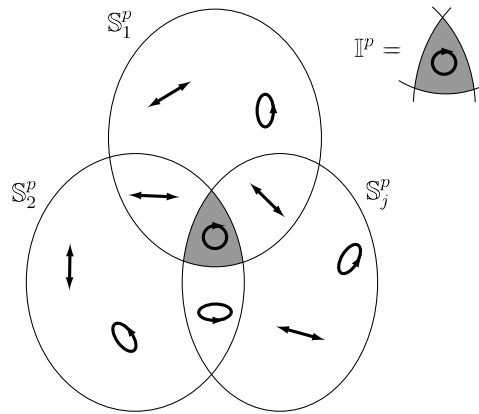


Figure 4.17: Venn diagrams representing equation (4.4).

⁷Transmittance value of an eigenpolarisation. This value is given by the eigenvalue corresponding to the eigenvector that describes such eigenpolarisation

Having found $\mathbb{S}_j^p$, the aim now is to find the polarisation states that all the $\mathbb{S}_j^p$ have in common to construct⁸

$$\mathbb{I}^p = \mathbb{S}_1^p \cap \mathbb{S}_2^p \cap \dots \cap \mathbb{S}_j^p \cap \dots \cap \mathbb{S}_n^p \quad (4.4)$$

for n pixels and each polarisation property p . The set $\mathbb{I}^p$ contains the polarisation states that have the greatest common influence throughout the whole image of the given object. The previously mentioned idea is depicted in figure (4.17) using Venn diagrams. In the case of $q(\theta, \epsilon) = 0$, it follows that the Müller matrices of the j pixels map all the $\vec{S}$ in $\mathbb{I}^p$ an equal distance within the Poincaré sphere. This means that the resultant $\vec{S}_j$ from $\vec{S}_j = \mathbf{M}_j \vec{S}$ are transformed an equal amount by all the pixels. To illustrate the previous statement take for example a linear polariser, i.e. a Glan–Thompson polariser. Since the eigenpolarisations of this polariser are two orthogonal Stokes vectors representing linearly polarised light, $\mathbb{I}^p$ then contains both right and left circular polarisation states. If any of the $\vec{S}$ in $\mathbb{I}^p$ is multiplied by the Müller matrix of the polariser, the resultant Stokes vector always lies on the equator of the Poincaré sphere. Please note that when choosing $q(\theta, \epsilon) = 0$ the unpolarised state is always contained in $\mathbb{I}^p$ despite of the object. Also note that the user can choose as well the $\vec{S}$ contained in the intersection between $\mathbb{I}^{retarder}$ and $\mathbb{I}^{diattenuator}$.

The ideas described in this section become more obvious in chapter 8, when engineering the illumination for the multiplexed optical data storage disk prototype.

4.4 Conclusions

This chapter has reviewed basic characteristics of confocal micropolarimetry such as resolution, extinction ratio, degree of purity and the theoretical polarisation PSF of the system. The chapter also explained Stokes and Müller imaging, illumination engineering, and the role of the confocal detector.

Due to small changes in both EDOP and ADOP for a single dipole, it is possible to conclude that neither the imaging system nor the pinhole affect significantly the acquisition of the polarisation images. The importance of the pinhole in high NA imaging polarimetry arises due to the object. If the imaging system is polarisation

⁸The symbol $\cap$ denotes intersection.

corrected, the role of the pinhole is to set the spatial resolution of the system. For systems not corrected and having a lens with an NA greater than 0.17, confocal imaging is the only reliable method to image the polarisation; providing the pinhole size is chosen appropriately. From the results given throughout this chapter, it appears that a pinhole of 1 Airy units in diameter yields acceptable results.

For finite sized pinholes, it was seen that the measured degree of polarisation is non-zero due to both temporal and spatial effects. Since the field is integrated spatially by the photodetectors in the PSA, any existent spatial distribution of polarisation in the beam affects the measurement of the degree of polarisation. A fully temporally polarised beam can thus appear as partially polarised.

From the calculated PSF matrix of the system (Müller matrix due to a free dipole), it is possible to conclude that a confocal micropolarimeter behaves as a linear analyser-polariser. The imaging system does not depolarise temporally but it does spatially. The only retardance introduced by the system corresponds to the π phase shift of the Airy pattern. The experimental proof of the latter is considered as future work.

Before proceeding to the results it is necessary to characterise and calibrate certain parts of the optical system. The next chapter deals with the characterisation of the objective lens and the scanning stage, and introduces the polarisation state analyser and the polarisation state generator. The calibration of the polarimetric part of the system is given in the subsequent chapter.

Chapter 5

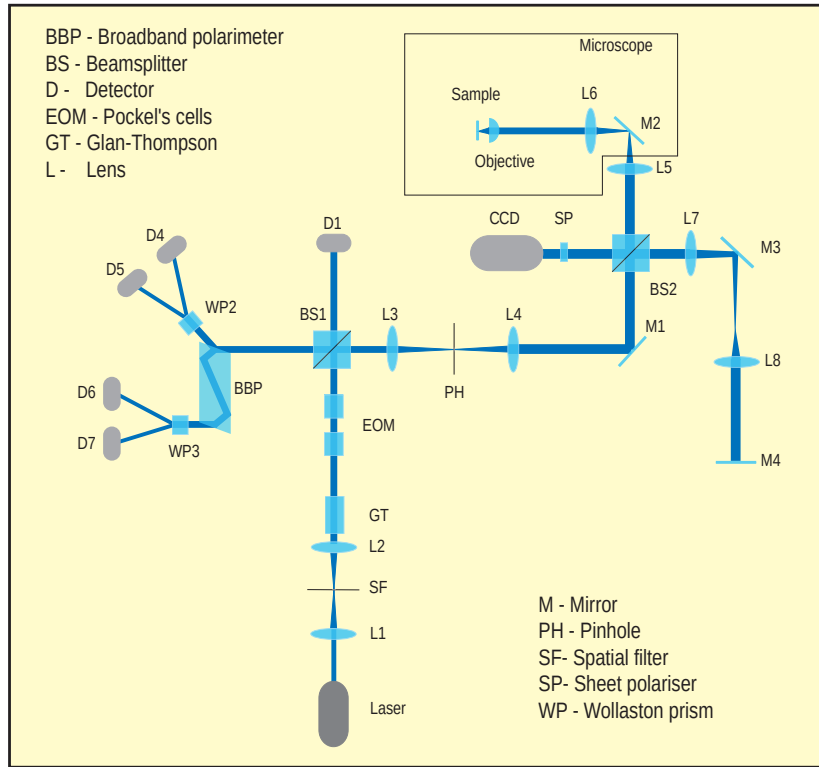
El experimento

This chapter presents the layout of the experimental system. Individual parts of the system are also addressed here in three main sections; namely scanning, imaging optics, and polarisation optics.

Regarding the imaging optics, the objective lens is described and characterised by means of interferometry. In addition, a measurement of the axial and transversal PSF of the system is shown and analysed. The introduction to the polarimetric part consists of the description of the polarisation state generator (PSG) and the polarisation state analyser (PSA). The device used for scanning is described here together with its calibration.

5.1 Experimental layout

Figure (5.1) shows the schematic of the confocal polarisation microscope built for this research. The light source is a highly distorted laser beam produced by a $405\pm 3\text{nm}$ Nichia laser diode (Coherent: Cube Laser). The beam is shaped using a single mode optical fibre (Point Source: KineFLEX). After the fibre, the beam is 0.7mm in diameter and aberration corrected. The beam passes through a Glan-Thompson polariser and then through a pair of electro-optical modulators (EOMs) manufactured by Leysop. The polariser sets the reference frame of the system, and the EOMs change the state of polarisation of the beam. After the EOMs, the light is split 50/50 by a custom made non-polarising beamsplitter (NPBS) from Lambda Photometrics. Half of the light is incident on a detector for monitoring purposes

**Figure 5.1:** Experimental setup

and the other half is focused ($f = 30\text{mm}$) to a pinhole. The light transmitted by the pinhole is collimated ($f = 150\text{mm}$) and passes through relay optics to then reach an objective lens of NA 0.95 (Zeiss: Epiplan-Apochromat). The purpose of the relay optics is to image the back focal plane of the objective lens into both the CCD camera and the diaphragm shown in the figure. The diaphragm is used to reduce the NA of the objective lens, and the camera forms part of a Twyman-Green interferometer that characterises the aberrations present in the objective lens. The reference path of the interferometer is created by a 70/30 NPBS¹. The light reflected from the sample is collected by the high NA objective lens, returns through the relay optics to then reach the pinhole. A portion of the light is filtered by the pinhole and then collimated to finally enter the polarimeter.

A more detailed introduction of individual elements of the setup is given in the sections to follow.

¹Transmission/reflection percentage.

5.2 Scanning

In confocal microscopy the image of an object can be formed by either scanning the beam across the object using a mirror or by scanning the object across a fixed beam using a stage. Nonetheless, being optical data storage the original motivation of this research, it was desired that the system resembles as much as possible a disk read-out system². The chosen scanning unit was a 3D nanopositioning stage (PIMars: P-563 stage).

The chosen stage has 3 channels x, y, z ; the z axis being the optical axis. Each axis has a $300\mu\text{m}$ range of movement. The stage is driven by varying the voltage in each channel between $0V$ and $10V$, and the actual position is monitored using capacitive sensors integrated in the stage. The position is both controlled and monitored by a National Instrument card. According to the manufacturer of the card the noise in the input and the readout signal is equivalent to $\pm 4\text{nm}$. The resonance frequencies of the x, y, z axes loaded with 200mg are 106Hz , 109Hz and 140Hz , respectively. These frequencies correspond to the frequency of the ripples created in the scanner when driven by a step pulse of $1V$ amplitude.

During scanning tests it was found that the stage depends on the load, and on the frequency and the amplitude of the driving waveform. The stage was also found to low-pass filter and delay the throughput signal. As explained next however, these features were accounted for with the characterisation of the stage.

5.2.1 Pre-shaping

The response of the stage was characterised by a method called pre-shaping. In this method the driving waveform is shaped in such a way that the readout signal has the same shape as a desired waveform. If for example the x axis is driven with a triangle waveform and the stage moves sinusoidally, the driving signal is shaped so the stage movements follow the triangle waveform.

The pre-shaping is carried out using the “instant” transfer function of the stage as follows. Let $A(v, t)$ to describe a given waveform and $B(v, t)$ the response of the stage to $A(v, t)$, where v is the driving voltage of the stage. The transfer function

²For example a CD player

that characterises this particular response is then given by

$$\mathcal{H}_B(v, \omega) = \frac{\mathcal{B}(v, \omega)}{\mathcal{A}(v, \omega)}, \quad (5.1)$$

where $\mathcal{A}(v, \omega)$ and $\mathcal{B}(v, \omega)$ are the time Fourier transforms of $A(v, t)$ and $B(v, t)$, respectively. Assuming now that $C(v, t)$ is the waveform that forces the stage to follow $A(v, t)$, it is possible to write

$$\mathcal{A}(v, \omega) = \mathcal{H}_B(v, \omega)\mathcal{C}(v, \omega) \quad (5.2)$$

being $\mathcal{C}(v, \omega)$ the time Fourier transform of $C(v, t)$. It then follows that the pre-shaped signal is given by

$$C(v, t) = \mathfrak{F}^{-1} \left\{ \frac{\mathcal{A}^2(v, \omega)}{\mathcal{B}(v, \omega)} \right\}, \quad (5.3)$$

where $\mathfrak{F}^{-1}$ denotes inverse Fourier transform.

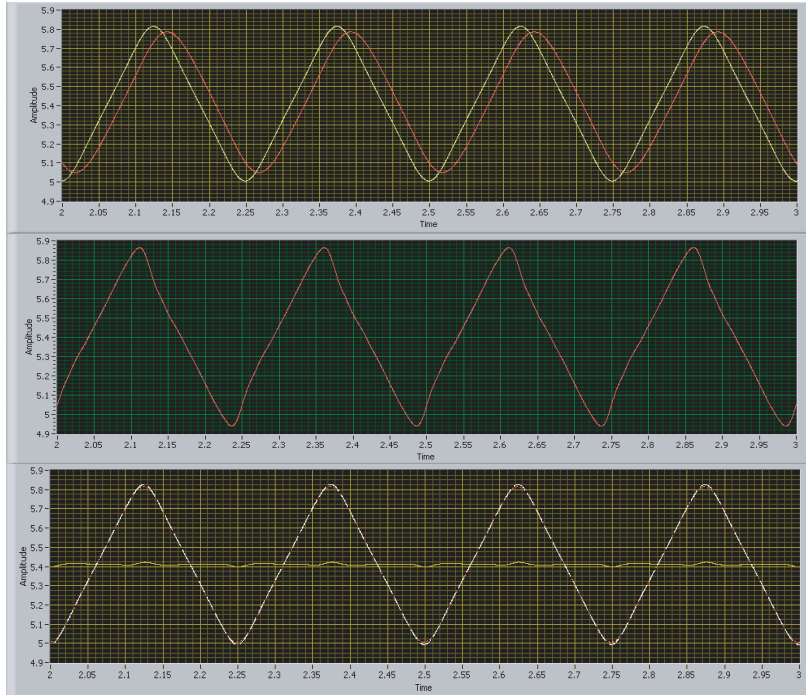


Figure 5.2: Top: Desired waveform and response of the stage. Centre: Shaped waveform. Bottom: comparison between the shaped waveform and the desired waveform

Figure (5.2) shows an example of the pre-shaping method. The top plot is a

comparison between the desired waveform (yellow) and the response of the stage to such a waveform (red), the centre plot is the waveform computed from equation (5.3), and the bottom plot is a comparison between the new output waveform (red) and the desired waveform (white).

After characterisation the stage is able to scan any 3D functional path with a standard deviation of 5nm. However this accuracy is achieved when scanning with a triangle waveform low-pass filtered at 10Hz and driven at 2Hz. As the frequency content, the amplitude, and the driving frequency of the waveform increases, the accuracy at the crests and troughs decreases due to a non linear response of the stage. The characterisation of non linearities in the stage is left as future work.

The pre-shaping method explained above was implemented in a LabVIEW program written to control the experiment. The next section explains more about this program and how an image is being acquired.

5.3 LabVIEW software

Most of the components in the setup are controlled by a program written in LabVIEW. The program reads out the measured irradiances by the PSA, sets the voltage input to the Pockel's cells, and controls the stage (including the pre-shaping method). Within the control of the stage the user can choose between a 1D, 2D or 3D scan.

During an image acquisition the software drives simultaneously all three axes with the given pre-shaped waveforms. In a 2D scan for example, the x axis is scanned with a ramp while the y axis is scanned with a triangular waveform and the z axis is kept constant. As a result of this simultaneous driving the image is stored in a vector, where each element of the vector has an assigned x, y, z position. The confocal image is then formed by plotting the image vector against the x, y, z waveforms. In the case of Stokes imaging four vector images are stored, i.e. one per detector. For Müller matrix imaging sixteen images are stored, i.e. one per input polarisation and per detector.

5.4 Imaging optics

The objective lens used in this research is an EC Epiplan-Apochromat 100x/0.95 HD DIC M27 objective manufactured by Zeiss. The lens was corrected by the manufacturer for 405nm wavelength illumination and optimised for polarisation imaging.

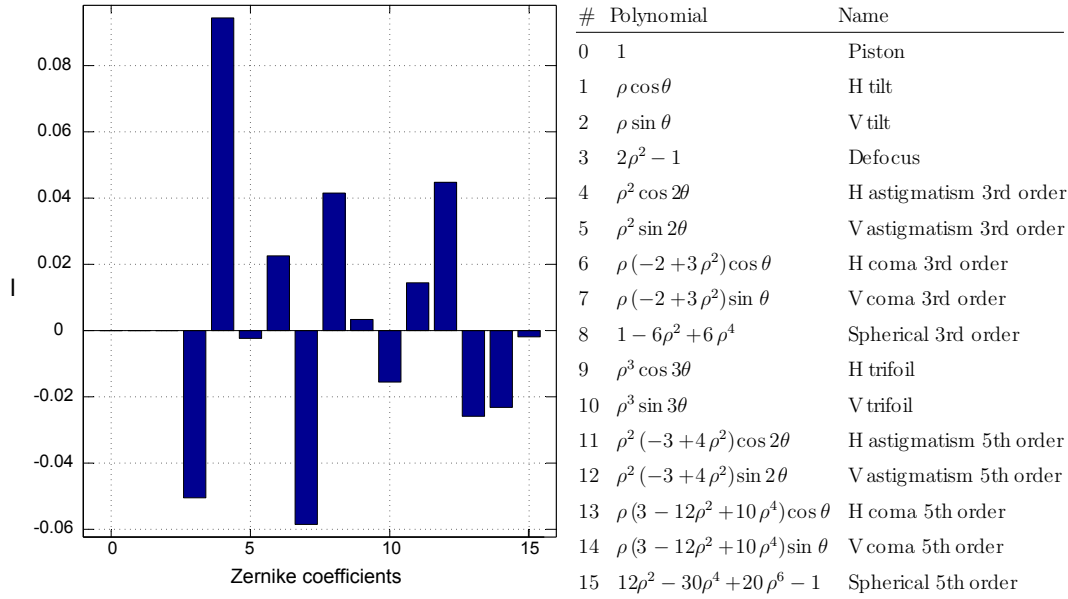


Figure 5.3: Zernike coefficients of the aberrations present in the objective lens (left), and the corresponding Zernike polynomials (right). H stands for horizontal and V for vertical.

