Spatial mode multiplexing : from fundamental concepts to applications

Pauline Boucher

Thèse de doctorat de
Sorbonne Université

présentée par

Pauline Boucher

pour obtenir le grade de Docteur de Sorbonne Université
sur le sujet:

Spatial mode multiplexing: from fundamental concepts to applications

Laboratoire Kastler Brossel
Physique quantique et applications

Membres du jury :

Pr. Luis SÁNCHEZ-SOTO	Rapporteur
Dr. Vincent MICHAU	Rapporteur
Dr. Anne SENTENAC	Membre du jury
Pr. Agnès MAÎTRE	Membre du jury
Pr. Claude FABRE	Membre invité
Dr. Guillaume LABROILLE	Membre invité
Pr. Nicolas TREPS	Directeur de thèse

À mon grand-père Jean Boucher,

À mon père François Boucher.

Contents

Introduction

"There is a crack, a crack in everything
That's how the light gets in."

– Leonard Cohen

If the study of light as a physical phenomenon is quite old, its use to extract information about our environment might be even older. When we learn about how Thales measured the size of the Kheops pyramid, beyond teaching us about the similar triangle theorem which carries his name, this measurement uses two properties of light. First, Thales uses the fact that light travels in a straight line. Second, what allows him to measure the size of the pyramid is not the measure of any amount of light, but rather the measure of the absence of light: the shadow of the pyramid. Discriminating between the presence or absence of light is a simple measurement which can easily reach good sensitivities. These two aspects make the calculation conducted by Thales possible.

Among the different types of light which can be produced, we will be particularly interested in laser light. One of the unique features of laser light is termed *coherence*. The electromagnetic field emitted by a laser is spatially and temporally coherent: the beam which is emitted or amplified by a laser has remarkable properties of directionality, spectral purity and intensity. Ever since the first laser light emission realized by Theodore H. Maiman in 1964[1], the variety of wavelength at which laser light can be emitted has greatly increased. The operating regimes of lasers (in the pulsed regime for example, the control of the pulse duration or of the repetition rate) can also now be controlled with very good precision. The development of such remarkable light sources is the fruit of parallel efforts from fundamental and industrial research. Conversely, laser light has allowed for tremendous progress both for fundamental research but also for countless new and unexpected functions in many different types of industry [Siegman 86].

[1] We cannot fail to mention the fact that at that time, Maiman worked in an industrial laboratory.

As we presented with the example of the size of the Kheops pyramid, light is an instrument of choice to perform measurements and gain information. A first aspect of measurements performed with laser light is that they allow to reach exceptionnal precisions. For each type of measurement one wishes to perform, the type of light used is tuned. Lasers operating in the pulsed regime are used to perform time of flight measurements. Such a measurement technique allows for instance to estimate the distance between earth and the Moon with an accuracy of a few millimeters [Murphy Jr. 08]. Time-of-flight measurements are also the working principle of the lidar technology (LIght Detection And Ranging) which is now used in a huge range of fields such as topography, defense or archeology. As an example, the European Space Agency launched this summer the Aeolus observation satellite whose mission is the observation of the atmosphere's dynamic. It carries a Doppler-lidar which uses near ultraviolet light pulses to probe the wind currents of the atmosphere. Lengths can be measured using interferometry, which relies on the temporal coherence of the electric fields used. Interferometry was developed at the end of the 18^{th} century with the pioneering works of Michelson and Morley [Michelson 87, Gouy 88, Michelson 90]. Using the extraordinary coherence properties of laser light, the precision of the distance measurements achievable with interferometry has reach truly extraordinary precision: just before the start of this Ph.D, the interferometers of the Virgo and LIGO collaborations were able to measure the coalescence of two black holes at a distance of ~ 410 megaparsec (or 13 billions of billions of billions of kilometers) [Abbott 16]. In this signal, the peak displacement measured by the interferometers was of the order of 2×10^{-18} m. But light is also more simply used to image objects which may be too small, too far away, etc. Using the coherence properties of laser light, phenomena which were beforehand untracktable can now be imaged such as the dynamics of chemical reactions [Lindinger 04]. In many fields — for instance medical imaging — the superiority of information retrieval using imaging is the fact that imaging is a non-destructive measurement. The manipulation of the spatial profiles of electromagnetic fields has also allowed to make great progress in the field of imaging — in particular in situations where traditional imaging techniques failed. Structured illumination has allowed the advent of some super-resolution microscopy techniques [Hell 94, Gustafsson 01]. Adaptive optics has allowed to overcome the limitations that the atmosphere fluctuations imposed on astronomical imaging [Beckers 93, Roddier 99, Tyson 10]. More generally the field of wavefront shaping constituted a major turning point for the imaging of opaque media [Vellekoop 15] for instance in biomedical imaging [Yu 15] but also in quantum information [Defienne 16] and communication technologies [Schwartz 09].

If light allows us to gain information by measurement, it also a very powerful tool to transmit information. Smoke signals were used by native americans as an effective long-distance communication system. The Chappe telegraph system was developed at the end of the 18^{th} century and used for over 50 years. Using this system, a complete message could be transmitted over more than 200 km in less than one hour. But it is laser light which has introduced a complete upheaval in the field of telecommunications with the advent of fiber-optic communications. Developed in the 1970s, fiber-optics has played a major role in the advent of the information revolution which characterizes the last 50 years. The first transatlantic cable to use optical fiber went in operation in 1988. To this date, it is estimated that 448 optical cables lie on oceans floors. They make up for a total length of 1.2 million km [TeleGeography 18]. To increase the data rate of the information transmitted in one optical fiber, multiplexing is used. Multiplexing techniques use all degrees of freedom of the electric field: wavelength, polarization, time and phase. All these multiplexing techniques are now well established. The shaping of the intensity profile of the laser light launched into some specific types of optical fibers is the last degree of freedom left to allow the increase of the data rate transmitted through these physical channels. The current world record of data transfer in an optical fiber is of 10 Pb/s (10 millions of Gb/s) [Soma 18] and was achieved using specially engineered fiber combined with all the different types of multiplexing techniques. The spatial multiplexing was realized using the multi-plane light conversion technology (MPLC) [Morizur 10, Morizur 11b, Labroille 14] which we will introduce in this manuscript (see figure 1). To give a sense of the order of magnitude of the data rate which was achieved, 10 Pb/s, Cisco estimated in 2017 that the total IP traffic for the year 2018 would be of 102 960 Pb/month [Cisco 17], which translates to 0.04 Pb/s.

Light does not only allow to gain or transport information, it can also be used to directly manipulate matter. Modification of the intensity profile of laser beams proves useful for the optimization of a number of processes. In atomic physics for instance, two dimensional "light sheets" allow to trap atoms in a plane [Davidson 95]. Other configurations allow to arrange them in arrays [Verkerk 92, Grynberg 93] or even arbitrary 3D structures [Barredo 18]. Atoms are not the only objects which can be trapped and moved using light beams: optical tweezzers [Ashkin 86] are for instance used to manipulate or probe the mechanical properties of polymers, proteins or enzymes [Moffitt 08]. At much higher powers, the interaction between laser light and matter has found a great variety of applications in industry. The use of lasers in a number of processes often allows higher fabrication rates and improved energy efficiencies. The automotive, aerospace, medical, semiconductor and

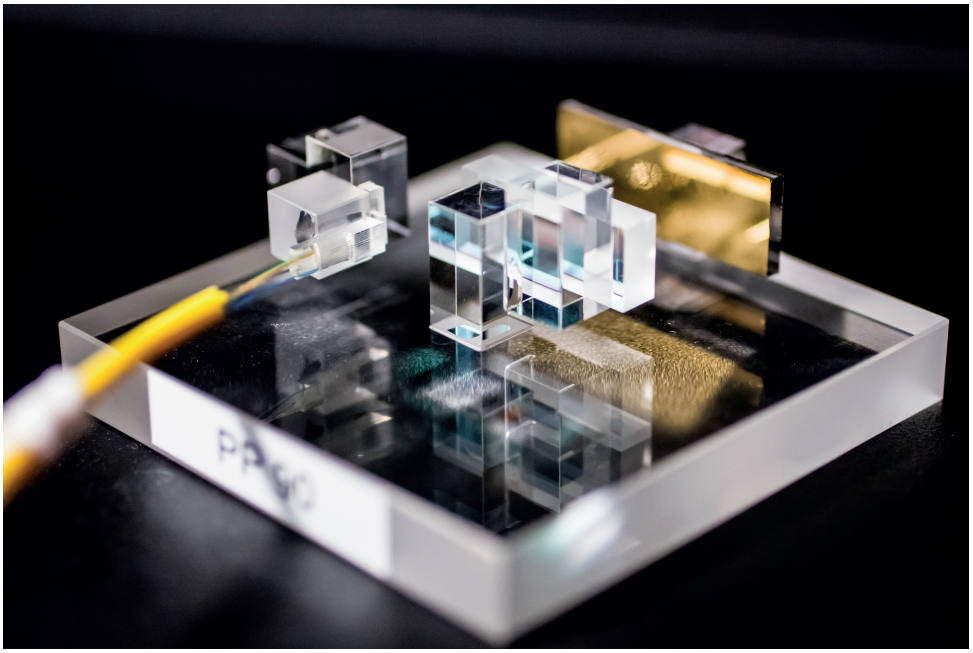

Figure 1: Physical implementation of the MPLC system presented in figure 2.9. The golden plate is the collection of phase plates while the piece of glass facing it is a plane mirror. ©CAILabs.

electrical industries are some of those who make use of these high intensity lasers. In laser processing, the light-matter interaction efficiency is affected by the laser wavelength, the power or quantity of light as well as the operating regime of the laser [Steen 91]. The engineering of the transverse profile distribution also leads to improved efficiencies [Dickey 05]. For instance, in the context of this thesis, we used shaped light pulses of femtosecond durations to drill sub-100 nm channels in glass.

The transverse intensity profile of a laser beam can be engineered in a number of ways which we will develop in this manuscript. Among them the MPLC technology will be our privileged object of study. This tool was developed as part of studies of the spatial properties of light at the quantum level and of the quantum protocols which can be implemented using spatial quantum fluctuations [Morizur 11b]. Following this work, the company CAILabs was founded in 2013 to exploit this technology. Applications range from spatial mode multiplexing in telecommunications to spatial profile shaping for laser processing. The MPLC system also finds its place in research laboratories, for instance in the context of free-space optical communications [Li 17] or quantum optics [Morizur 11a]. While it was developed in a quantum optics context, the MPLC tool is a classical object. Indeed, most of the basic concepts which are needed to describe laser physics — and all the practical details — can be described using classical tools [Siegman 86]. This also applies to the description of laser light. For this reason, except for a few specific details, we will not use in this manuscript the quantum description of the electromagnetic field and will stick to a purely classical description. This thesis was conducted within the framework of a CIFRE collaboration between the Laboratoire Kastler-Brossel — where the MPLC technology was invented — and the company CAILabs which possesses and develops it. Two different lines of research were pursued during the course of this thesis. The first one concerns the development of new use cases of the MPLC technology, which should be relevant both as relevant scientific work as well as for future industrial applications. The second aims at gaining a deeper understanding of the fundamental phenomena which underlie the MPLC technology efficiency.

This thesis is organized as follows. In the first chapter we introduce the notion of spatial modes, which can be derived from Maxwell's equations and will be the main tool used in the rest of the manuscript. We present how spatial modes can be modified and in particular introduce the MPLC technology. In the second chapter, we consider the problem of the spatial resolution of imaging systems. We then introduce the theory of parameter estimation and derive a fundamental limit to spatial resolution. We present the fact that modal decomposition is an optimal measurement scheme for the problem of the estimation of the separation of two

incoherent point sources. We finally present the experiment we conducted to validate such results. In the third chapter, we show how spatial mode basis change can be tailored so as to allow the manipulation of one spatial degree of freedom of the electromagnetic field by another. In the last chapter, we present how we modeled the transmission properties of the MPLC system. We use tools from random matrix theory to derive models for the distribution of the transmission coefficients of the MPLC and compare our results to other systems.

Chapter 1

Modes of the electromagnetic field

"It's a dangerous business, Frodo, going out your door. You step into the Road, and if you don't keep your feet, there is no knowing where you might be swept off to."

– Bilbo Baggins

Contents

The aim of this chapter is to introduce the concepts and notations necessary to describe classical electromagnetic fields. We begin with the description of modes of the electromagnetic field and linger on spatial modes, which are essential to the understanding of the remainder of this thesis. We then describe the different techniques which allow for the manipulation such spatial modes. Finally, we present the multi-plane light conversion (MPLC) technology exploited by CAILabs — the study of which is the heart of this work.