The aberrations in the objective lens were measured using a Twyman-Green interferometer and a flat mirror. Ideally a spherical mirror should be used when measuring the aberrations of a lens[16]; however, it was not possible to find a parabolic mirror of $\text{NA} \geq 0.95$. As a consequence only the even aberrations should be considered as correct. The measurement consisted of maximising the amount of light passing through the pinhole by adjusting the axial position of the lens in respect to the flat mirror. This in consequence yielded the right amount of defocus that would correct for spherical aberration if present. The interference pattern produced by the light reflected from the mirror and the light from the reference path was then recorded using a CCD camera. The pattern was processed by a wavefront analysis software developed by Rami Saab and Arthur van de Nes. In this software the pattern is Fourier-transformed, the first diffraction order is filtered and then

inverse—Fourier—transformed. The result of this process is a wrapped wavefront that is unwrapped and fitted to Zernike polynomials. Figure (5.3) shows the measured coefficients of the first 15 polynomials and their names. The coefficients were normalised to the wavelength. It is possible to see from this figure that the greatest aberration in the objective lens is 0.095 wavelengths of horizontal astigmatism.

5.4.1 Point spread function of the system

The measurement of the PSF of the system consisted on measuring the Müller matrix of a 20nm gold bead. A microscope slide was impregnated with a solution of HCl and gold particles diluted in water to 0.1 mol. The slide was in contact with the resulting solution and then rinsed with water to remove any loose beads.

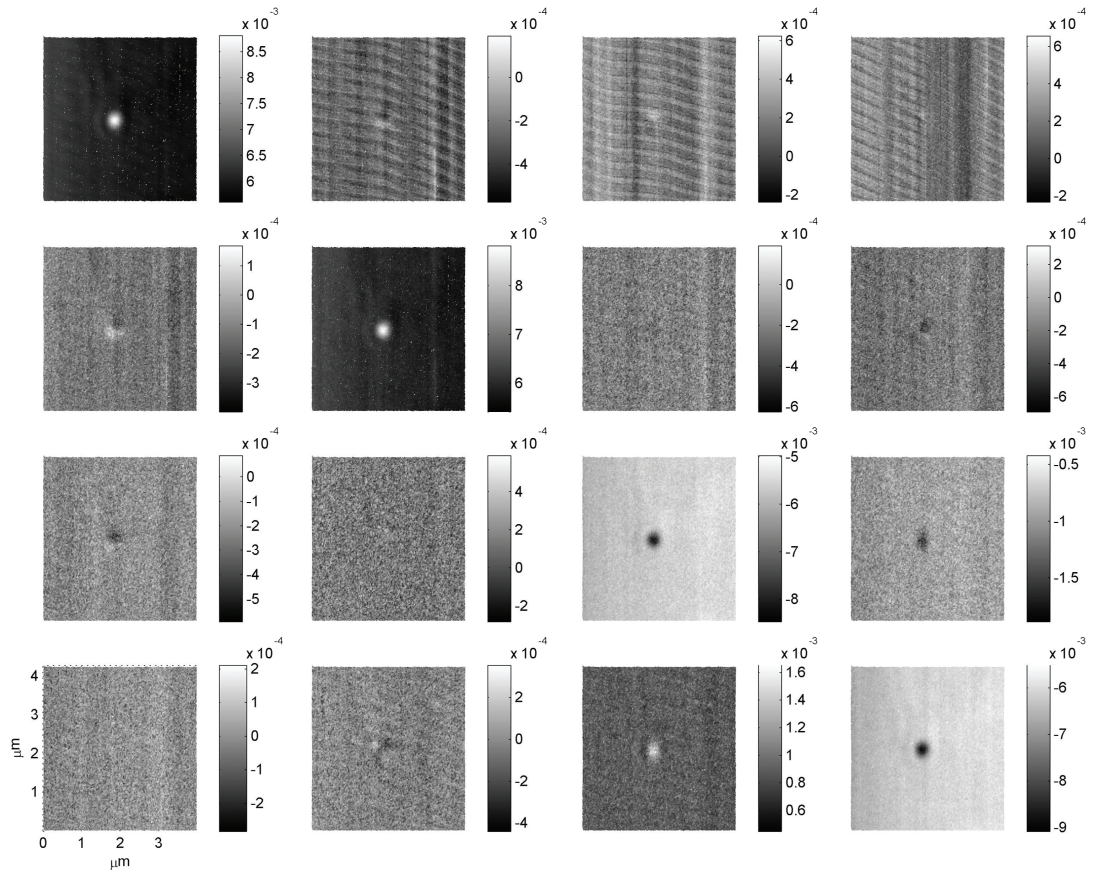


Figure 5.4: Confocal signal from a low reflecting surface

Figure (5.4) shows the Müller matrix of the transverse PSF, and figure (5.5) is the Müller matrix of the axial PSF of the system. The axial and the transversal

resolution according to the Rayleigh criteria³ are 437nm and 352nm, respectively. A PSF with similar characteristics can be created with an objective lens of 0.71 NA.

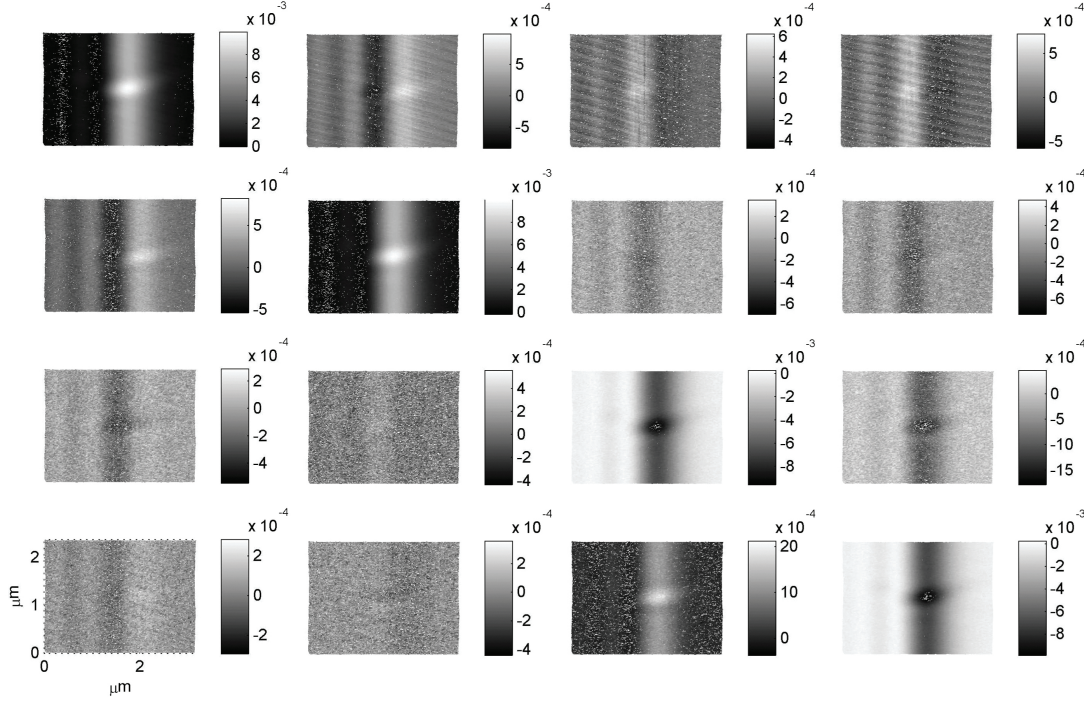


Figure 5.5: Confocal signal from a low reflecting surface

Due to the noise present in figure (5.4), it is not possible to draw reliable conclusions regarding the polarisation characteristics of the gold bead. A measurement of a larger bead which in principle should reflect more light is left as future work. Figure (5.5) shows a different axial PSF than the conventional measurement where only a central and side lobes can be seen. This is due to the interactions of the light between the bead and the coverslip where the bead is deposited.

Note that the measurement of the 20nm bead pushes the capabilities of the experimental system to the limit. A pinhole of 1/3 of the Airy disk was used for this measurement, and the greatest amount of light entering the PSA was approximately 645 nanowatts. Since the responsivity of the photodiodes used is approximately 0.15 amps per watt, this means that each detector was generating less than 100 nanoamps in the best case scenario. The light levels were measured using a calibrated Newport powermeter from the laboratory of Dr. Carl Paterson.

³The beam diameter at 80% of the total irradiance

5.5 Polarisation optics

5.5.1 Broadband polarimeter

The PSA used in this work is the broadband polarimeter (BBP) introduced by Compain and Drevillon[71]. Figure (5.6) shows the schematic of the BBP designed for 405nm wavelength. The beam is incident to the top face of the prism at 78.4° and split into two paths. The reflected portion is diattenuated by the surface of the glass and passes through a Wollaston prism with fast axis at 45° . The beam is further split by the Wollaston prism into two and the irradiance of each beam is measured using a photodiode. The transmitted portion is also diattenuated by both the entrance and exit surfaces and retarded $\pi/2$ by two total internal reflections. The transmitted beam is further split by a Wollaston prism with fast axis at -45° , and the resulting beams are finally measured individually using a photodiode. Note that since each one of the four paths have different polarisation properties, they all act as different analysers.

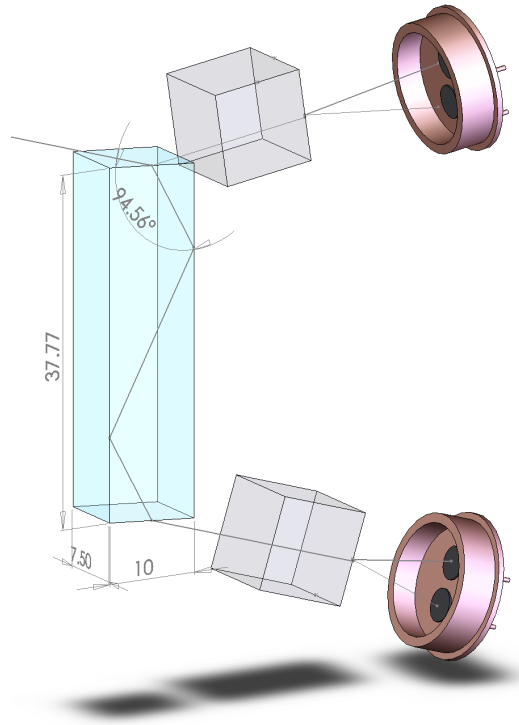


Figure 5.6: Schematic of the broadband polarimeter

When a beam with a given polarisation state passes through the polarimeter, the four detectors yield a unique set of readings (for a noise free system). Given now that changing the state of polarisation changes linearly the detected irradiances, it is possible to find a linear transformation between Stokes vectors and measured irradiances. This transformation is represented in this case by the 4×4 instrument matrix⁴.

The BBP was designed and optimised using the condition number⁵ κ of the instrument matrix. The Müller matrix of each analyser path is first found by multiplying the Müller matrix of an ellipsometric surface and the matrix of a polariser. The Müller matrices of the reflected paths have the Fresnel reflection coefficients as parameters. The polariser is set at 45° for one of the paths and at -45° in the other path. Given now that only irradiances are being detected, all the first rows of each matrix are stacked to construct a parameterised instrument matrix $\mathbf{T}$. By then using an algorithm coded in Matlab, the condition number of $\mathbf{T}$ is minimized by changing the given parameters. Once the optimum condition number is obtained, the code yields a set of Fresnel coefficients that can be used to find the optimum angle of incidence and the angle of the total internal reflection. The remaining geometrical parameters of the prism follows from these two angles and the size of the beam. The resulting instrument matrix from this optimisation is given by

$$\mathbf{T}_{theory} = \begin{pmatrix} 0.9681 & 0.5524 & 0.7951 & 0 \\ 0.9681 & 0.5524 & -0.7951 & 0 \\ 1 & -0.6111 & 0.0009 & 0.7915 \\ 1 & -0.6111 & -0.0009 & -0.7915 \end{pmatrix}$$

with $\kappa = 1.761$; being $\sqrt{3}$ the optimum possible[57].

The BBP polarimeter is fixed on the “auto-aligning” mount shown in figure (5.7). This mount was designed with the aim of facilitating the experimental optimisation of the angle of incidence. When moving the micrometer the prism pivots around the point where the light is incident on the prism, and at the same time the Wollaston prisms move following the trajectory of the light. The BBP prism was

⁴See section 3.2 for a more general description of instrument matrices

⁵See section 3.2.1 for condition number

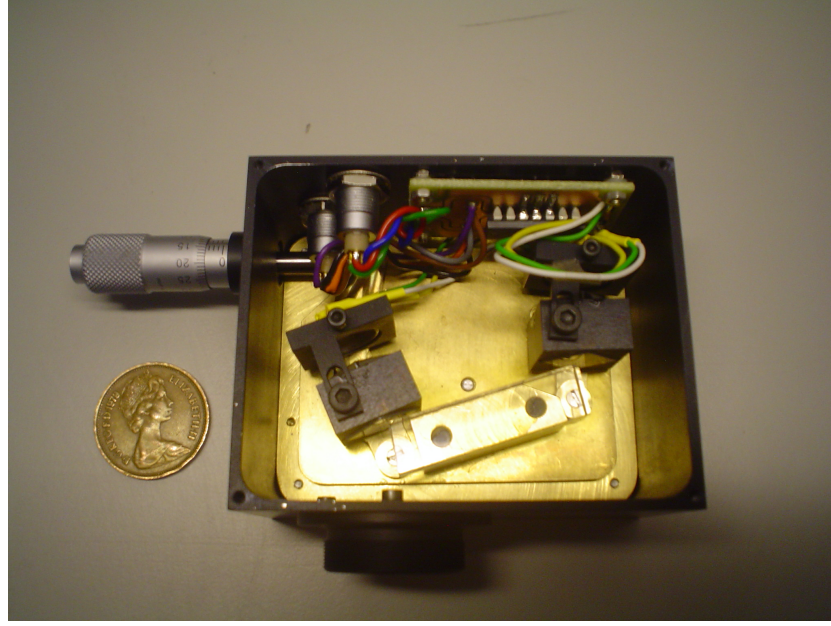


Figure 5.7: The polarimeter

polished by James Stone from N-LAF34 glass, the electronics were built by Brian Burrows and the mechanics constructed by Simon Johnson.