1.1 Modes of the electromagnetic field

The description of the electromagnetic field uses the concept of modes which we detail in this section (see [Fabre 11] for a reference manual on the subject).

1.1.1 Maxwell's equations

Maxwell's equations encompass the wave-like behaviour of light. We consider a real electric field $\mathbf{E}(\boldsymbol{r}, t)$, and magnetic field $\mathbf{B}(\boldsymbol{r}, t)$[1]. The *frequency* Fourier decomposition of the electric field is expressed as:

$$\mathbf{E}(\boldsymbol{r}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widetilde{\mathbf{E}}(\boldsymbol{r}, \omega)\, e^{-i\omega t} d\omega. \tag{1.1}$$

The complex field is composed of the sum of positive frequency $\mathbf{E}^{(+)}$ and negative frequency $\mathbf{E}^{(-)}$ components of the real field. The positive frequency component is defined as:

$$\mathbf{E}^{(+)}(\boldsymbol{r}, t) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \widetilde{\mathbf{E}}(\boldsymbol{r}, \omega)\, e^{-i\omega t} d\omega. \tag{1.2}$$

Both components carry the same information and are linked by the following equation:

$$\mathbf{E}(\boldsymbol{r}, t) = \mathbf{E}^{(+)}(\boldsymbol{r}, t) + \mathbf{E}^{(+)*}(\boldsymbol{r}, t) = \mathbf{E}^{(+)}(\boldsymbol{r}, t) + \mathbf{E}^{(-)}(\boldsymbol{r}, t). \tag{1.3}$$

In vacuum, and in the absence of currents and charges, the complex field follows Maxwell's equations:

$$\begin{aligned} \nabla \cdot \mathbf{E}^{(+)} &= 0 & \nabla \times \mathbf{E}^{(+)} &= -\frac{\partial \mathbf{B}^{(+)}}{\partial t} \\ \nabla \cdot \mathbf{B}^{(+)} &= 0 & \nabla \times \mathbf{B}^{(+)} &= \mu_0 \epsilon_0 \frac{\partial \mathbf{E}^{(+)}}{\partial t} \end{aligned} \tag{1.4}$$

[1] Bold characters designate vectors.

where ϵ_0 is the permittivity of free space and μ_0 the permeability of free space. They are linked to the speed of light in vacuum c by the relation $c = 1/\sqrt{\epsilon_0 \mu_0}$.
Maxwell's equations can be combined to find the wave equation:

$$\nabla^2 \mathbf{E}^{(+)} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}^{(+)}}{\partial t^2}. \tag{1.5}$$

1.1.2 Traveling plane waves

We make the assumption that the system we describe lies in a finite volume. We then define another finite volume V, of cubic shape and side length L, which encompasses the former. We define the *spatial* Fourier components of the complex field as:

$$\mathscr{E}_{\boldsymbol{n}}(t) = \frac{1}{L^3} \int_V d^3\boldsymbol{r}\, \mathbf{E}^{(+)}(\boldsymbol{r}, t) e^{-i\boldsymbol{k}_n \cdot \boldsymbol{r}} \tag{1.6}$$

where the wave vector $\boldsymbol{k}_{\boldsymbol{n}}$ is defined as:

$$\boldsymbol{k}_{\boldsymbol{n}} \cdot \mathbf{x} = n_x \frac{2\pi}{L}; \qquad \boldsymbol{k}_{\boldsymbol{n}} \cdot \mathbf{y} = n_y \frac{2\pi}{L}; \qquad \boldsymbol{k}_{\boldsymbol{n}} \cdot \mathbf{z} = n_z \frac{2\pi}{L}. \tag{1.7}$$

with $(n_x, n_y, n_z) \in \mathbb{Z}^3$. $\boldsymbol{k}_{\boldsymbol{n}}$ in turns define the angular frequency $\omega_{\boldsymbol{n}} = c\,|\boldsymbol{k}_{\boldsymbol{n}}|$. Maxwell's equations impose:

$$\boldsymbol{k}_{\boldsymbol{n}} \cdot \mathscr{E}_{\boldsymbol{n}}(t) = 0. \tag{1.8}$$

which means the spatial Fourier components of the electromagnetic field are thus necessarily orthogonal to the wave vector in Fourier space. They are called *transverse fields* and lie in a two dimensional space. One can choose a real basis $\{\boldsymbol{\epsilon}_{\boldsymbol{n},s}\}_{s\in 1,2}$ of this space, and the projection of the field on this basis gives the polarization components of the field[2]. We can finally rewrite, using a common index $i = \{\boldsymbol{n}, s\}$ the expression of the field:

$$\mathbf{E}^{(+)}(\boldsymbol{r}, t) = \sum_i \mathscr{E}_i(t)\, \boldsymbol{\epsilon}_i e^{i\boldsymbol{k}_n \cdot \boldsymbol{r}}. \tag{1.9}$$

In the formalism we derived, normalized solutions to the wave equation can be written as:

$$\boldsymbol{u}_i(\boldsymbol{r}, t) = \boldsymbol{\epsilon}_i e^{i(\boldsymbol{k}_{\boldsymbol{n}} \cdot \boldsymbol{r} - \omega_{\boldsymbol{n}} t)} \tag{1.10}$$

[2]The vectors $\epsilon_{\boldsymbol{n},s}$ can be chosen to be complex — in which case they need to be orthogonal with respect to the Hermitian product.

with $i = \{\boldsymbol{n}, s\}$. The norm is defined using the following scalar product:

$$\frac{1}{L^3}\int d^3\boldsymbol{r}\, \boldsymbol{u}_i \boldsymbol{u}_{i'}^*. \tag{1.11}$$

Hence, the electric (and magnetic) field can be described using the scalar components $\mathscr{E}_i$ on the mode basis of the *polarized traveling plane waves*

$$\mathbf{E}^{(+)}(\boldsymbol{r}, t) = \sum_i \mathscr{E}_i(t)\, \boldsymbol{u}_i(\boldsymbol{r}, t)\,. \tag{1.12}$$

1.1.3 Mode basis change

The decomposition of equation 1.9 is not the only possible one. One can use any ensemble of vector functions $\{\boldsymbol{u}_i(\boldsymbol{r})\}$ which satisfy the following equations:

$$\frac{1}{L^3}\int_V d^3\boldsymbol{r}\, \boldsymbol{u}_i^*(\boldsymbol{r}, t) \cdot \boldsymbol{u}_j(\boldsymbol{r}, t) = \delta_{i,j} \tag{1.13}$$

$$\sum_i \boldsymbol{u}_i^*(\boldsymbol{r}, t) \cdot \boldsymbol{u}_i(\boldsymbol{r}', t) = \delta(\boldsymbol{r} - \boldsymbol{r}')\,. \tag{1.14}$$

Let $\{\boldsymbol{w}_i(\boldsymbol{r}, t)\}$ be another suitable mode basis. We can define the unitary matrix U as:

$$U_{i,j} = \frac{1}{L^3}\int_V d^3\boldsymbol{r}\, \boldsymbol{u}_i^*(\boldsymbol{r}, t)\, \boldsymbol{w}_j(\boldsymbol{r}, t) \tag{1.15}$$

which allows to switch from one mode basis to another

$$\boldsymbol{w}_j(\boldsymbol{r}, t) = \sum_i U_{i,j}\, \boldsymbol{u}_i(\boldsymbol{r}, t)\,. \tag{1.16}$$

This mode basis change matrix will be studied in more details in section 1.3.

1.1.4 Temporal and spectral modes

The decomposition presented above is called the normal mode decomposition. Each mode consists of an independent monochromatic polarized wave. This representation, while convenient for instance in the case of a monochromatic beam of light, is not very well suited to pulses of light for instance. If we consider a field of a single polarisation where the spectrum is centered around an angular frequency ω_0 and small compared to this frequency ($\Delta\omega \ll \omega_0$), we can write

$$\mathbf{E}^{(+)}(\boldsymbol{r}, t) = \sum_i \mathscr{E}_i \boldsymbol{u}_i(t - r_i/c)\, e^{i(\boldsymbol{k_i}\cdot\boldsymbol{r} - \omega_0 t)} \tag{1.17}$$

where $\{u_i(t)\}$ is a set of orthonormal temporal modes, whose time derivatives varies slowly with respect to ω_0 and r_i is the projection on $\boldsymbol{r}$ on the unit vector defined by $\boldsymbol{k}_i$. The same reasoning can be applied to the Fourier expression of the field:

$$\mathbf{E}^{(+)}(\boldsymbol{r},\omega) = \sum_i \mathscr{E}_i \boldsymbol{u}_i(\omega-\omega_0)\, e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}} \tag{1.18}$$

where $\{\boldsymbol{u}_i(\omega)\}$ are referred to as spectral modes. This concept of temporal modes, while central in the study of ultra-short laser pulses, will not be developed further in this thesis.

1.1.5 Spatial modes

The concept of spatial modes is the central object of this thesis. This concept proves itself very useful in the study of a number of light beams, in particular laser beams. Indeed, in the plane normal to their direction of propagation, laser beams have a finite extent. As an example, beams propagating inside optical fibers are, by design, localized to the physical dimensions of the fiber. The goal of this section is to derive the expression of spatial modes bases $\{u_i\}$ which may describe a monochrome linearly polarized field.

We will use in the remainder of this thesis, except on a few occasions, the *paraxial approximation*. This approximation is applied when considering beams which are composed of plane waves with wave vectors close to the propagation axis of the beam — the angle between the wave vector and the direction of propagation should be small. Using the expression of the field we derived in 1.12, in the case of a monochromatic, linearly polarized field, we can rewrite the wave equation 1.5 as the *Helmholtz* equation — a more complete derivation of this equation can be found in [Siegman 86]:

$$\left[\nabla^2 + k^2\right] E(x,y,z) = 0. \tag{1.19}$$

By definition, plane waves are solutions of this equation. We choose z as the propagation axis and look for solutions which have the following form:

$$E(x,y,z) = u(x,y,z)\, e^{-ikz}. \tag{1.20}$$

The $\exp(-ikz)$ components oscillates rapidly along z with a period equal to the wavelength of the light λ ($k = 2\pi/\lambda$). By inserting this expression in 1.19, and factorizing the $\exp(-ikz)$ term, we can rewrite the wave equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - 2ik\frac{\partial u}{\partial z} = 0. \tag{1.21}$$

The paraxial approximation states that the transverse spatial profile of the beam evolves *slowly* along z compared to one wavelength, but also compared to the variations caused by the finite transverse size of the beam. It can be expressed as follows:

$$\left|\frac{\partial^2 u}{\partial z^2}\right| \ll \left|2k\frac{\partial u}{\partial z}\right| \quad \text{and} \quad \left|\frac{\partial^2 u}{\partial z^2}\right| \ll \left|\frac{\partial^2 u}{\partial x^2}\right| \quad \text{and} \quad \left|\frac{\partial^2 u}{\partial z^2}\right| \ll \left|\frac{\partial^2 u}{\partial y^2}\right|. \tag{1.22}$$

We rewrite equation 1.21 using this approximation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - 2ik\frac{\partial u}{\partial z} = 0. \tag{1.23}$$

We finally derive the paraxial wave equation

$$\nabla_T^2 u - 2ik\frac{\partial u}{\partial z} = 0. \tag{1.24}$$

$\nabla_T^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator in the transverse plane. The paraxial approximation is commonly accepted to be valid for beams whose plane wave components propagate with angles up to 0.5 rad ($\sim 29°$) [Siegman 86]. In the paraxial approximation, we can write the electric field in the following form:

$$\mathbf{E}^{(+)} = e^{i(k_0 z - \omega_0 t)} \sum_i \mathscr{E}_i \tilde{u}_i. \tag{1.25}$$

Modes $\tilde{u}_j$ can be separated in transverse (u_i) and longitudinal (v_k) parts:

$$\tilde{u}_j = u_i v_k \tag{1.26}$$

The transverse part $u_i(x, y, z)$ obeys

$$\frac{1}{L^2}\iint dxdy\, u_i u_j^* = \delta_{ij} \quad \forall z \in \mathbb{R} \tag{1.27}$$

while the longitudinal part $v_k(t - z/c) = v_k(\tau)$ obeys:

$$\frac{c}{L}\int d\tau\, v_k v_l^* = \delta_{kl} \tag{1.28}$$

Gaussian beam One exact solution to the scalar wave equation is a spherical wave:

$$E(x,y,z) = \frac{\exp\left(-ik\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}\right)}{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}}. \tag{1.29}$$