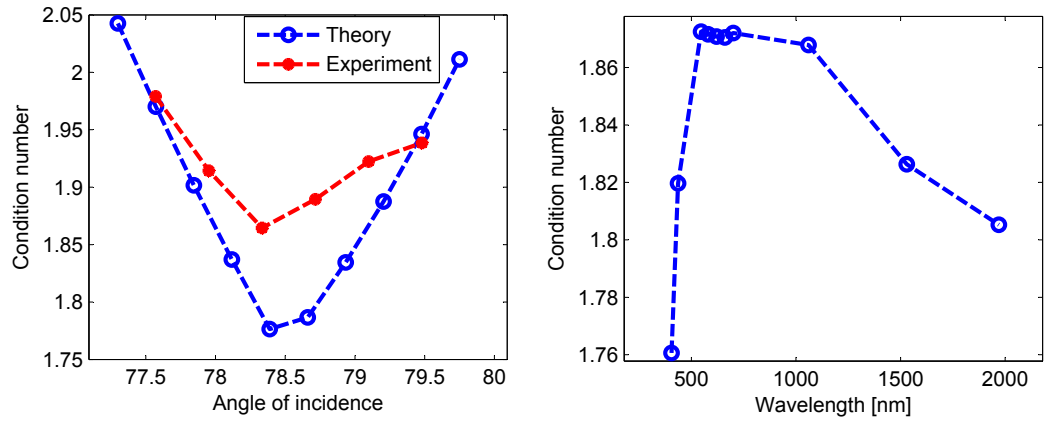


Figure 5.8: Angle (left) and wavelength (right) dependence of the condition number

The wavelength dependence of κ is shown in the right hand side plot of figure (5.8), and the angular dependence is given in the left hand side plot. As it can be seen from the latter the optimum condition number is in practice higher than the

expected one. This is because the instrument matrix in reality is

$$\mathbf{T}_{\text{experiment}} = \begin{pmatrix} 0.9524 & 0.5979 & 0.7625 & 0.0301 \\ 0.8630 & 0.4577 & -0.7277 & -0.0292 \\ 0.9113 & -0.5172 & 0.0093 & 0.7341 \\ 1 & -0.6260 & -0.0238 & -0.7688 \end{pmatrix}$$

with $\kappa = 1.86$. The difference between theory and practice is due to the entrance and exit faces not being parallel to each other. This feature helped to deviate the parasitic reflection from the exit surface.

5.5.2 Pockel's cells

Pockel's cells proved to be the most appropriate PSG when using a 405nm laser illumination⁶. Formed by a lithium niobate crystal placed between a pair of electrodes, these cells act as a variable retarder since the birefringence of the crystal varies proportionally to the voltage applied. If two of these cells are placed in an appropriate configuration, it is possible to produce a set of four optimum polarisation states needed for polarimetry (see section 3.2). Similarly to [5], the cells were placed with their fast axis at 27.4° and 72.3° from the reference axis⁷, and the polarisation states generated were the resulting states from the four possible retardance permutations of $7\pi/4$ and $3\pi/4$ radians. The voltages corresponding to these retardances were found by pre-calibrating the cells.

The Pockel's cells are pre-calibrated using Malus' law and a pair of Glan-Thompson polarisers. Each one of the cells is first placed individually with its fast axis at 45° between crossed polarisers, and the applied voltage to the cell is then varied linearly across a full range, i.e. $-2V$ to $+2V$. By varying the applied voltage, the amount of light that passes through the second polariser changes according to

$$I(v) = I_o \cos^2 \delta(v), \quad (5.4)$$

where I_o is the amplitude, and $\delta(v)$ is the retardance induced by the Pockel's cells in

⁶A pair of liquid crystals where burnt by the laser

⁷Set by the angular position of the first polariser

function of the applied voltage v . Once the transmitted irradiance $I(v)$ is measured as a function of the voltage, the required retardance and hence the voltages can be found by solving Eq. (5.4) for $\delta(v)$. In practice however the Pockel's cells respond non-linearly to the applied voltage. The set of voltages found during pre-calibration were used only as a reference and then optimised after calibrating for polarisation. The condition number obtained was 1.739.

The matrices that actually represent the PSA and the PSG were found experimentally by calibrating the system for polarisation. As it is seen in the next chapter, the calibration of polarisation features in the system played an important role during this research.

5.6 Conclusion

Individual parts and the layout of the system were introduced in this chapter. An interferometric characterisation of the objective lens was given together with the measurement of the axial and transversal PSF of the system. The scanning device was introduced and a calibration procedure was shown. Regarding the polarimetric part of the system the PSA and the PSG were explained in detail.

The robustness of the experimental setup was also quantified in this chapter. It was shown that it is possible to perform a stage scan with a precision of $\pm 5\text{nm}$ by pre-shaping the driving waveform. It was also mentioned that the measured axial and transversal resolutions are 437nm and 352nm , respectively. Regarding the polarimetric part of the system it was possible to reach a condition number of 1.739 for the PSG and of 1.86 for the PSA, where $\sqrt{3}$ is the best possible. Finally from the measurement of the 20nm gold bead, it was proved that it is still possible to comment on the resolution of the system despite of the ~ 645 nanowatts entering the PSA.

The next chapter deals with the calibration of the polarimetric part of the system. Since the system is formed by optical elements that can change the polarisation of the light in an unknown way, it is necessary to account for this.

Chapter 6

Calibration

The objective of this chapter is to explain the calibration of a confocal polarisation microscope, in particular the polarisation sensitive part of the apparatus. Polarisation sensitive systems such as Müller matrix polarimeters are commonly calibrated using the Eigenvalue Calibration Method (ECM)[\[72\]](#). Due however to the layout of the experimental setup, three calibration procedures which derive from the ECM are considered here. These are a two-step, a three-step, and the double pass calibration from reference [\[64\]](#). The user can choose any of these calibration procedures depending on the polarisation properties of the system to be calibrated.

The double pass, the two-, and the three-step calibration procedures are explained here after introducing the ECM. To validate the use of the three-step calibration method, the polarisation properties of the confocal polarisation microscope are measured for different pinhole sizes. The parameters from the Lu-Chipmann decomposition discussed in section [3.1.2](#) are used to quantify these properties.

6.1 Unveiling the ECM

The eigenvalue calibration method (ECM) was originally introduced by Compain and Drevillon [\[72\]](#) in 1999. Their motivation was to calibrate efficiently a polarimeter. During the course of almost 20 years several methods were introduced for example by Azzam [\[73, 74\]](#), Sabatke [\[57\]](#) and Brudzewski [\[68\]](#). All with the disadvantage of requiring a priori information of the system such as the generated polarisation states or at least the angular position of the polarisation optics. The ECM on the

other hand proved to be more general and reliable.

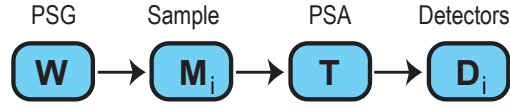


Figure 6.1: Block diagram of Müller matrix polarimetry

The ECM is an ingenious method based on control theory and linear algebra. The objective of the method is to find $\mathbf{W}$, $\mathbf{T}$ and $\mathbf{M}$ in the linear system of figure (6.1) when only $\mathbf{D}$ is known. To do this the polarisation element $\mathbf{M}$ used as calibration sample is changed until the system of equations is overdetermined and the solution is unique. To reach a solution nevertheless, three assumptions need to be made: the noise in the system is negligible, any polarisers used as calibration samples must be perfect, and in general the calibration samples used must have a given Müller matrix. The reason for these assumptions becomes more obvious when introducing the method in more detail.

6.1.1 The Heart of the Method

The Müller matrix polarimeter depicted in figure (6.1) can be mathematically expressed as

$$\mathbf{D}_i = \mathbf{T}\mathbf{M}_i\mathbf{W}, \quad (6.1)$$

where $\mathbf{M}_i$ is the Müller matrix of the i th polarisation element used as a calibration sample, $\mathbf{D}_i$ is in general a $q \times n$ matrix with irradiance readings as entries, $\mathbf{T}$ is the $q \times 4$ matrix known as instrument matrix, and $\mathbf{W}$ is a $4 \times n$ matrix formed by n generated Stokes vectors. For simplicity however, it is assumed that $q = 4$ and $n = 4$. If equation (6.1) is left multiplied by the inverse of the measurement of air, i.e.

$$\mathbf{D}_{air}^{-1} = (\mathbf{T}\mathbf{M}_i\mathbf{W})^{-1} = (\mathbf{T}\mathbf{W})^{-1}, \quad (6.2)$$

one obtains

$$\mathbf{C}_i = \mathbf{D}_{air}^{-1}\mathbf{D}_i = \mathbf{W}^{-1}\mathbf{M}_i\mathbf{W}, \quad (6.3)$$

where $\mathbf{C}_i$ and $\mathbf{M}_i$ are similar matrices. This means that $\mathbf{C}_i$ and $\mathbf{M}_i$ have the same characteristic polynomial, and hence the same eigenvalues[47]. Given now that $\mathbf{M}_i$

can be written in general as (see section 3.1.2)

$$\mathbf{M}_i = \mathbf{R}(\theta) \begin{pmatrix} 1 & -\cos 2\Psi_i & 0 & 0 \\ -\cos 2\Psi_i & 1 & 0 & 0 \\ 0 & 0 & \sin 2\Psi_i \cos \Delta_i & \sin 2\Psi_i \sin \Delta_i \\ 0 & 0 & -\sin 2\Psi_i \sin \Delta_i & \sin 2\Psi_i \cos \Delta_i \end{pmatrix} \mathbf{R}(-\theta), \quad (6.4)$$

and that the eigenvalues of $\mathbf{M}_i$ are given by

$$\mu_1 = 2\tau \sin^2(\Psi_i), \quad \mu_2 = 2\tau \cos^2(\Psi_i),$$

$$\mu_3 = \tau \sin(2\Psi_i) \exp(i\Delta_i), \quad \mu_4 = \tau \sin(2\Psi_i) \exp(-i\Delta_i),$$

it is thus possible to partly reconstruct $\mathbf{M}_i$ from the eigenvalues of $\mathbf{C}_i$. As it is seen later, to fully reconstruct $\mathbf{M}_i$ the angle θ is found by optimising the solution of the system of equations. Note that it is only possible to write equation (6.2) with the assumption that the Müller matrix of air is the identity matrix. It is here where noise can become an issue. With the assumption of a noise free system, the problem is now simplified to finding $\mathbf{W}$ from equation (6.3), and then using equation (6.2) to find $\mathbf{T}$.

6.1.2 Solving a particular case of the Sylvester equation

Multiplying equation (6.3) by $\mathbf{W}$ on both sides and rearranging yields

$$\mathbf{M}_i \mathbf{W} - \mathbf{W} \mathbf{C}_i = 0, \quad (6.5)$$

which is a particular case of the Sylvester equation $\mathbf{M} \mathbf{W} - \mathbf{W} \mathbf{C} + \mathbf{K} = \mathbf{0}$ when $\mathbf{K} = \mathbf{0}$. The Sylvester equation is used in control theory [75], neural networks [76] and in optics for wavefront reconstruction [77]. A more complete literature review can be found in reference [78]. Common methods to solve equation (6.5) are by means of Schur decomposition with backward substitution [79], by diagonalisation [80], and by column stacking. The ECM uses the last two methods in the following way.

To solve equation (6.5) using the diagonalisation and the column stacking method,

it is necessary to first introduce the vec operator[81]. This is an operator that stacks the columns of a $n \times m$ matrix to form a $1 \times n * m$ vector, and it has the property

$$\text{vec}(\mathbf{J}\mathbf{K}\mathbf{F}) = (\mathbf{F}^T \otimes \mathbf{J})\text{vec}(\mathbf{K}), \quad (6.6)$$

where $\mathbf{J}$, $\mathbf{F}$, and $\mathbf{K}$ are arbitrary matrices, T represents transpose, and $\otimes$ denotes Kronecker product[35, 36, 37]. It is possible to use equation (6.6) to rewrite equation (6.5) as[82]

$$(\mathbf{I} \otimes \mathbf{M}_i - \mathbf{C}_i^T \otimes \mathbf{I}) \text{vec}(\mathbf{W}) = 0 \quad (6.7)$$

$$\mathbf{H}_i \text{vec}(\mathbf{W}) = 0, \quad (6.8)$$

where $\mathbf{I}$ is the 4×4 identity matrix and $\mathbf{H}_i$ is a 16×16 matrix. It then follows as shown in Appendix (B) that if $\mathbf{H}_i$ is Hermitian, $\mathbf{W}$ exists in the eigenspace of $\mathbf{H}_i$. However, given that in general $\mathbf{H}_i$ is not Hermitian, to satisfy this condition it is necessary to multiply by $\mathbf{H}_i^T$ on both sides of equation (6.8) yielding

$$\mathbf{H}_i^T \mathbf{H}_i \text{vec}(\mathbf{W}) = 0. \quad (6.9)$$

In addition given that $\mathbf{W}$ can be anywhere in the 16×16 —dimensional eigenspace of $\mathbf{H}_i^T \mathbf{H}_i$, the next step to solve equation (6.9) is to use the null space method[83]. In this method n measurements are taken with different polarisation elements as calibration samples. An analogous equation to equation (6.9) is written for each measurement and then linearly combined to form

$$\mathbf{L} \text{vec}(\mathbf{W}) = 0, \quad (6.10)$$

where

$$\mathbf{L} = \mathbf{H}_1^T \mathbf{H}_1 + \dots + \mathbf{H}_n^T \mathbf{H}_n. \quad (6.11)$$

As also shown in Appendix (B), it then follows that if the n calibration samples are chosen appropriately, $\mathbf{W}$ is in fact the eigenvector of $\mathbf{L}$ corresponding to the null eigenvalue of $\mathbf{L}$ itself. The matrix $\mathbf{L}$ however must have only one null eigenvalue to ensure that the solution is unique. If for example two eigenvalues of $\mathbf{L}$ are

zero, $\text{vec}(\mathbf{W})$ lies in a 16-dimensional hyperplane defined by the two eigenvectors corresponding to the null eigenvalues. If three eigenvalues are zero, then $\text{vec}(\mathbf{W})$ lies in a 16-dimensional hypervolume and so forth.

To ensure that $\mathbf{W}$ is the unique solution of equation (B.7), it is needed to choose n calibration samples that reduce the dimension of the null space of $\mathbf{L}$ to one¹. In theory[72] it is possible to use a perfect polariser as long as $\theta \neq 0$ or $\theta \neq \pi/2$ and an ellipsometric surface with $\Phi \neq 0$ or $\pi/4$ and $\Delta \neq 0$ or π . Given now that the use of these two polarisation elements as calibration samples ensures that only one eigenvalue of $\mathbf{L}$ is zero, it then follows that the angles θ_i of these elements can be found by minimizing

$$\epsilon = \sqrt{\frac{\lambda_{16}}{\lambda_{15}}}. \quad (6.12)$$

That is the ratio of the smallest over the second smallest eigenvalue. This minimization however, only yields the relative angular orientation of the polarisation elements. Since the eigenvalues of a matrix are invariant to rotation, it is possible to write $\mathbf{M}'_i = \mathbf{R}(\theta')\mathbf{M}_i\mathbf{R}(-\theta')$ for the n calibration elements and still obtain the same eigenvalues but a different $\mathbf{W}$. In consequence the frame of reference must be set by the user. This is done by fixing the angular orientation of one the polarisation elements. To avoid experimental errors the frame of reference should be given by the angular position of the first polariser after the source. Once $\mathbf{W}$ is found, $\mathbf{T}$ is finally obtained from

$$\mathbf{T} = \mathbf{D}_{air}\mathbf{W}^{-1} = (\mathbf{TW})\mathbf{W}^{-1}. \quad (6.13)$$

Note that equation (6.12) is a metric that quantifies the error in the calibration. The main sources of error are imperfections in the polarisation elements, laser noise, and electronic noise. A statistical comparison of different measurements of $\mathbf{D}_{air}$ can quantify the contribution of the laser and electronic noise to ϵ .

The ECM as described so far can be used to calibrate a setup laid out in a transmission configuration. To calibrate the setup built for this research (see section 5.1), it is necessary to complement the calibration with a few more measurements as explained next.

¹The number of solutions is given by the dimension of the null space of the problem[83], see Appendix (B).

6.2 Calibration of the system

Due to the layout of the setup, see figure (5.1), three calibration procedures are considered. These are a two-step, a three-step, and a double pass calibration[64]. A brief introduction to these procedures is given next, and a more detailed description is given in the corresponding sections.

The experimental setup used in this research can be depicted by its block diagram shown in figure (6.2). In this figure the bold letters represent the Müller matrix of each unit: **N** represents the confocal microscope, formed by the imaging optics **F** (including the pinhole) and a flat mirror **E** as sample; **B** is the Müller matrix of the beamsplitter, and **Q** represents the setup excluding the PSA **T** and the PSG **W**.