In the context of the paraxial approximation, we take the situation in which the spherical wave is considered on a transverse plane which is far from z_0 along the z axis, and for coordinates x, y which are also not too far from the axis ($|x-x_0| \ll |z-z_0|$ and $|y-y_0| \ll |z-z_0|$). The Fresnel approximation to diffraction theory states that we can make the following approximations:

$$\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2} \simeq z - z_0 + \frac{(x-x_0)^2+(y-y_0)^2}{2(z-z_0)} \tag{1.30}$$

$$\frac{1}{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}} \simeq \frac{1}{z-z_0} \tag{1.31}$$

In order to simplify expressions and without losing generality, we set $x_0 = 0$, $y_0 = 0$. Using the Fresnel approximation and 1.29, we derive the expression of the *paraxial* spherical wave $E(x,y,z) = u(x,y,z)e^{-ik(z-z_0)}$ with:

$$u(x,y,z) = \frac{1}{z-z_0}\exp\left(-ik\frac{x^2+y^2}{2(z-z_0)}\right). \tag{1.32}$$

This expression is not satisfactory as such, since the "beam" is not localized along the propagation axis. We set $z_0 = 0$ and introduce a complex term iq_0 with $q_0 \in \mathbb{R}$: we replace the $z - z_0$ term by the function $q(z) = z + iq_0$. This function is called the complex radius of curvature. This allows us to define the transverse spatial profile of a gaussian spherical beam — which is still a solution of the paraxial wave equation (in the paraxial and Fresnel approximations):

$$u(x,y,z) = \frac{1}{q(z)}\exp\left(-ik\frac{x^2+y^2}{2q(z)}\right). \tag{1.33}$$

Using the complex radius of curvature, we define the following quantities:

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i\frac{\lambda}{\pi w^2(z)}, \tag{1.34}$$

$$\tan\psi(z) = \frac{\pi w^2(z)}{\lambda R(z)} = \frac{z}{z_R}, \tag{1.35}$$

$$z_R = \frac{\pi w_0^2}{\lambda}. \tag{1.36}$$

Equation 1.34 for $z = 0$ gives $q_0 = z_R$. We can rewrite equation 1.33 as:

$$u(x, y, z) = \sqrt{\frac{2}{\pi}} \frac{1}{w(z)} \exp\left(-\frac{x^2 + y^2}{w^2(z)}\right) \exp\left(-ik\frac{x^2 + y^2}{2R(z)} + i\psi(z)\right) \tag{1.37}$$

which is the general expression of a *Gaussian* beam. Using this expression, we can give physical meanings to the quantities introduced above. $R(z)$ is the radius of curvature of the phase of the beam and $w(z)$ the waist size of the beam. We can rewrite these quantities as functions of the parameters w_0 and $z - R$:

$$w(z) = w_0 \sqrt{1 + \left(\frac{z}{z_R}\right)^2} \tag{1.38}$$

$$R(z) = z + \frac{z_R^2}{z}. \tag{1.39}$$

z_R is called the *Rayleigh length*: z_R gives the propagation distance after which the waist size has increased by a factor of $\sqrt{2}$. This quantity will be used extensively in chapter 3. $\psi(z)$ is called the *Gouy* phase. The Gaussian beam is a very important tool: it correctly describes the electric field in many optical systems. It is nonetheless not the only solution to the paraxial wave equation.

Hermite-Gaussian modes If one looks at solutions of the paraxial wave equation which are separable in the cartesian coordinate system ($u_{nm}(x, y, z) = u_n(x, z) \times u_m(y, z)$), one derives the following decoupled differential equations:

$$\frac{\partial^2 u_n}{\partial x^2} - 2ik\frac{\partial u_n}{\partial z} = 0 \tag{1.40}$$

$$\frac{\partial^2 u_m}{\partial y^2} - 2ik\frac{\partial u_m}{\partial z} = 0 \tag{1.41}$$

Solution to these equations is the Hermite-Gauss modes family which calls upon the Hermite polynomials. These functions are solution to the following differential equation: $\frac{d^2 f}{dx^2} - 2x\frac{df}{dx} = -2nf$ with $n \in \mathbb{N}$. The first Hermite polynomials read as

follow:

$$H_0(x) = 1 \tag{1.42}$$
$$H_1(x) = 2x \tag{1.43}$$
$$H_2(x) = 4x^2 - 2 \tag{1.44}$$
$$H_3(x) = 8x^3 - 12x \tag{1.45}$$
$$H_4(x) = 16x^4 - 48x + 12 \tag{1.46}$$

These polynomials are orthogonal with respect to the weight function e^{-x^2}:

$$\int_{-\infty}^{+\infty} H_m(x) H_n(x) e^{-x^2} dx = \sqrt{\pi} 2^n n! \delta_{nm} \tag{1.47}$$

The one-dimensional expression of a Hermite-Gaussian transverse profile is

$$u_n(x,z) = \left(\frac{2}{\pi}\right)^{1/4} \left(\frac{1}{2^n n! w_0}\right)^{1/2} \left(\frac{q(0)}{q(z)}\right)^{1/2} \left[\frac{q^*(0)}{q(0)} \frac{q^*(z)}{q(z)}\right]^{n/2} H_n\left(\frac{\sqrt{2}x}{w(z)}\right) \exp\left(-i\frac{kx^2}{2q(z)}\right). \tag{1.48}$$

Using the expressions of the Gouy phase and waist size derived in the previous section, we can write

$$\frac{q(0)}{q(z)} = \frac{w_0}{w(z)} \exp\left(i\psi(z)\right) \tag{1.49}$$

$$\left[\frac{q^*(0)}{q(0)} \frac{q^*(z)}{q(z)}\right]^{n/2} = \exp\left(in\psi(z)\right) \tag{1.50}$$

A Hermite-Gaussian transverse profile of index n in the x direction and index m in the y direction can be expressed as

$$u_{nm}(x,y,z) = \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{1}{2^{n+m} n! m! w^2(z)}\right)^{1/2} H_n\left(\frac{\sqrt{2}x}{w(z)}\right) H_m\left(\frac{\sqrt{2}y}{w(z)}\right) \times \exp\left[-\frac{\left(x^2+y^2\right)}{w^2(z)}\right] \exp\left[-ik\frac{\left(x^2+y^2\right)}{2R(z)} + i\left(n+m+1\right)\psi(z)\right] \tag{1.51}$$

and will be referred to as HG_{nm} modes. Their spatial extent is equal to that of the Gaussian mode up to a $\sqrt{n+m+1}$ proportionality factor.