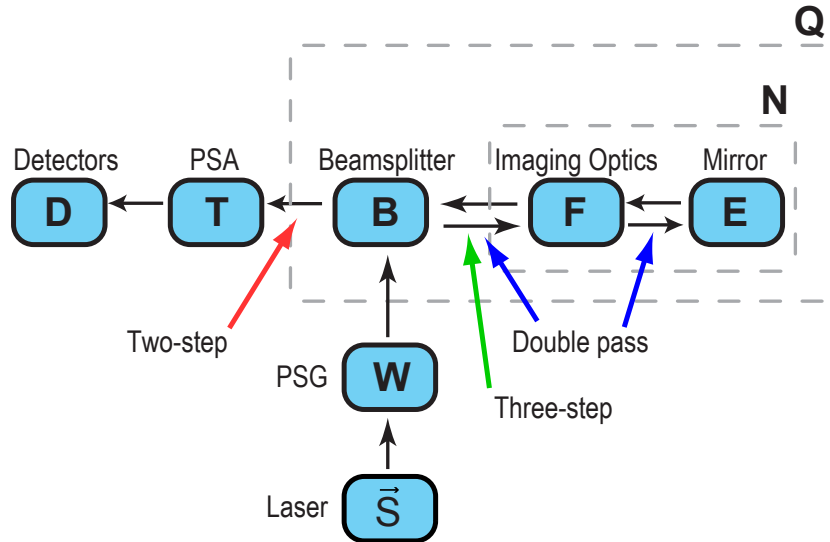


Figure 6.2: Simplified schematic of the confocal polarisation microscope.

The two-step calibration process yields **W**, **T**, and **Q**. This means that during a scan the polarisation is effectively being measured between the beamsplitter and the PSA, i.e. where the red arrow is pointing in figure (6.2).

The double pass procedure yields **W**, **T**, **B**, and **F** as long as it is assumed that **E** is a perfect mirror. The polarisation can effectively be measured during a scan either between the relay optics and the mirror, i.e. at the object plane, or between the beamsplitter and the relay optics.

The three-step procedure is a combination of the two-step, the double pass calibration, and a fitting algorithm. This procedure yields **W**, **T**, **B**, and **N** and it is

also possible to obtain $\mathbf{F}$ and $\mathbf{E}$. The polarisation can thus be measured during a scan in the same locations as the double pass procedure.

Regardless of the method used, the position in the setup where the polarisation is effectively being measured depends on where the calibration samples are placed. Due however to the field interactions at the pinhole (see section 1.4.1), it is imperative to calibrate the system at the image plane. The field distribution at the pinhole depends on the spatial variations of phase, amplitude, and polarisation at the back focal plane of the confocal lens. If a spatially homogenous calibration sample is used here, the field distribution at the pinhole is affected. The spatial imperfections in the imaging system can be classified as either phase[16] or polarisation aberrations[17], which should be corrected locally using for example an imaging polarimeter with a CCD camera, a wavefront sensor, and a spatial light modulator. The image plane is located at the pinhole and hence is inaccessible in practice. If the pinhole however is sufficiently small, the calibration samples can be placed between the pinhole and the beamsplitter (green arrow).

6.2.1 Two-step calibration

The two-step calibration is suitable for cases when $\mathbf{Q}$ only shows retardance and when one is solely interested in the relative polarisation properties of the sample. As long as $\mathbf{Q}$ does not show diattenuation or depolarisation, the information in the polarisation is mapped by the retardance in $\mathbf{Q}$ and not lost. The efficiency of this method can be assessed by doing a Lu-Chipman decomposition on $\mathbf{Q}$ after calibration. It is not possible to use the two-step process if the resulting polarisance, diattenuation, and depolarisation are high.

This calibration is a two step process. First $\mathbf{W}$ and the product $\mathbf{TQ}$ are obtained by calibrating with the polarisation elements $\mathbf{M}_i$ placed between the PSG and the beamsplitter. Note that it is not yet possible to obtain $\mathbf{T}$ alone since $\mathbf{D}_{air} = \mathbf{TQW}$. In the second step of the calibration the polarisation elements are placed between the PSA and the beamsplitter. This step yields $\mathbf{T}$ and the product $\mathbf{QW}$. The Müller matrix of the system $\mathbf{Q}$ then follows from $\mathbf{T}^{-1}\mathbf{TQ}_1$ or $\mathbf{Q}_2\mathbf{W}\mathbf{W}^{-1}$. In theory $\mathbf{Q} = \mathbf{Q}_1 = \mathbf{Q}_2$ regardless of the chosen product, however, due to electronic noise they are not in practice. The norm of the difference between $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can thus be

used as an error estimator. It is important to mention that the sample that defines the reference axis must be the same in both steps.

The efficiency of the calibration procedure explained above was assessed using a flat mirror in the object plane of the high NA objective lens. The chosen polarisation elements for the calibration were[5] a polariser at $= 0^\circ$, one at $\approx 90^\circ$, and a quarter waveplate at $\approx 30^\circ$. The pinhole used was filtering 2/3 of the Airy disc. The measured $\mathbf{Q}$ showed a retardance of $\lambda/24$, polarisance and diattenuation of 0.015, and 0.008 of depolarisation. The relative error in $\mathbf{Q}$ was 8.5×10^{-4} which corresponds to the order of magnitude of the electronic noise.

6.2.2 Double pass calibration

The double pass calibration[64] is a variant of the ECM introduced for systems where the light passes twice through a polarisation element $\mathbf{M}_i$. To use this calibration it is necessary to assume that the light passing back through $\mathbf{M}_i$ is reflected by a perfect mirror at normal incidence, and that the behaviour of the polarisation elements is independent of the direction of propagation of the light. The validity of these assumptions can be assessed using an algebraic procedure as explained later.

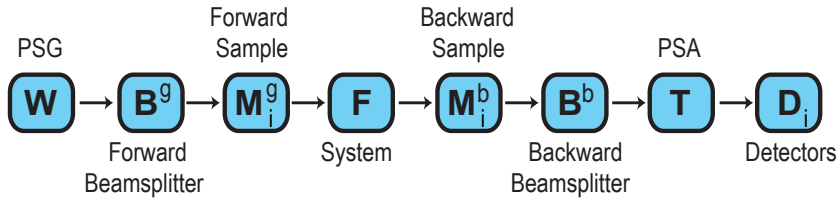


Figure 6.3: Block diagram representing the path of the light in the confocal polarisation microscope built for this research

In systems such as the one depicted in figure (6.2) light passes twice through the same polarisation element $\mathbf{M}_i$, and in consequence equation (6.1) is no longer of the same form. Adapting figure (6.2) to figure (6.3) and following the direction of propagation, it can be seen that equation (6.1) is replaced by

$$\mathbf{D}_i = \mathbf{T} \mathbf{B}^b \mathbf{M}_i^b \mathbf{N} \mathbf{M}_i^f \mathbf{B}^f \mathbf{W}, \quad (6.14)$$

where f and b denote the forward and backward direction of propagation, respec-

tively. The analogous equation to equation (6.3) is

$$\mathbf{C}_i^{dp} = \mathbf{W}^{-1}(\mathbf{B}^f)^{-1}\mathbf{K}\mathbf{B}^f\mathbf{W}, \quad (6.15)$$

where $\mathbf{K} = \mathbf{N}^{-1}\mathbf{M}_i^b\mathbf{N}\mathbf{M}_i^f$. Note from equation (6.15) that $\mathbf{C}$ and $\mathbf{K}$ are similar matrices, however, $\mathbf{K}$ and $\mathbf{M}_i^b$ are not. The latter means that the eigenvalues of $\mathbf{K}$ are no longer independent of the angular orientation of $\mathbf{M}_i$. Unless $\mathbf{M}_i^b$ and $\mathbf{N}$ commute, it is not possible to find $\mathbf{M}_i$ and hence $\mathbf{H}$ can not be reconstructed.

To find $\mathbf{W}$ using equation (6.15) let one assume that $\mathbf{N} = \mathbf{E}$, where $\mathbf{E}$ represents a perfect mirror at normal incidence given by

$$\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (6.16)$$

Also assume that the polarisation elements used for calibration are independent of the propagation direction. The latter assumption means that the Müller matrices of each polarisation element can be written as

$$\mathbf{M}_i^f = \mathbf{R}(\theta)\mathbf{M}_i\mathbf{R}(-\theta), \quad (6.17)$$

and

$$\mathbf{M}_i^b = \mathbf{R}(-\theta)\mathbf{M}_i\mathbf{R}(\theta) \quad (6.18)$$

Substituting equations (6.16), (6.17) and (6.18) into equation (6.15) yields

$$\mathbf{C}_i^{dp} = \mathbf{W}^{-1}(\mathbf{B}^f)^{-1}\mathbf{R}(-\theta)\mathbf{M}_i\mathbf{M}_i\mathbf{R}(\theta)\mathbf{B}^f\mathbf{W}, \quad (6.19)$$

if $\mathbf{E}$ is chosen to commute to the left, or

$$\mathbf{C}_i^{dp'} = \mathbf{W}^{-1}(\mathbf{B}^f)^{-1}\mathbf{E}^{-1}\mathbf{R}(\theta)\mathbf{M}_i\mathbf{M}_i\mathbf{R}(-\theta)\mathbf{E}\mathbf{B}^f\mathbf{W}, \quad (6.20)$$

if $\mathbf{E}$ is chosen to commute to the right. In both cases it is possible to reconstruct $\mathbf{H}$ and solve the problem using the null space method described in section (6.1.1).

The choice of the direction of commutation depends on the angular position of the polarisation elements. If the elements are placed at θ , $\mathbf{E}$ commutes to the left. For $\mathbf{E}$ to commute to the right the polarisation elements are still placed at θ , but the data is processed as if the element was placed at $-\theta$. In the first case the calibration yields $\mathbf{B}^f \mathbf{W}$ and $\mathbf{T} \mathbf{B}_i^g$. In the second case however, the calibration yields $\mathbf{T} \mathbf{B}_i^g \mathbf{E}$ and $\mathbf{B}^f \mathbf{W}$. In theory both cases should yield the same matrices unless the assumptions were invalid. It is thus possible to verify the validity of the assumptions by comparing $\mathbf{T} \mathbf{B}_i^g \mathbf{E}$ and $\mathbf{T} \mathbf{B}_i^g$ for example.

Two cases were used to test the double pass calibration method. One case consisted of the calibration of the system using a flat mirror but removing the pinhole and the objective lens. In this case the assumptions mentioned above were valid. The relative error between the matrices obtained from the left and right commutation is 1.2×10^{-4} . The second test consisted of calibrating the system with a flat mirror but with the objective lens in place and a pinhole of $2/3$ of the Airy disc. In this case the relative error was 0.03. Even though the first case showed satisfactory results, it is not always possible to guarantee that the path of the light is the same with and without the pinhole and the objective lens. The results from the second case were unsatisfactory, however, these results can be improved with the three-step calibration process.

6.2.3 Three-step calibration

The three-step calibration method can be used if the behaviour of the polarisation elements $\mathbf{M}_i$ depends on the direction of propagation of the light and even if the mirror is not perfect. This calibration method is a combination of the double pass calibration, the two-step method and a fitting algorithm. The aim is to find $\mathbf{T}$, $\mathbf{W}$, the polarisation elements $\mathbf{M}_i^f$ and $\mathbf{M}_i^b$ used as calibration samples, $\mathbf{B}^f$, $\mathbf{B}^b$, and $\mathbf{N}$.

The calibration procedure is as follows. First a set of measurements are taken placing the n calibration samples in three different places of the setup: between the PSG and the beamsplitter, between the beamsplitter and the PSA, and between the beamsplitter and the pinhole. Since the first two positions correspond to where the measurements of the two-step calibration are taken, it is possible to obtain the Müller matrices of the calibration samples in both directions $\mathbf{M}_i^f$ and $\mathbf{M}_i^b$, $\mathbf{T}$, $\mathbf{W}$,

and $\mathbf{Q}$. Note that $\mathbf{Q} = \mathbf{B}^b \mathbf{N} \mathbf{B}^f$ and can be written using equation (6.6) as

$$\text{vec}(\mathbf{Q}) = \mathcal{B} \text{vec}(\mathbf{N}), \quad (6.21)$$

where $\mathcal{B} = (\mathbf{B}^f)^T \otimes \mathbf{B}^b$. The measurements taken between the beamsplitter and the pinhole correspond to the double pass method, and these are given by

$$\mathbf{D}_i = \mathbf{T} \mathbf{B}^b \mathbf{M}_i^b \mathbf{N} \mathbf{M}_i^f \mathbf{B}^f \mathbf{W}. \quad (6.22)$$

By now multiplying by the inverse of $\mathbf{T}$ and $\mathbf{W}$, equation (6.22) can be simplified to

$$\mathbf{J}_i = \mathbf{T}^{-1} \mathbf{D}_i \mathbf{W}^{-1} = \mathbf{B}^b \mathbf{M}_i^b \mathbf{N} \mathbf{M}_i^f \mathbf{B}^f. \quad (6.23)$$

Equation (6.6) is again used together with the multiplicative property² of the Kronecker product to write equation (6.23) as

$$\text{vec}(\mathbf{J}_i) = [(\mathbf{B}^f)^T \otimes \mathbf{B}^b] \left[(\mathbf{M}_i^f)^T \otimes \mathbf{M}_i^b \right] \text{vec}(\mathbf{N}). \quad (6.24)$$

The linear combination of the n measurements can thus be given by

$$\mathcal{H} = \mathcal{B} (\mathcal{I} + \dots + \mathcal{M}_i + \dots + \mathcal{M}_n) \text{vec}(\mathbf{N}), \quad (6.25)$$

where $\mathcal{M}_i = (\mathbf{M}_i^f)^T \otimes \mathbf{M}_i^b$, and the 16×16 identity matrix from the measurement of air is represented by $\mathcal{I}$. Note the correspondence between equations (6.21) and (6.25). The remaining matrices $\mathbf{B}^f$, $\mathbf{B}^b$, and $\mathbf{N}$ can be found from these two equations using a Levenberg-Marquardt algorithm[84]. Dr. van de Nes developed this algorithm which was capable of solving equations (6.21) and (6.25) using 4×4 random matrices as entries with an accuracy comparable with the electronic noise of the experiment.

The three-step calibration method was tested using four polarisation elements as calibration samples. These were a polariser at 0° , one at $\approx 90^\circ$, a quarter waveplate at $\approx 30^\circ$, and a $\lambda/8$ waveplate at $\approx 30^\circ$. The relative error between the measured $\mathbf{Q}$ and the retrieved product $\mathbf{B}^b \mathbf{N} \mathbf{B}^f$ was 6.5×10^{-6} .

To approve the reliability of the three-step method it is necessary to carry out

² $\mathbf{AB} \otimes \mathbf{CD} = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$

further experiments. For example using an object whose Müller matrix is known instead of a flat mirror. Later in chapter 7 the Müller matrix of the MODS³ pits is compared to an FDTD⁴ simulation. For a purely empirical demonstration, the next section shows the dependance of the polarisation in the system due to the size of the pinhole.

6.3 Pinhole polarisation dependance

The polarisation dependance of the pinhole was measured to corroborate the calibration of the system when using the three-step method. With a flat mirror in the object plane of the objective lens, the Müller matrix $\mathbf{N}$ was measured for different sizes of the pinhole. Figure (6.4) shows the results from the polar decomposition of the different $\mathbf{N}$ measured. The top left and right plots show the polarisation and diattenuation, respectively. The bottom left and right plots show the retardance and the depolarisation, respectively.

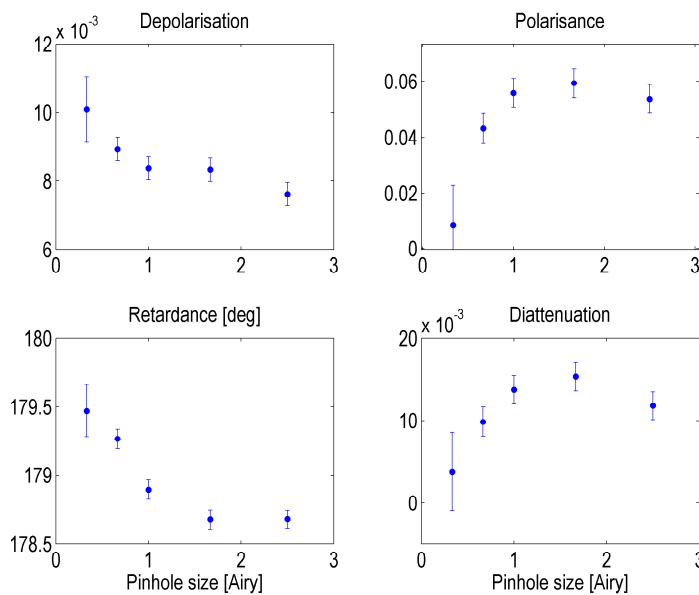


Figure 6.4: Parameters from the polar decomposition of the confocal signal produced by a mirror for different pinhole diameters

It is possible to see from figure (6.4) that apart from the depolarisation, the trend

³Multiplexed optical data storage

⁴Finite Difference Time Domain

of the plots show that as the pinhole is smaller, the measurement is more of a perfect mirror. Given that the level of light decreases as the size of the pinhole decreases, the contribution of the noise to the measurement increases and the depolarisation thus increases.