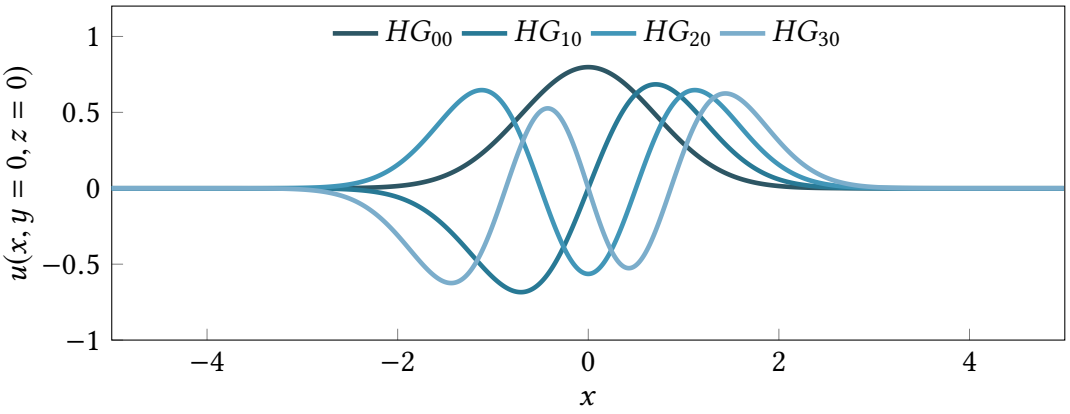

Figure 1.1: Amplitude of the first four Hermite-Gaussian modes.

Laguerre-Gaussian modes The same derivation can be conducted in the case of separable functions in the cylindrical coordinate system. The solution found defines the family of Laguerre-Gaussian transverse profiles and uses the generalized Laguerre polynomials, which read as follows:

$$L_0^m(x) = 1 \tag{1.52}$$

$$L_1^m(x) = -x + m + 1 \tag{1.53}$$

$$L_2^m(x) = \frac{x^2}{2} - (m+2)x + \frac{(m+2)(m+1)}{2} \tag{1.54}$$

$$L_3^m(x) = \frac{-x^3}{6} + \frac{(m+3)x^2}{2} - \frac{(m+2)(m+3)x}{2} + \frac{(m+1)(m+2)(m+3)}{6} \tag{1.55}$$

$$L_{p+1}^m(x) = \frac{(2p+1+m-x)\,L_p^m(x) - (p+m)\,L_{p-1}^m(x)}{p+1} \tag{1.56}$$

A Laguerre-Gaussian beam of radial index $p \in \mathbb{N}$ and of azimuthal index $m \in \mathbb{Z}$ can be expressed as

$$u_{pm}(r,\theta,z) = \sqrt{\frac{2p!}{(1+\delta_{0m})\,\pi\,(m+p)!}}\,\frac{1}{w(z)}\left(\frac{\sqrt{2}r}{w(z)}\right)^{|m|} L_p^{|m|}\left(\frac{2r^2}{w^2(z)}\right)\times$$
$$\exp\left[-\frac{r^2}{w^2(z)}\right]\exp\left[-ik\frac{r^2}{2R(z)} + i\,(2p+m+1)\,\psi(z) + im\theta\right] \tag{1.57}$$

One particularly interesting feature of Laguerre-Gaussian beams is the presence of the $\exp[im\theta]$ term. Indeed, for $m \neq 0$, the mode possesses an orbital momentum along z: its phase describes helices. We will refer to them as LG_p^m modes.

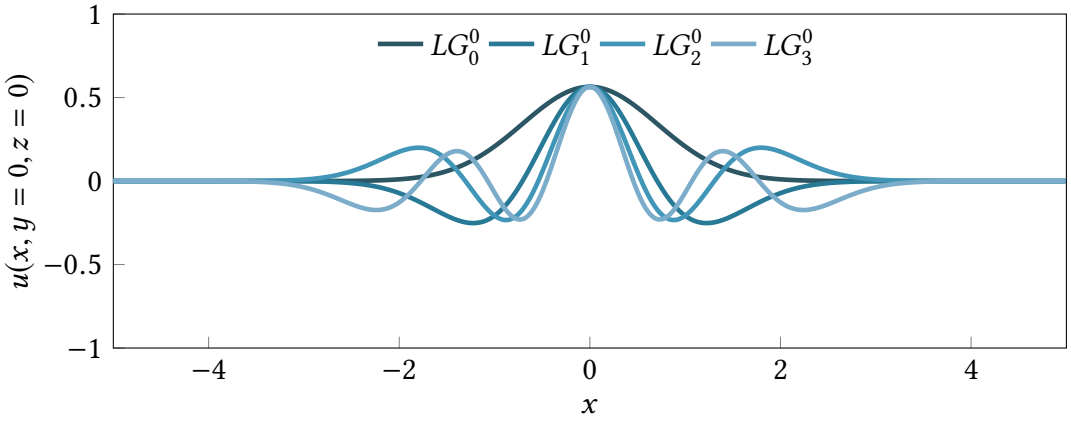

Figure 1.2: Amplitude of the first four Laguerre-Gaussian modes.

Other types of free-space modes The types of modes which can exist is as large as the number of types of solutions to the wave equation given the constraints of the medium the light propagates in. In free space, an interesting type of modes we wish to dwell upon is that of non-diffracting beams. Non diffracting beams were first introduced in [Durnin 87a, Durnin 87b], using general solutions of the wave equation derived in [Whittaker 03]. Fields of the type

$$E(x, y, z, t) = \exp[i(\beta z - \omega t)] \int_0^{2\pi} A(\phi) \exp[i\alpha(x \cos\phi + y \sin\phi)] d\phi \tag{1.58}$$