6.4 Conclusion

It was seen in this chapter that it is possible to calibrate a confocal polarisation microscope in reflection configuration. The calibration can be performed using any of the three methods introduced which derive from the eigenvalue calibration method (ECM). As it was seen the choice of the method depends on the polarisation properties of the system.

The two-step method is ideal for systems that only show retardance and if the user is solely interested in the relative polarisation properties of the sample. This method can be used for systems configured in reflection or in transmission.

The double pass calibration can be used for systems where the light passes twice through the calibration samples $\mathbf{M}_i$, i.e. systems in reflection configuration. To use this calibration it is necessary to assume that the light passing back through $\mathbf{M}_i$ was reflected by a perfect mirror at normal incidence. It is also necessary to assume that the behaviour of the polarisation elements is independent of the direction of propagation of the light.

The three-step calibration method can be used if the behaviour of the polarisation elements $\mathbf{M}_i$ depends on the direction of propagation of the light and even if the mirror is not perfect. This calibration method is a combination of the double pass calibration, the two-step method and a Levenberg-Marquardt fitting algorithm.

It was also mentioned that regardless of the chosen calibration method, it is imperative to calibrate the system at the image plane. This is due to the interactions of the field happening at the pinhole. The spatial imperfections in the imaging system classify as either phase or polarisation aberrations, and these aberrations should be corrected locally using for example an imaging polarimeter with a CCD camera, a wavefront sensor, and a spatial light modulator.

Finally the three-step method was corroborated by measuring the Müller matrix

of the imaging system. The results showed that the calibration method is acceptable.

The optical data storage results are given in the next chapter. That is the climax of the thesis.

Chapter 7

MODS

The research culminates in this chapter. The answers to the questions that motivated the building of a state-of-the-art confocal polarisation microscope are given next. Is it possible to increase the current optical storage capacity? Can one store information using the polarisation of the light? If so what storage increase one can expect? The short answer to these questions is yes, yes, and at least 7.16 times more than any existing non-volumetric optical technology.

The following pages explain the background and the principle behind the storage method termed MODS (multiplexed optical data storage). Müller matrix images of the prototype and the method to retrieve the data are given. The method is reviewed for three different incident polarisation states.

The first optical data storage (ODS) device was commercially introduced in a large scale in 1982 even though it was patented much earlier in 1952[85]. The CD-ROM was the first non-volumetric optical device capable of storing up to 640Mb of information. Later on in 1995 the single layer DVD was introduced increasing the storage capacity to 4.7Gb. The dual layer version came later with 8.5Gb of storage space[86, 87, 88, 89, 90]. Nowadays it is possible to find a double sided and dual layer DVD that can store 15.9Gb. Despite of this significant increase of storage capacity two new technologies were introduced during 2007. Those were the Blu-ray Disc (BD) with a capacity of 25Gb and the now superseded HD-DVD with 15Gb.

The evolution of the optical disk has been possible mainly thanks to three breakthroughs in engineering. These are the improved design and manufacturing techniques of higher NA objective lenses, the availability of low cost lasers with a shorter

wavelength, and the scope of writing/reading media. Since a disk read-out system is a diffraction limited system, the storage capacity was increased by increasing the NA of the objective lens and decreasing the wavelength. Encoding also played an important role in the storage increase. The technology however seems to have reached a point where storing more information in the same physical space is no longer possible, at least by using conventional methods and without sacrificing costs. Decreasing the wavelength furthermore requires the use of CaF lenses which increases the production costs. Increasing the NA even more requires an immersion technique[1], which can yield potentially difficult technical challenges. For example, an air gap between the objective lens and the cover layer of 30nm would need to be maintained with an accuracy of $\pm 3\text{nm}$. Temperature changes would be enough to destabilise the system. A different approach is needed to overcome these technical issues.

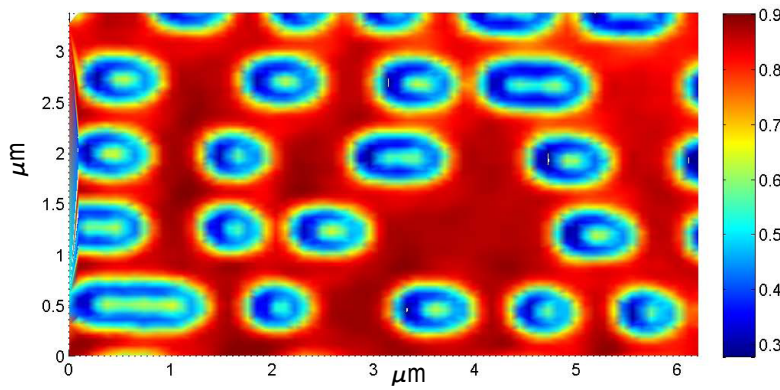


Figure 7.1: Confocal image of the tracks in a DVD

New ideas were developed in the beginning of the millennium with the aim of increasing the storage capacity of the optical disk. The SLAM project (Super Laser Array Memory)[2] was initiated in 2001 with the objective of finding methods to multiplex information. Instead of storing the information in pits and lands as in figure (7.1), i.e. 0 and 1, the aim was to introduce a range of logical levels. With this approach it is possible to increase the storage capacity and keep the same storage area. The number of data bits that can be stored in a pit follows the relationship $\text{Log}_2(n)$, where n are the available levels. If for example 64 levels are available, then every pit can store 6 bits of information. The storage capacity is thus increased 6 times.

Different ideas were considered in the SLAM project[91]. These were plasmon resonance encoding using metallic nanoparticles, phase singularity encoding from trenches and ridges, encoding via polarisation conversion from a reflective step, helical structure encoding, and gray scale encoding using an asymmetric pit. In most of these approaches several disadvantages were found such as high production cost, complexity and high sensitivity. The most convenient method identified was polarisation conversion. This method consisted of using a shaped pit that would exhibit form birefringence¹

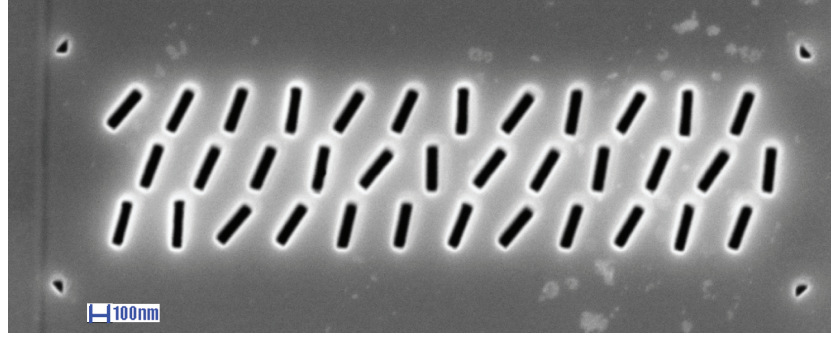


Figure 7.2: MODS Pits

The features considered in this work for data storage are the rectangular pits shown in figure (7.2). These pits were engraved by Dr Attila Tóth² in a silicon substrate using electron beam milling. The picture was taken by Dr Tóth using an scanning electron microscope. Each groove is 50nm wide, 200nm long, $\lambda/4$ deep, and 200nm apart. The corresponding angular orientations given in degrees and with respect to the horizontal axis are

50, 63, 71, 82, 58, 66, 88, 54, 81, 60, 86, 70,
 70, 66, 67, 79, 51, 89, 53, 60, 81, 68, 56, 85,
 75, 87, 50, 57, 77, 79, 69, 51, 72, 60, 76, 70.

The angles range from 50° to 89° and where chosen with the objective of maximising cross-talk.

The motivation for the rectangular shape originates from studying the behaviour

¹Property of an object of changing the state of polarisation of the light depending on the geometry of the object.

²Research Institute for Technical Physics and Materials Science, Budapest, Hungary

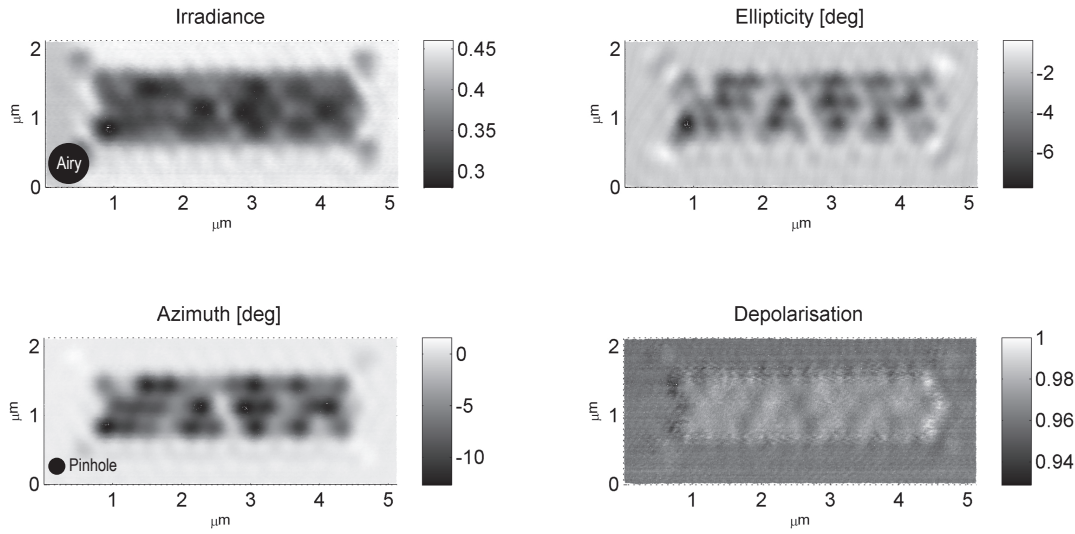


Figure 7.3: Stokes parameters measured when illuminating the engraved pits with linearly polarised light

of nanorods. It has been shown[92] that only the electric field perpendicular to the rod is coupled and re-emitted. This means that if the state of polarisation of the re-emitted field can be related to the orientation of the rod, information can then be stored using the angle at which the rod is placed. The practical realisation and use of nanorods in optical data storage is not feasible. This is true despite the recent publication [93] which claimed that such a solution is feasible. As stated before, the SLAM project considered this solution about 8 years prior to the publication of [93] and concluded that mass manufacturing of a disk containing nanorods would not be possible. A subsequent patent by some of the original researchers of SLAM considered the construction of WORM or WMRM disks using nanorods embedded in gel. It was suggested that a laser beam would melt and liquefy the gel and another separate beam would then trap and rotate the nanorods into the required orientation. However, the solution turned out not to be feasible.

It was found that it is possible to retrieve the angular position of the pits from the azimuthal angle of the re-emitted field. The dynamic range however depended on the polarisation state of the illumination. Take for example figure (7.3) and figure (7.4). The first one shows the Stokes parameters measured when illuminating the pits with linearly polarised light and the second one when using circularly polarised illumination. Note the improvement of the spatial resolution in azimuth and

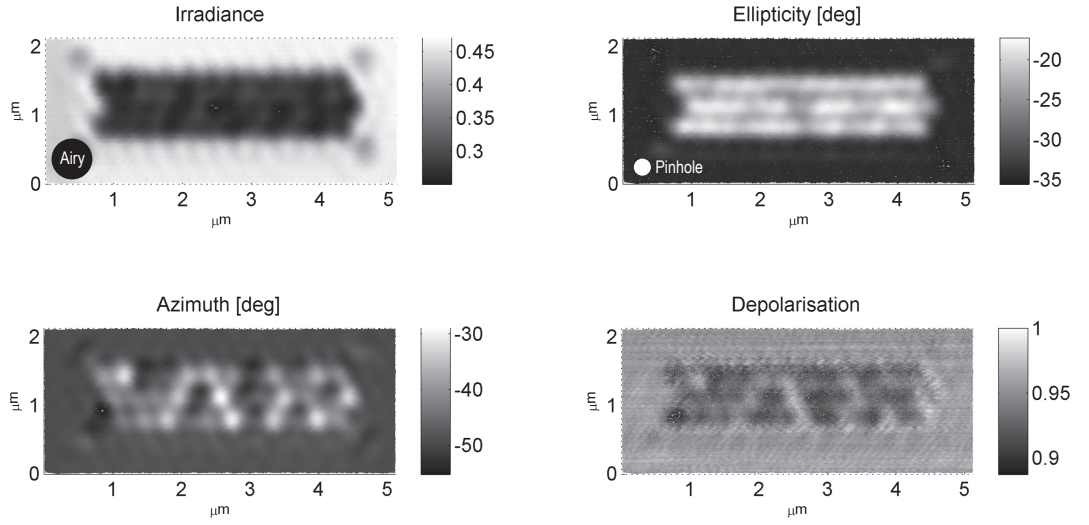


Figure 7.4: Stokes parameters measured when illuminating the engraved pits with circularly polarised light

ellipticity plots compared to the irradiance. It is possible to see from these figures that the gray scale modulation in the plots depends on the state of polarisation of the illumination. In the case of linearly polarised illumination the irradiance, azimuth angle, and ellipticity plots show modulation while the depolarisation does not. When using circularly polarised illumination the modulation appears mostly in the azimuth angle and depolarisation plots. In both cases however it is possible to measure the angle of the pits. Figure (7.5) shows a line scan of the azimuthal plots when using linearly (left) and circularly (right) polarised illumination. The blue and red markers show the experimental and expected values, respectively. The coincidence is remarkable. Note however that the angular range of the azimuthal plot is smaller when using the linearly polarised illumination. The greatest deviation of the linearly polarised case is $\epsilon_r = 1.03^\circ$ but the dynamic range of the measurement is $D_r = 11.3^\circ$. This means that in the case of having pits engraved at an angle between 0° and 180° , it is possible to resolve 50 levels which is equivalent to an increase of 5.65 times. This increase can be found explicitly from

$$\text{Increase} = \log \left(\frac{180^\circ \times D_r}{39^\circ \times \epsilon_r} \right), \quad (7.1)$$

where the 39° is the angular range at which the pits were engraved.

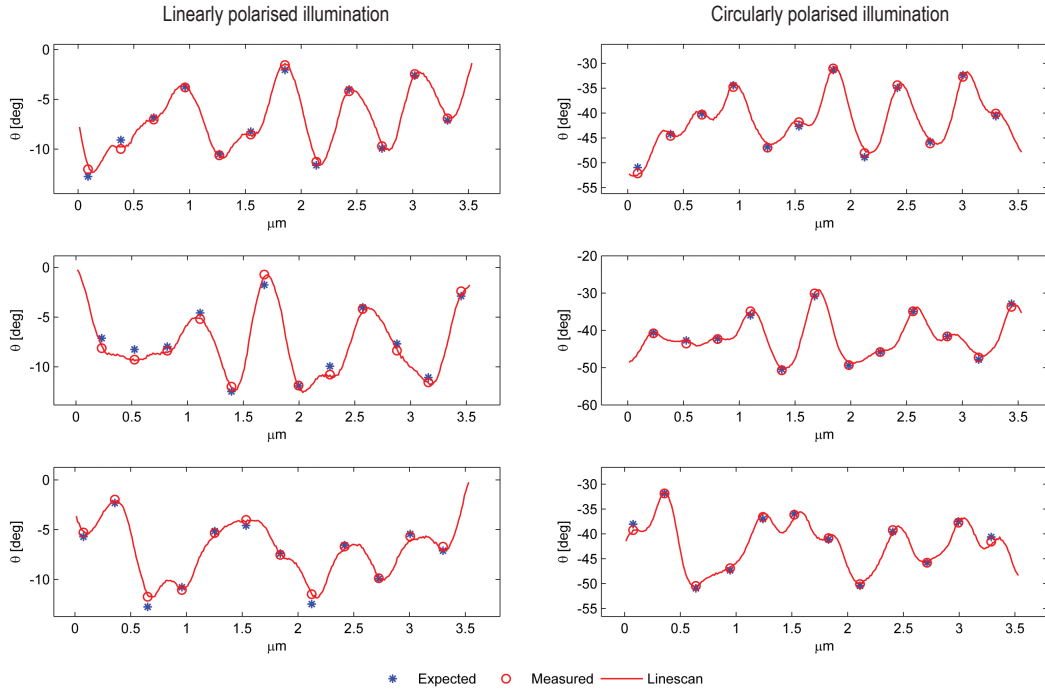


Figure 7.5: Linescans of the polarisation images obtained when illuminating the pits with linearly (left column) and circularly (right column) polarised illumination. Each row corresponds to a rows in figure (7.2) and each marker to a pit.

In the case of circularly polarised light the dynamic range is 22.1° and the error is 1.22° . This is equivalent to 84 levels and an increase of 6.39 times. The error is greater in pits which are not perfect rectangles but rhomboids. The direct relationship between the measured range of azimuth angles and the polarisation state of the illumination is left as future work. The circularly polarised illumination nevertheless corresponds to one of the states that yield the greatest storage increase.

A set of optimum polarisation states that yield the greatest increase in data storage can be found by using the illumination engineering method explained in section (4.3). In this method the aim is to find the set of polarisation states that are equally transformed by all the pixels of the image. To do so the Müller matrix of the pits is measured, decomposed, and analysed.

The Müller matrix of the engraved pits is shown in figure (7.6). This measurement was taken with a pinhole of $1/3$ of the Airy disk. The fact that elements off the diagonal are non zero suggests that the pits modify the state of polarisation of the light to a certain extent. To identify and quantify these changes it is convenient

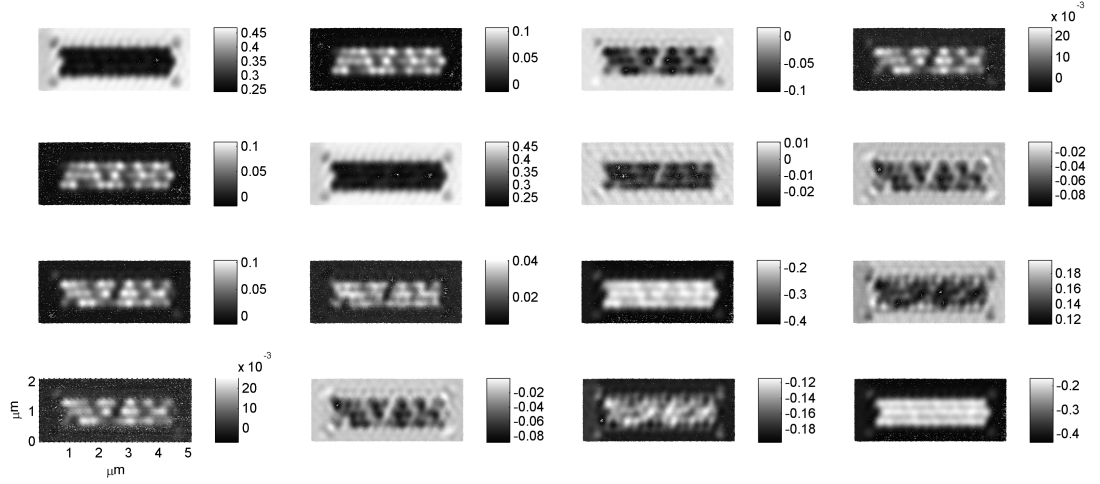


Figure 7.6: Müller matrix of the MODS pits

to decompose the Müller matrix into a retarder, a diattenuator, and a depolariser matrix and find the magnitude of each.

The Müller matrix decomposition considered here is the one proposed by Lu and Chipman (see section 3.1.2). Figure (7.7) shows the magnitude of the retardance (top left), depolarisation (bottom right), polarisance (bottom left), and diattenuation (top right) obtained from the decomposition. It can be seen that the pits polarise and diattenuate up to 50% of the light, depolarise 10%, and induce, at most, a retardance of $\lambda/20$. This means that the pits predominantly influence the degree of polarisation and the reflection of a certain component of the field.

Note from figure (7.7) the coincidence between polarisance and diattenuation. This is a typical behaviour of an imperfect analyser-polariser. By plotting the diattenuator and the polarisance vectors in the Poincaré sphere, it is possible to see which type of analyser-polariser the pits are. Figure (7.8) shows the Poincaré representation of the polarisance (left) and diattenuation (centre) vectors, and for sake of completeness the fast axis of the retarder matrix (right). Each marker in these plots represents the end point of the previously mentioned vectors from a pixel in the image. The colour represents the different orientations at which the pits were measured. The red markers denote the polarisation properties of the pits when placed perpendicular to the reference axis³. The green ones were obtained after rotating the entire set of pits approximately 45°. Blue and black represent a rotation of 90°

³Axis set by the first polariser

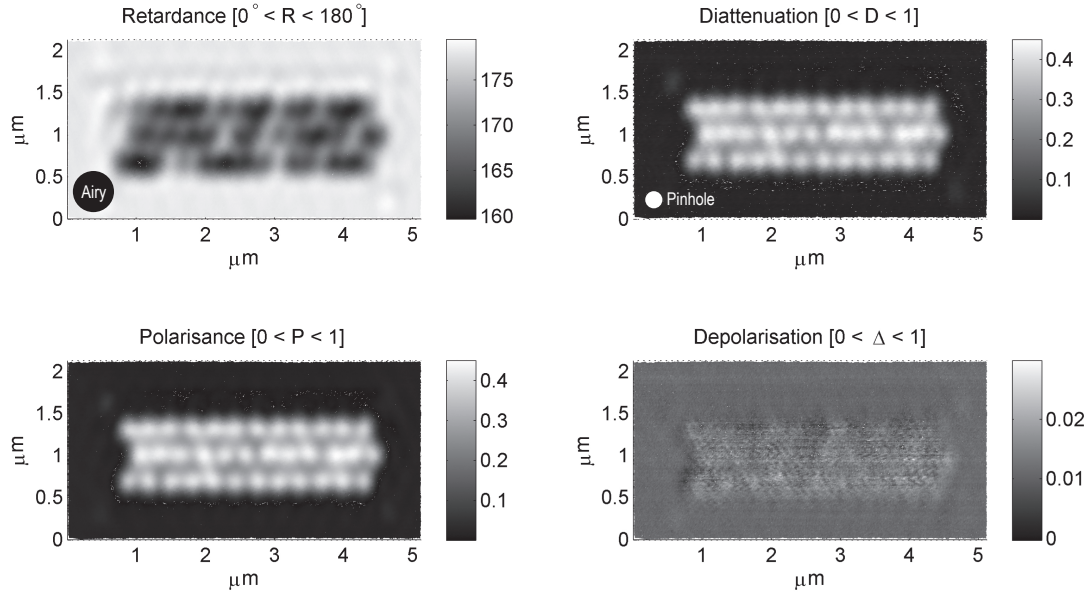


Figure 7.7: Magnitude of the retardance (top left), depolarisation (bottom right), polarisation (bottom left), and diattenuation (top right) induced by the MODS Pits to the state of polarisation of the illumination

and 135° , respectively. It is possible to conclude that the pits behave in general as an elliptical polariser-analyser, and they span a range of 180° . The markers however are closer to the equator than to the zenith of the Poincaré sphere. This suggest that the pits affect more the azimuth angle of the polarisation illumination than the ellipticity.

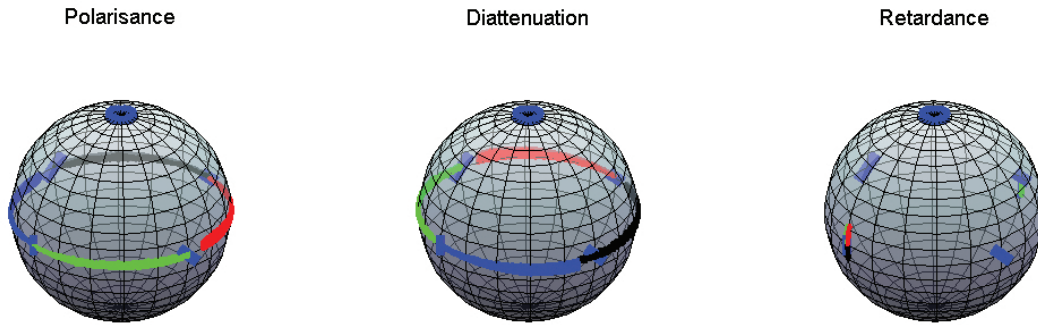


Figure 7.8: Poincaré sphere representation of the polarisance vector (left), diattenuator vector (centre), and fast axis (right) corresponding to each pixel of the image

Figure (7.8) also represents the eigenvectors of the respective Müller matrices. It thus follows from the illumination engineering method (see section 4.3) that right and left circularly polarised states form part of the optimum set. Given now that

by definition the null vector is contained in a set of this nature, unpolarised light is also considered as an optimum polarisation state for illumination.

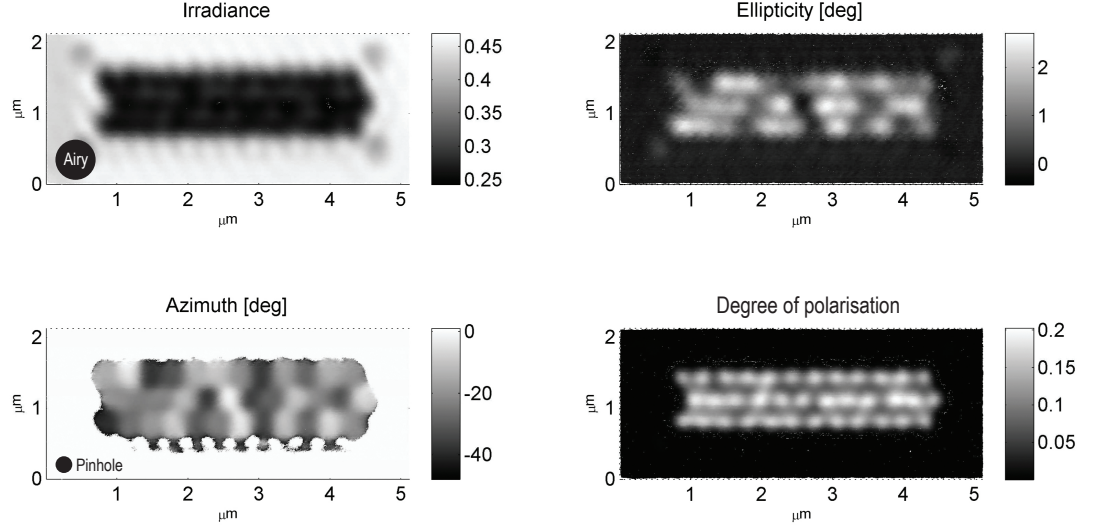


Figure 7.9: Stokes parameters measured when illuminating the pits with unpolarised light

The possibility of using unpolarised illumination for data storage is investigated next. The use of this type of illumination can reduce the costs of an optical disk drive and avoid the use of lasers. Figure (7.9) shows the Stokes parameters measured when illuminating the pits with unpolarised light. Note how the spatial resolution of the depolarisation plot is dramatically increased. The line scans obtained from the azimuth angle resulting from unpolarised illumination is shown in the left hand side of figure (7.10). The right hand side shows the line scans from the azimuth angle extracted from the diattenuation vector of the Müller matrix. In both cases the range is 36.3° , and the greatest deviation is 1.17° . This corresponds to an increase of 7.16 times since the number of available levels is 143.

The difference in the spatial resolution between the irradiance plots and the rest of the plots is due to the type of detection. Take again the example from section 4.1.1. Two dipoles positioned one Airy unit apart and emitting different fields. As it was shown in figure (4.8) the irradiance is homogeneous to a certain extent along the horizontal axis, however, the polarisation changes significantly along the same axis. The resolution increment then follows from the fact that the microscope can measure the state of polarisation of the field and measure in consequence these changes.

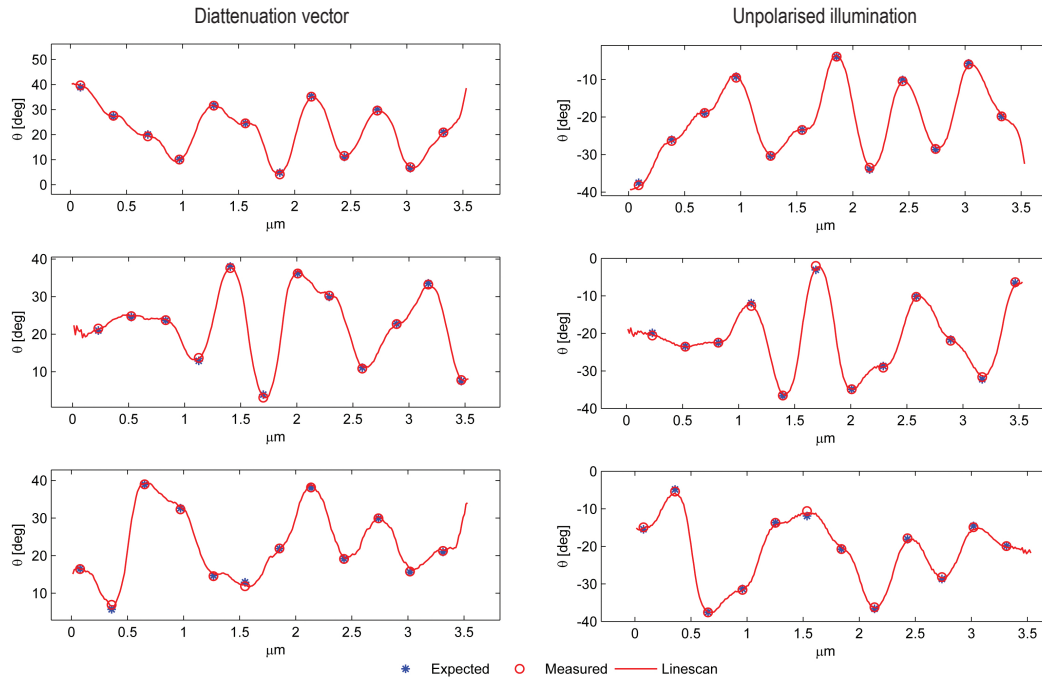


Figure 7.10: Left: Linescans obtained from the image representing the azimuth angle of the polarisation vector. Right: Linescans obtained from the image representing the azimuth angle of the diattenuation vector

The difference in the angular range obtained with unpolarised and with polarised light is due to the degree of polarisation and the type of detection. The polarisation resolution of the system decreases as two dipoles emitting different polarisation states are brought together. This only happens when the re-emission process is fully polarised. In the case of unpolarised light the field of the dipoles add up in phase and in intensity, however, the contribution between the phase of the fields of the dipoles averages to zero over time. Note in addition that the actual angular range of the pits is 39° and the measured one is 36.15° . It appears intuitively that the size of the pinhole is the reason for this difference. A concise investigation using finite difference time domain (FDTD) simulations is left as future work.

Apart from the illumination it was also necessary to investigate common readout errors such as tracking, and windowing. The tracking error quantifies the freedom in the system of not reading out the data from the exact centre of the pit. The windowing error is defined in case the user needs to set an integration time during scanning. Figure (7.11) shows the storage increase as a function of the tracking error for circularly polarised (right) and unpolarised (left) illumination. These plots

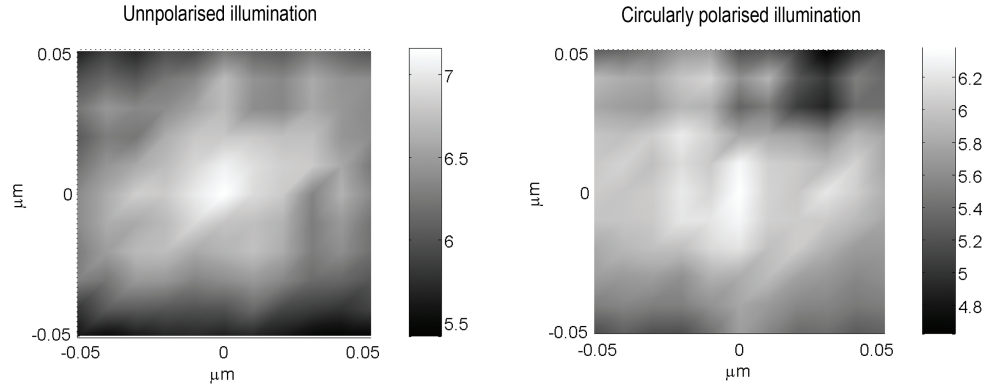


Figure 7.11: Storage capacity as a function of the tracking error when illuminating with circularly polarised light (right), and with unpolarised light (left).

where obtained by displacing the readout signal from the centre of the pits and using the largest standard deviation obtained to compute the new storage increase. The vertical and horizontal axis represents such displacement in micrometers.