(with $\beta^2 + \alpha^2 = (\omega/c)^2$ and $A(\phi)$ any continuous complex function of ϕ) are solutions of the wave equation. If $\beta > 0$, the time-average intensity at (x, y) is constant: $\forall z \in \mathbb{R}$, $I(x, y, z = 0) = I(x, y, z)$. This property means these solutions do not diffract. As an example, if one takes $A(\phi) = 1/(2\pi)$, we find

$$\begin{aligned} E(x, y, z, t) &= \exp[i(\beta z - \omega t)] \int_0^{2\pi} \exp[i\alpha(x \cos\phi + y \sin\phi)] \frac{d\phi}{2\pi} \\ &= \exp[i(\beta z - \omega t)] J_0\left[\alpha\sqrt{x^2 + y^2}\right] \end{aligned} \tag{1.59}$$

where J_0 is the zero-order Bessel function of the first kind. We define an ideal Bessel beam in cylindrical coordinates (r, θ, z) as

$$E(r, \theta, z, t) = \exp[i(\beta z - \omega t)] J_n[\alpha r] \exp(in\theta) \tag{1.60}$$

with J_n the nth order Bessel function. This "non-diffracting" property of such modes unfortunately renders them unphysical: indeed, the energy of such modes is infinite

[McGloin 05]. However, finite-energy realization of Bessel beams can take the form of Bessel-Gauss beam — Bessel beams with a Gaussian envelope [Gori 87], which are discussed in appendix A.

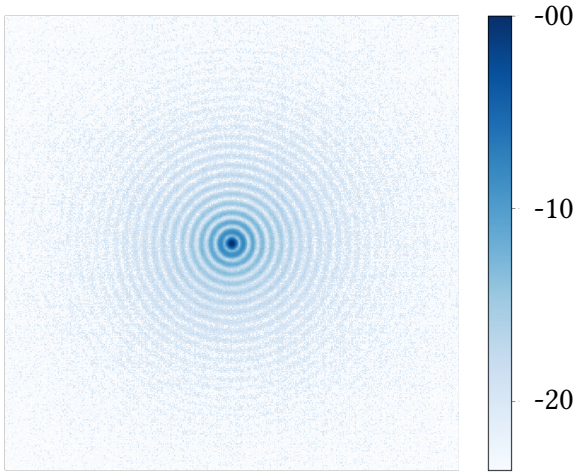

Figure 1.3: Experimental measurement (dB scale) of the intensity of a Bessel-Gauss beam — the image was taken during the preparatory works of A.

Non-diffracting beams also contain the families of Airy and Mathieu beams, which both carry the name of the equation their envelope is a solution of [Mazilu 10, Berry 79, Siviloglou 07, Gutiérrez-Vega 00].

Prolate spheroidal wave functions Although we do not aim at making a complete zoology of free-space spatial modes, we finally mention the circular prolate spheroidal wave functions (PSWF) which are used in the context of hard-aperture imaging since they possess the property of being an orthonormal basis of bandwidth-limited functions [Landau 61, Slepian 61]. These functions arise in two different contexts: as solutions of a second order Sturm-Liouville equation and as the spectrum of an integral operator. A good review of these functions is [Wang 17] — which we make a brief summary of in this paragraph. First, we consider the subclass of functions $f(t) \in L^2(\mathbb{R})$ whose Fourier transform $\tilde{f}(\omega)$ is zero for $\omega > W$ (W is the bandwidth parameter). Such a function is called W-bandlimited. The space of all W-bandlimited functions is known as the Paley-Wiener space. The time-frequency problem which was addressed by Slepian *et al.* [Slepian 83] can be stated as follow:

to what extent can the energy of a function belonging to the Paley-Wiener space be concentrated in a finite interval (which means there exists a value T such that $f(t) = 0$ for $|t| > T/2$)? This question amounts to studying the Fredholm integral equation of the second kind: find $\mu \in \mathbb{R}$ and function ϕ such that

$$B\phi(x) = \int_{-1}^{+1} \frac{\sin c(x-t)}{\pi(x-t)} \phi(t)\, dt = \mu\phi(x) \text{ with } c > 0 \tag{1.61}$$

where $x \in [-1, 1]$ and $c = \pi WT$. Spectral theory indicates that the operator B admits a countable set of eigenvalues and eigenfunctions $\{\mu_n, \phi_n\}_{n\in\mathbb{N}}$:

$$\int_{-1}^{+1} \frac{\sin c(x-t)}{\pi(x-t)} \phi_n(t;c)\, dt = \mu_n(c)\phi_n(x;c) \tag{1.62}$$

The eigenvalues are strictly positive and form a decreasing series $\mu_0(c) > \mu_1(c) > \ldots > 0$. The eigenfunctions are orthogonal and complete in $L^2([-1,1])$. The function $\phi_n(x;c)$ are the PSWF of degree n and bandwidth c. If we define the operator F as

$$F_c\phi(x;c) = \int_{-1}^{+1} e^{icxt} \phi(t;c)\, dt \tag{1.63}$$

one can verify that

$$B\phi(x;c) = \frac{2\pi}{c} \left(F_c^* \circ F_c\right) \phi(x;c) \tag{1.64}$$

This equation implies that B and F_c have the same eigenfunctions $\{\phi_n\}$ and eigenvalues $\{i^n\lambda_n(c)\}$ with the relation

$$\lambda_n(c) = \sqrt{\frac{2\pi}{c}\mu_n(c)} \tag{1.65}$$

The integral operators F_c commute with the second-order differential operator defined by:

$$D_x^c = -\frac{d}{dx}\left(1-x^2\right)\frac{d}{dx} + c^2x^2 \tag{1.66}$$

which corresponds to the Sturm-Liouville equation. Indeed:

$$D_x^c \int_{-1}^{+1} \frac{\sin c(x-t)}{\pi(x-t)} dt = \int_{-1}^{+1} \frac{\sin c(x-t)}{\pi(x-t)} dt. \tag{1.67}$$

The PSWF $\{\phi_n\}$ are thus also eigenfunctions of the Sturm-Liouville problem

$$D_x^c\phi_n(x;c) = \chi_n\phi_n(x;c) \text{ for } x \in [-1,1] \tag{1.68}$$

This problem arises from solving the Helmholtz equation in prolate spheroidal coordinates by separation of variables. It is a special case of the spheroidal wave equation with $m = 0$ [Wang 17]:

$$\frac{d}{dx}\left[\left(1-x^2\right)\frac{d\phi}{dx}\right] - \left[\frac{m^2}{1-x^2} + c^2x^2\right]\phi = -\lambda\phi \text{ with } x \in [-1,1] \tag{1.69}$$

The functions $\{\phi_n\}$ we introduced are thus the PSWF of order 0. The essential property of the PSWF we are interested in is the fact that they are eigenfunctions of an operator (F_c) which is related to the finite Fourier transform. They form an orthogonal basis of the Paley-Wiener space of c-bandlimited functions and can be used as a basis for the decomposition of bandlimited functions. In the field of optics, these functions are particularly useful in imaging. Virtually all imaging systems are aperture-limited: the bandwidth of the collected signal is limited by the extent of the optical instruments of the system, which takes the form of a circular pupil. We will develop the use of the PSWF in this context in chapter 2.