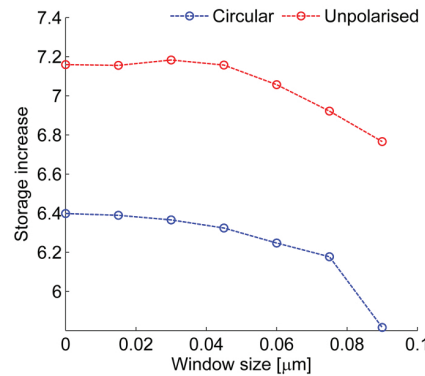


Figure 7.12: Storage capacity as a function of the windowing error obtained when illuminating with circularly polarised light(blue), and with unpolarised light (red). The vertical axis represents the storage increase, and the horizontal axis denotes half of the length of the window

The storage capacity as a function of the windowing error is shown in figure (7.12). This error was calculated by increasing the range of integrated points along the readout direction. For example in a window of 20nm the data obtained from the points within 10nm before and 10nm after the centre of the pit is integrated. The graph represents the storage increase obtained for different lengths of windows. It is possible to see that in the case of unpolarised illumination (red) the storage capacity increases to 7.18 when the window is 60nm long. This is an interesting fact

since the width of the pit is 50nm. The blue line represents the windowing error obtained with circularly polarised illumination.

It is possible to conclude that the polarisation encoding technology introduced in this chapter is promising. Up to 160Gb of information can be stored in a single layer and single sided disk when using circularly polarised light. In the case of using unpolarised light the storage capacity can be up to 179Gb. The use of unpolarised light in addition does not require the use of lasers which in principle could reduce the cost of disk read out systems. It is worth mentioning that the polarisation encoding technology is compatible with existing binary formats, even if the NA is further increased and the wavelength decreased. In fact the technology increases the signal to noise ratio of conventional binary formats. This can be seen from the fact that the polarisation images shown here have a greater spatial resolution and dynamic range than the irradiance images. The increase of the resolution and the range is even greater in the depolarisation plots when using unpolarised illumination.

Conclusion

The main focus of this research was to build a confocal polarisation microscope that could reveal if it is possible to utilise multiplexed information in optical data storage using the polarisation of the light. The development and construction of the microscope resulted in a number of interesting findings that were discussed in this thesis.

Multiplexed optical data storage was shown to be a viable alternative for the next generation of optical data storage. It is possible with this technology to store up to 7.16 times more information than any other non-volumetric solutions. Taking the Bluray for example, the storage capacity of a single layer and single sided disk can reach up to 179Gb if unpolarised light is used. If circularly polarised light is used, the storage capacity can reach 160Gb.

An efficient method to measure the longitudinal component of a dipole field in the image plane were proposed. It was found that by breaking the symmetry of the back focal plane field distributions resulting from an electric dipole it is possible to determine the longitudinal component of the dipole moment. Future work is planned to include experimental verification of the principles of operation of the method.

A transfer function was introduced with the aim of explaining the propagation of spatial frequencies through a high NA system. This function termed HTF showed that the resolution in the spatial frequency domain depends strongly on the state of polarisation emitted by the imaged object. The HTF can explain resolution in function of the magnitude of the field. The concept of polarisation resolution still remains undefined in a rigorous way. It was shown however, that it is possible to observe in the polarisation images shown in chapter 8 a greater spatial resolution and dynamic range than in the irradiance images. It was also shown that the resolution and the range is even greater in the depolarisation plots when using unpolarised

light for illumination.

The concept of degree of purity was introduced in order to understand the concept of resolution in the polarisation domain. The understanding was aided by the investigation of the extinction ratio for different polarisation states. It was concluded that neither the imaging system nor the pinhole affect significantly the acquisition of the polarisation images of a single dipole. A pinhole of 1 Airy units in diameter yields acceptable results. The importance of the pinhole in high NA imaging polarimetry arises when imaging more complex objects. It was seen that in the case of finite pinhole the measured degree of polarisation is non-zero due to both temporal and spatial effects. A fully temporally polarised beam can appear as partially polarised if the beam contains a spatial polarisation distribution.

A possible solution to conclude the ongoing discussion in the literature regarding polarisation and partial coherence was given. It was proposed that the term “coherence” should refer only to the scalar properties of a beam. Only in this context it is possible to guarantee that a measure exists to solely quantify temporal and spatial wavefront fluctuations. A novel metric to account for the spatial variations of the polarisation was introduced. This metric was termed the degree of spatial polarisation. The quantity commonly known as the degree of polarisation was termed the degree of temporal polarisation and derived in a novel way.

From the calculated PSF matrix of the system (Müller matrix due to a free dipole), it is possible to conclude that a confocal micropolarimeter behaves as a linear analyser-polariser. The imaging system does not depolarise temporally but it does spatially. The only retardance introduced by the system corresponds to the π phase shift of the Airy pattern. The experimental proof of the latter is considered as future work.

Three algorithms derived from the eigenvalue calibration method were introduced to calibrate the microscope. Those were a two-step, double-pass, and three-step method. It was seen that the choice of the method depends on the polarisation properties of the system. The two-step method is ideal if one is solely interested in the relative polarisation properties of the sample and for systems that only show retardance. This method can be used for systems configured in reflection or in transmission. The double pass calibration can be used for systems where the light

passes twice through the calibration samples $\mathbf{M}_i$, i.e. systems in reflection configuration. To use this calibration it is necessary to assume that the light passing back through $\mathbf{M}_i$ was reflected by a perfect mirror at normal incidence. It is also necessary to assume that the behaviour of the polarisation elements is independent of the direction of propagation of the light. The three-step calibration method can be used if the behaviour of the polarisation elements $\mathbf{M}_i$ depends on the direction of propagation of the light and even if the mirror is not perfect. This calibration method is a combination of the double pass calibration, the two-step method and a Levenberg-Marquardt fitting algorithm. It is imperative to calibrate the system at the image plane, e.g. after the pinhole.

Optical and mechanical parts of the system were introduced individually. An interferometric characterisation of the objective lens was given. It was shown that it is possible to perform a stage scan with a precision of $\pm 5\text{nm}$ by pre-shaping the driving waveform. It was also mentioned that the measured axial and transversal resolutions are 437nm and 352nm respectively. Regarding the polarimetric part of the system it was possible to reach a condition number of 1.739 for the PSG and of 1.86 for the PSA.

The idea of using a Müller matrix polarimetry to read out the information from an optical disk is nowadays unpractical. However using a Wollaston prism to read out the data is ideal as a first step. As future work, it is possible to investigate the influence of the geometry of the pit to the ellipticity of the illuminating field, and then use a complete polarimeter. Adding a new variable would increase further the storage capacity of the optical disk. It is also possible to investigate the use of the degree of polarisation and the irradiance. A total of four independent variables could then be used to store information. If one day Müller matrix polarimetry proves more practical, it could be possible to use the eight independent variables of a Müller matrix to store information. It is worth considering the latter as future work since the idea of having only eight levels in eight variables would result in 24 times more storage. The read out of data from a write once disk is also left as future work.

Bibliography

- [1] S. M. Mansfield and G. S. Kino. Solid immersion microscope. *Appl. Phys. Lett.*, 57:2615–2616, 1990.
- [2] S. F. Pereira et al. Report on reading/recording results with most promising method. *Internal report*, 2003.
- [3] A.S. van de Nes P. Török C. Rockstuhl, S. F. Pereira and M. Salt. Investigations on different multiplexing approaches in optical data storage systems. *SLAM*, 2000.
- [4] P. Munro (12.5) C Rockstuhl (12.5) PT (25), Martin Salt (25) and HP Herzig (12.5). High density optical data storage. *International patent application*, 2003.
- [5] A. De Martino, K. Yong-Ki, E. Garcia-Caurel, B. Laude, and B. Drevillon. Optimized mueller polarimeter with liquid crystals. *Optics Letters*, 28(8):616–618, 2003.
- [6] T. Wilson and C. Sheppard. Theory and practice of scanning optical microscopy. 1984.
- [7] Richard Feynman. *The Feynman Lectures on Physics: Volume 2*, volume 2 of *The Feynman Lectures on Physics*. Addison-Wesley, Boston, 1963.
- [8] J. D. Jackson. *Classical Electrodynamics*. WILEY, 1998.
- [9] M. Born and E. Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, Cambridge, England, 1999.

- [10] R. M. A. Azzam and N. M. Bashara. *Ellipsometry and Polarized Light*. Amsterdam: North-Holland, 1977.
- [11] G. W. Raymond. *Fourier Series and Optical Transform Techniques in Contemporary Optics: An Introduction*. WILEY, London, England, 1996.
- [12] P.D. Higdon P.Török and T. Wilson. On the general properties of polarising conventional and confocal microscopes. *Opt. Commun.*, 148:300–315, 1998.
- [13] E. Hecht. *Optics (4th Edition)*. Addison Wesley, 2001.
- [14] A.S. van de Nes. *Rigorous Electromagnetic Field Calculations for Advanced Optical Systems*. PhD thesis, Delft University of Technology, 2005.
- [15] P. R.T. Munro. *Application of numerical methods to high numerical aperture imaging*. PhD thesis, Imperial College of Science, Technology and Medicine, London, England, 2006.
- [16] D. Malacara. *Optical shop testing*. John Wiley, 2007.
- [17] R. A. Chipman. *Polarization Aberrations*. PhD thesis, THE UNIVERSITY OF ARIZONA., 1987.
- [18] R. M. A. Azzam. *M. A. Azzam, Selected Papers on Ellipsometry*, volume 27. Washington: Bellinham, 1991.
- [19] B. R. Masters. Selected papers on confocal microscopy. 1996.
- [20] P. Török. *DSc thesis: An imaging theory for advanced, high numerical aperture optical microscopes*. PhD thesis, Budapest, Hungary, 2004.
- [21] J. W. Goodman. *Introduction to Fourier Optics*. Roberts & Company Publishers, 2004.
- [22] C. J. R. Sheppard, M. Gu, Y. Kawata, and S. Kawata. Three-dimensional transfer functions for high-aperture systems. *J. Opt. Soc. Am. A*, 11(2):593–598, 1994.
- [23] G. B. Arfken and H. J. Weber. *Mathematical Methods for Physicists*, volume Fifth. Academic Press, San Diego, 2001.

-
- [24] R. N. Bracewell. *Fourier Transform and Its Applications*. McGraw-Hill Education, 1980.
- [25] M.R. Foreman, C.Macías-Romero, and P. Török. Determination of the three-dimensional orientation of single molecules. *Opt. Lett.*, 33(9):1020–1022, 2008.
- [26] P. Török. Propagation of electromagnetic dipole waves through dielectric interfaces. *Opt. Lett.*, 25:1463, 2000.
- [27] P. Török, P.D. Higdon, and T. Wilson. Theory for confocal and conventional microscopes imaging small dielectric scatterers. *J. Mod. Opt.*, 45:1681, 1998.
- [28] D J. Innes and A.L. Bloom. Design of optical systems for use with laser beams. *Spectra-Physics Laser Technical Bulletin*, 5:1, 1966.
- [29] B. Richards and E. Wolf. Electromagnetic diffraction in optical systems - ii structure of image field in an aplanatic system. *Proc. Roy. Soc. A-Math Phy*, 253:358–379, 1959.
- [30] S.S. Sherif M.R. Foreman and P. Török. Photon statistics in single molecule orientational imaging. *Opt. Express*, 15:13597, 2007.
- [31] H.C. van de Hulst. *Light scattering by small particles*. Dover Publications, 1981.
- [32] R. Dorn, S. Quabis, and G. Leuchs. Sharper focus for a radially polarized light beam. *Phys. Rev. Lett.*, 91(23):233901, 2003.
- [33] N. Wiener. Generalized harmonic analysis. *Acta Mathematica*, 55(1):117–258, 1930.
- [34] J. J. Gil. Polarimetric characterization of light and media. physical quantities involved in polarimetric phenomena. *Eur. Phys. J. Appl. Phys.*, 40:1–47, 2007.
- [35] H. Neudecker. A note on kronecker matrix products and matrix equation systems. *Journal on Applied Mathematics*, 17 (3):603–606, 1969.
- [36] J. W. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions On Circuits And Systems*, 25 (9), 1978.

- [37] A. Graham. *Kronecker Products and Matrix Calculus with Applications*. Chichester: Ellis Horwood Limited, 1981.
- [38] U. Fano. A stokes-parameter technique for the treatment of polarization in quantum mechanics. *Phys. Rev.*, 93(1):121–123, 1954.
- [39] E. Wolf. Optics in terms of observable quantities. *Il Nuovo Cimento*, 12(6):884–888, 1954.
- [40] E. Wolf. Coherence properties of partially polarized electromagnetic radiation. *Il Nuovo Cimento*, 13(6):1165–1181, 1959.
- [41] G. B. Parrent and P. Roman. On the matrix formulation of the theory of partial polarization in terms of observables. *Il Nuovo Cimento*, 15(3):370–388, 1960.
- [42] R. Barakat. Degree of polarization and the principal idempotents of the coherency matrix. *Optics Communications*, 23(2):147 – 150, 1977.
- [43] E. Wolf. *Introduction to the Theory of Coherence and Polarization of Light*. New York: Cambridge University Press, 2001.
- [44] A. A. Michelson. *Experimental determination of the velocity of light*. Washington, 1882.
- [45] A. Fresnel. *Considérations mécaniques sur la polarisation de la lumière*. USA: Dowden, Hutchinson and Ross, 1975.
- [46] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*. UK:Cambridge University Press, 1995.
- [47] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [48] R. Clark Jones. A new calculus for the treatment of optical systems v. a more general formulation, and description of another calculus. *J. Opt. Soc. Am.*, 37(2):107, 1947.
- [49] G. G. Stokes. *On the composition and resolution of streams of polarized light from different sources*. USA: Dowden, Hutchinson and Ross, 1975.

-
- [50] H. Poincaré. *Concerning theoretical descriptions of Polarized light*. USA: Dowden, Hutchinson and Ross, 1975.
- [51] J. S. Tyo. Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic errors. *Appl. Opt.*, 41:619–630, 2002.
- [52] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. Philadelphia: SIAM, 1997.
- [53] R. M. A. Azzam. Division-of-amplitude photopolarimeter (doap) for the simultaneous measurement of all four stokes parameters of light. *Opt. Acta*, 29: 685–689, 1982.
- [54] Jacques Hadamard. Sur les problemes aux derivees partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.
- [55] G. W. Stewart A. K. Cline, C. B. Moler and J. H. Wilkinson. An estimate for the condition number of a matrix. *SIAM J. Numer. Anal.*, 16(2):368–375, 1979.
- [56] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Baltimore: Johns Hopkins University Press, 1996.
- [57] M. R. Descour W. C. Sweatt J. P. Garcia E. L. Dereniak S. A. Kemme D. S. Sabatke, A. M. Locke and G. S. Phipps. Figures of merit for complete stokes polarimeter optimization. *Proc. SPIE*, 4133:75, 2000.
- [58] G. W. Stewart. *Introduction to Matrix Computations*. New York: Academic Press, 1973.
- [59] A. Ruz. Merits for computing. *Proc. SPIE*, 43:25, 2009.
- [60] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. Philadelphia: Society for Industrial and Applied Mathematics, 2002.
- [61] S. N. Savenkov. Optimization and structuring of the instrument matrix for polarimetric measurements. *Opt. Eng.*, 41:965–972, 2002.

- [62] A. van der Sluis. Condition numbers and equilibration of matrices. *Numer. Math.*, 14:14–23, 1969.
- [63] J. M. Bueno and M. C. W. Campbell. Confocal scanning laser ophthalmoscopy improvement by use of mueller-matrix polarimetry. *Opt. Lett.*, 27(10):830–832, 2002.
- [64] D. Lara and C. Dainty. Double-pass axially resolved confocal mueller matrix imaging polarimetry. *Opt. Lett.*, 30(21):2879–2881, 2005.
- [65] P. R. T. Munro and P. Török. Properties of high-numerical-aperture mueller-matrix polarimeters. *Opt. Lett.*, 33(21):2428–2430, 2008.
- [66] S. Inoue. Studies on depolarization of light at microscope lens surfaces i. the origin of stray light by rotation at lens surfaces. *Exp. Cell Res.*, 3:199–208, 1952.
- [67] J. L. Pezzaniti and R. A. Chipman. Mueller matrix imaging polarimetry,. *Opt. Eng.*, 34:1558–1568, 1995.
- [68] K. Brudzewski. Static stokes ellipsometer: general analysis and optimization,. *J. Mod. Opt.*, 38:889–896, 1991.
- [69] S. Lu and R. A. Chipman. Interpretation of mueller matrices based on polar decomposition. *J. Opt. Soc. Am. A*, 13(5):1106–1113, 1996.
- [70] T. Setälä A. Shevchenko A. T. Friberg K. Lindfors, A. Priimagi and M. Kaivola. Local polarization of tightly focused unpolarized light. *Nature photonics*, 1:228 – 231, 2007.
- [71] E. Compain and B. Drevillon. Broadband division-of-amplitude polarimeter based on uncoated prisms. *Appl. Opt.*, 37:5938–5944, 1998.
- [72] S. Poirier E. Compain and B. Drevillon. General and selfconsistent method for the calibration of polarization modulators, polarimeters, and mueller matrix ellipsometers. *Appl. Opt.*, 38:3490–3502, 1999.

-
- [73] I. M. Elminyaw R. M. A. Azzam, E. Masetti and A. M. El-Saba. Construction, calibration, and testing for a four-detector photopolarimeter. *Rev. Sci. Instrum.*, 59:84–88, 1988.
 - [74] R. M. A. Azzam and A. G. Lopez. Accurate calibration of the four-detector photopolarimeter with imperfect polarizing elements. *J. Opt. Soc. Am. A*, 6:1513–1521, 1989.
 - [75] M. Darouach. Solution to sylvester equation associated to linear descriptor systems. *Systems Control Lett.*, 55(10):835–838, 2006.
 - [76] V Vapnik C J C Burges. A new method for constructing artificial neural networks,. *Bell Laboratories report*, 1995.
 - [77] H. Ren and R. Dekany. Fast wave-front reconstruction by solving the sylvester equation with the alternating direction implicit method. *Opt. Express*, 12: 3279–3296, 2004.
 - [78] B. Zhou and G.R. Duana. On the generalized sylvester mapping and matrix equations. *Systems Control Lett.*, 57 (3):200–208, 2008.
 - [79] F. R. Gantmacher. *Theory of Matrices*. New York, N.Y. : Chelsea Publishing Co., 1977.
 - [80] H. T. Burgess. Solution of the matrix equation $x^{-1}ax = n$. *The Annals of Mathematics*, 19(1):30–36, 1917.
 - [81] H. V. Henderson and S. R. Searle. Vec and vech operators for matrices, with some uses in jacobians and multivariate statistics,. *JSTOR archive*, 1979.
 - [82] M.A. Epton. Methods for the solution of $axd-bxc=e$ and its application in the numerical solution of implicit ordinary differential equations,. *BIT*, 20:341–345, 1980.
 - [83] A. Pothen T. F. Coleman. The null space problem. i. complexity. *SIAM J. Algebraic Discrete Methods*, 7:527–537, 1986.

- [84] W. H. Press and Teukolsky. *Numerical recipes in C. The art of scientific computing*. Cambridge: University Press, 1992.
- [85] J. D. Zook D. Chen. An overview of optical data storage technology. *IEEE, Proceedings.*, 63:1207–1230, 1975.
- [86] A. Huijser J.Pasman G. van Rosmalen K. Schouhamer Immink G. Bouwhuis, J. Braat. *Principles of optical disc systems*. Bristol and Boston: Adam Hilger, 1985.
- [87] A. Marchant. *Optical Recording: a technical overview*. Boston: Addison Wesley, 1990.
- [88] J.P Huignard E. Spitz and C. P. Puech. Early experiments on optical disc storage. *IEEE*, 6:1413, 2000.
- [89] J. Zavislan G. Sincerbox. *Selected Papers on Optical Storage*. SBellinham: PIE Optical Engineering Press, 1992.
- [90] M. Mansuripur and G. Sincerbox. Principles and techniques of optical data storage. *Proc. IEEE*, 85(11):1780–1796, 1997.
- [91] S. F. Pereira et al token. Report on reading/recording results with most promising method. *Internal report*, 2003.
- [92] C. Rockstuhl and H. P. Herzig. Calculation of the torque on dielectric elliptical cylinders. *J. of the Opt. Soc. of A*, 22:109–116, 2005.
- [93] J. Chon P. Zijlstra and M. Gu. Five-dimensional optical recording mediated by surface plasmons in gold nanorods. *Nature*, 459:410–413, 2009.
- [94] C. Jones. *The Jones Calculus*. USA: Dowden, Hutchinson and Ross, 1975.
- [95] A. J. Laub. Matrix analysis for scientists and engineers. *Society for Industrial and Applied Mathematics*, 143, 2005.

Appendix A

Jones calculus

The Jones calculus was first derived by Clark Jones[94] in 1941. Nowadays it is an important theoretical tool in polarisation optics. This calculus describes the interaction of fully polarised monochromatic light with non depolarising media. This appendix introduces Jones vectors and basic Jones matrices. An extensive introduction of the subject can be found in reference[10].

A Jones vector is nothing more than a simplified description of the electric vector introduced in chapter 1. Assuming that $\vec{E}(\mathbf{r}, t)$ is a collimated beam propagating parallel to the z axis, i.e. $A_z(\mathbf{r}, t)e^{i\delta_z(\mathbf{r}, t)} = 0$, the electric field can be written as

$$\vec{E}(\mathbf{r}, t) = \begin{pmatrix} A_x(\mathbf{r}, t)e^{i\delta_x(\mathbf{r}, t)} \\ A_y(\mathbf{r}, t)e^{i\delta_y(\mathbf{r}, t)} \end{pmatrix}. \quad (\text{A.1})$$

Here A_i is the complex amplitude of the i th component and δ_i the relative phase of the $i = x, y$ component. When such a field is incident upon a non depolarising object its state of polarisation is modified according to[10]

$$E_x^{out} = g_{11}E_x^{in} + g_{12}E_y^{in}, \quad (\text{A.2a})$$

$$E_y^{out} = g_{21}E_x^{in} + g_{22}E_y^{in}, \quad (\text{A.2b})$$

where $E_i^{in} = A_i^{in}e^{i\delta_i^{in}}$ and $E_i^{out} = A_i^{out}e^{i\delta_i^{out}}$ are the $i = x, y$ components of the input and the output field, respectively. Equations (A.2) can be written equivalently as

$$\vec{E}^{out} = \mathbf{G} \vec{E}^{in}, \quad (\text{A.3})$$

where the matrix $\mathbf{G}$ is known as a Jones matrix. This matrix represents the instantaneous or time invariant polarisation changes induced by a given optical element. For the purposes of this work, the matrix $\mathbf{G}$ is written explicitly^[10] in terms of its eigenpolarisations¹, and eigentransmittances² as

$$\mathbf{G} = (\chi_1 - \chi_2)^{-1} e^{-(\alpha_o + i\delta_o)} \begin{pmatrix} \psi_2 \chi_1 - \psi_1 \chi_2 & \psi_1 - \psi_2 \\ -\chi_1 \chi_2 (\psi_1 - \psi_2) & \psi_1 \chi_1 - \psi_2 \chi_2 \end{pmatrix}, \quad (\text{A.4})$$

where α_o and δ_o represent the polarisation independent attenuation and global phase, respectively. The eigentransmittances are denoted by $\psi_{1,2}$ and the eigenpolarisations by $(1, \chi_1)$ and $(1, \chi_2)$. In the case of a Cartesian coordinate system the eigensystem takes the following explicit form

$$\chi_{1,2} = \frac{\tan \theta_{1,2} + i \tan \epsilon_{1,2}}{1 - i \tan \theta_{1,2} \tan \epsilon_{1,2}}, \quad (\text{A.5a})$$

$$\psi_1 = e^{(\alpha + i\delta)}, \quad (\text{A.5b})$$

$$\psi_2 = \psi_1^{-1}. \quad (\text{A.5c})$$

Here $\theta_{1,2}$ and $\epsilon_{1,2}$ are the ellipsometric parameters³ corresponding to the eigenpolarisations given by $(1, \chi_1)$ and $(1, \chi_2)$, respectively. The relative attenuation between the eigenpolarisations is given by α , and δ represents the relative phase difference. A polarisation sensitive element is thus considered to be a “pure” diattenuator when $\delta = 0$ or a “pure” retarder when $\alpha = 0$. The following list summarises the main properties of these two typical polarisation elements.

Diattenuator. This polarisation element is also known as a polariser. It has the capability of absorbing partially or totally a component of the field. The absorption is termed diattenuation and defined as

$$D = \frac{||\lambda_1|^2 - |\lambda_2|^2|}{|\lambda_1|^2 + |\lambda_2|^2}, \quad 0 \leq D \leq 1, \quad (\text{A.6})$$

where λ_1 and λ_2 are the eigenvalues of the Jones matrix of the diattenuator.

¹The states of polarisation that will remain in the same state at the output yet attenuated by a constant factor

²Transmission coefficient corresponding to an eigenpolarisation

³Introduced earlier in chapter 2

Diattenuation appears in reflections and transmissions at oblique incidence, dichroic materials and wave guiding. The Jones matrix of a diattenuator is Hermitian, and the eigenvector of the matrix corresponding to the largest eigenvalue is called the diattenuator vector. The latter defines the diattenuation axis.

Retarder. Element with the ability to introduce a phase shift between the components of the electric field without significant attenuation. Retardance is usually achieved by birefringence, the Faraday effect, wave guiding or total internal reflection. The Jones matrix of a retarder is unitary⁴ and its eigenvectors define the fast and slow axes of the optical element.

In general either of these pure elements are said to be linear, circular or elliptical depending on the geometrical shape described by their eigenpolarisations. Pure elements in addition can be used to build more complex elements. If for example a system is formed by a variable retarder followed by a perfect polariser, a composite matrix is found by multiplying the individual matrices in the reverse order as the light is propagating through. The polarisation element of this particular example is known as an elliptical perfect analyser; this element has the ability to extinguish completely an elliptical polarisation state.

By using Jones calculus, it is now possible to determine the polarisation properties of a fully polarised beam and non depolarising elements. However, due to the lack of statistical information in the calculus, neither the Jones vectors nor the Jones matrices can describe temporal or spatial fluctuations in the field. For example the Jones calculus cannot describe the case of a field reflected from an optically varying material, i.e. an inhomogeneous material. To describe this case (often found in this work), it is necessary to consider the statistics of the beam and the inhomogeneities of the material as shown in chapters 2 and 3, respectively.

⁴Matrix whose inverse is equal to its conjugate transpose

Appendix B

ECM proofs

This appendix contains the proofs to various lemas of the Eigenvalue Calibration Method. The aim is to aid in the understanding of how to find $\mathbf{W}$ from $\mathbf{H} = 0$, where $\mathbf{H} = \mathbf{M} \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{C}^T$ (see chapter 6).

It is proven first that $\mathbf{W}$ exists in the eigensystem of $\mathbf{H}$. When $\mathbf{W}$ is written as a 1×16 vector it is in fact the eigenvector of $\mathbf{H}$ that corresponds to the null eigenvalue. It is then shown that $\mathbf{W}$ can be written as a 1×16 vector if at least $\mathbf{C}$ is Hermitian. The vector form of $\mathbf{W}$ is in fact constructed by the Kronecker product of one eigenvector of $\mathbf{M}$ with one of the eigenvectors of $\mathbf{C}$. The last proof shows that a system of equations with a null space of dimension one and containing $\mathbf{W}$ in its eigensystem can be constructed with $\mathbf{H}$.

Lemma The matrix $\mathbf{W}$ written as a 1×16 vector is the eigenvector that corresponds to the null eigenvalue of $\mathbf{H}$ [95].

Proof Let $\mathbf{X} = \{\mathbf{x}_1^T, \dots, \mathbf{x}_m^T\}$ be the eigenvectors of $\mathbf{M}$ with corresponding eigenvalues $\mu_1, \dots, \mu_m$, and $\mathbf{Y} = \{\mathbf{y}_1^T, \dots, \mathbf{y}_m^T\}$ be the eigenvectors of $\mathbf{C}$ with corresponding eigenvalues $\kappa_1, \dots, \kappa_m$. Since $\mathbf{H} = \mathbf{M} \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{C}^T$, it then follows that[37]

$$\begin{aligned} (\mathbf{M} \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{C}^T)(\mathbf{x} \otimes \mathbf{y}) &= (\mathbf{M}\mathbf{x} \otimes \mathbf{y}) - (\mathbf{x} \otimes \mathbf{C}^T\mathbf{y}) \\ &= (\mu\mathbf{x} \otimes \mathbf{y}) - (\mathbf{x} \otimes \kappa\mathbf{y}) \\ &= (\mu - \kappa)(\mathbf{x} \otimes \mathbf{y}). \end{aligned} \tag{B.1}$$

Therefore the non trivial solution is given when $(\mu - \kappa)$ is zero.

Lemma The matrix $\mathbf{W}$ can be written as $(\mathbf{x} \otimes \mathbf{y})$ if $\mathbf{C}$ is Hermitian.

Proof Note that $\mathbf{M}$ and $\mathbf{C}$ are similar matrices and assume that at least $\mathbf{C}$ is Hermitian. Let $\mathbf{M} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ and $\mathbf{C} = \mathbf{Y}\mathbf{\Lambda}\mathbf{Y}^{-1}$ be the spectral decomposition of $\mathbf{M}$ and $\mathbf{C}$, respectively. The matrix $\mathbf{\Lambda}$ is diagonal and contains the eigenvalues of $\mathbf{M}$ and $\mathbf{C}$. Rewriting $\mathbf{C} = \mathbf{W}^{-1}\mathbf{M}\mathbf{W}$ as

$$\mathbf{Y}\mathbf{\Lambda}\mathbf{Y}^{-1} = \mathbf{W}^{-1}\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}\mathbf{W} \quad (\text{B.2})$$

follows that

$$\mathbf{W} = \mathbf{X}\mathbf{Y}^{-1}. \quad (\text{B.3})$$

By then writing equation (B.3) in terms of Kronecker products as

$$\text{vec}(\mathbf{W}) = (\mathbf{X} \otimes \mathbf{Y}^{-1\text{T}})\text{vec}(\mathbf{I}), \quad (\text{B.4})$$

and since the eigenvectors of a Hermitian matrix are orthogonal, i.e. $\mathbf{Y}^{-1} = \mathbf{Y}^{\text{T}}$, then

$$\begin{aligned} \text{vec}(\mathbf{W}) &= (\mathbf{X} \otimes \mathbf{Y})\text{vec}(\mathbf{I}), \\ &= (\mathbf{x} \otimes \mathbf{y}). \end{aligned}$$

Lemma A linear system of equations denoted by $\mathbf{L}$ can be constructed to contain $\mathbf{W}$ as the unique solution.

Proof Note that the 16 eigenvalues of $\mathbf{H}$ are given by the possible combinations of $\mu_{1..4} - \kappa_{1..4}$. The number of times that $\mu - \kappa = 0$ appears can thus be given by [72]

$$\aleph = \sum_{i=1..4} \alpha^2, \quad (\text{B.5})$$

where α is the multiplicity of each eigenvalue, and $\aleph$ represents the dimension of the null space of $\mathbf{H}$. Consider now n different measurements denoted by

$$\mathbf{C}_{\mathbf{M}_i} = \mathbf{W}^{-1}\mathbf{M}_i\mathbf{W}, \quad (\text{B.6})$$

for $i = 1, \dots, n$. It is then possible to produce n different $\mathbf{H}_{\mathbf{M}_i}$ and write

$$\mathbf{L} = \mathbf{H}_{\mathbf{M}_1}^T \mathbf{H}_{\mathbf{M}_1} + \dots + \mathbf{H}_{\mathbf{M}_n}^T \mathbf{H}_{\mathbf{M}_n}. \quad (\text{B.7})$$

It then follows from equation (B.1) that $\mathbf{W}$ is now in the eigensystem of $\mathbf{L}$, and that the dimension of the null space of $\mathbf{L}$ can be reduced to one as long as $\mathbf{W}$ and $\mathbf{M}_i$ does not commute.

In summary to find $\mathbf{W}$ it is necessary to construct a system of equations $\mathbf{L}$ whose null space has dimension one. The eigenvector of $\mathbf{L}$ corresponding to the null eigenvalue is in fact $\text{vec}(\mathbf{W})$.