Waveguide modes Free-space propagation is only one of the many propagation regimes in which one might be interested when looking at light propagation. One other particularly interesting regime is that of light propagation in waveguides, such as optical fibers to only name one, where we are interested in modes which are bound to the waveguide. The model we choose to present here is that of the step-index fibers — which allows for solutions to be derived analytically — although the reasoning is similar for graded index fibers, or other types of index profiles. Most types of waveguides and waveguide modes are presented in detail in [Snyder 83]. The transverse profile of an ideal step-index fiber is composed of two parts: a core of radius r_{core} and refractive index n_{co} and a cladding of theoretically infinite extent and refractive index n_{cl}, with $n_{co} > n_{cl}$. The refractive index is defined by the phase velocity v of the light in the unguided medium: $n = c/v$ (c is the speed of light in vacuum) and is thus a function of the wavelength λ or equivalently the angular frequency ω. We also assume the fiber to be infinite in the longitudinal direction ($\boldsymbol{z}$ is the unit vector of this axis) in order to have translational invariance in that direction:

$$\mathbf{E}_j(x,y,z) = \boldsymbol{e}_j(x,y)\exp\left(i\beta_j z\right) = \left[\boldsymbol{e}_{tj}(x,y) + e_{zj}(x,y)\,\boldsymbol{z}\right]\exp\left(i\beta_j z\right) \tag{1.70}$$

where we separate the transverse and longitudinal components of the field, and β_j is the *propagation constant* of mode j.

We look at modes which do not attenuate as they propagate ($\beta_j \in \mathbb{R}$) and which are bounded — which translates into $\beta_j > 0$. When taking the temporal dependence

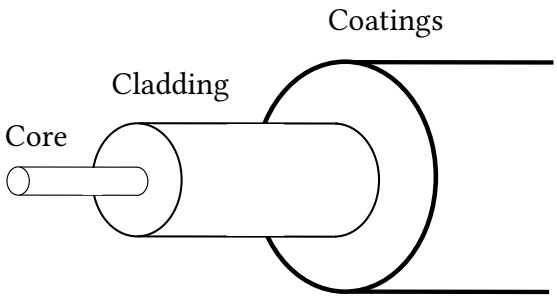

Figure 1.4: Sketch of the structure of an optical fibre — not to scale. The coating layers protect the physical fiber but play no role in the confinement of light in the fiber.

into account, the field varies as $\exp\left(i\beta_j z - i\omega t\right)$. The *phase velocity* of a mode is given by $v_{pj} = \omega/\beta_j$. Another useful notion is that of *group velocity*. Under certain conditions (see [Smith 70] for a detailed account), for instance in absence of absorption, the group velocity of a mode gives the speed at which the power in a given mode is transmitted, and it is given by $v_{gj} = d\omega/d\beta_j$. When replacing the expression of 1.70 in the source-free vector wave equation, one can derive [Snyder 83]:

$$\left(\nabla_T^2 + n^2\frac{\omega^2}{c^2} - \beta_j^2\right)\boldsymbol{e}_j(x,y) = -\left(\nabla_T + i\beta_i \boldsymbol{z}\right)\left(\boldsymbol{e}_{tj}(x,y)\cdot\nabla_T \ln n^2\right). \tag{1.71}$$

In a step-index fiber, except at the interface between core and cladding, we have $\nabla_T \ln n^2 = 0$ which means that

$$\left(\nabla_T^2 + n^2\frac{\omega^2}{c^2} - \beta_j^2\right)\boldsymbol{e}_j(x,y) = 0 \tag{1.72}$$

everywhere, except at the interface. The requirement that the fields should (i) have a finite norm everywhere, (ii) be bounded to the waveguide and (iii) satisfy the boundary condition of the core-cladding interface (continuity as well as continuity of the first derivative at the interface) leads to an eigenvalue equation for the modal propagation constant β_j — whose derivation is beyond the scope of this thesis. The fiber parameter $V = kr_{core}\sqrt{n_{co}^2 - n_{cl}^2}$ governs the number of solutions to this eigenvalue equation and their value. In a simple step-index fiber, when $0 < V < 2.405$, the fiber operates in the single mode regime, meaning only one spatial mode is guided by this fiber. Other types of optical fibers include *few-mode fibers* where a small number of modes are supported, which have non-degenerate group velocities, and

multi-mode fibers where a large number of modes can be supported, with degeneracies of the group velocities. One last element worth mentioning in the derivation of modes of waveguides is that its cross-sectional geometry/refractive-index profile induce specific polarization properties[3].

1.1.6 Typically available spatial modes

The spatial modal description of an electromagnetic field lies primarily in the type of source where it comes from. Thermal light emitted by a light bulb cannot be described by a single spatial mode, but rather by an infinite sum over a mode basis — for instance, the plane wave mode basis. We will focus in this thesis almost entirely on laser light. Laser light is — in a majority of cases — produced by fields in cavities [Grynberg 10]. The geometry of these cavities imposes constraints on the spatial modes of the field for it to be resonant. The properties of the cavity will also determine whether the laser is *single-mode* or *multi-mode*, that is if it can be described as a single spatial mode, or as a statistical superposition of several spatial modes. The same analysis holds true for temporal modes. For usual laser cavities, the output of the cavity is a Gaussian beam. By playing with the cavity parameters, Hermite-Gaussian beams, as well as Laguerre-Gaussian beams, can be produced. Fibered lasers produce in most cases beams which are composed of the eigenmodes of the fiber used.

1.2 Shaping the light

Different spatial modes have different properties that might be interesting depending on the desired application. It is thus of great interest to be able to shape a laser beam in order to change its spatial properties. Several lossless techniques exist to carry out this operation, which we will present here. An even larger number of beam shaping techniques involving loss exist — among which are slides and intensity masks. We will however not detail the existing lossy techniques in this section. The shaping techniques are *mode-independent* in the sense that all the modes of

[3] Indeed, we mentioned previously that the solutions to Maxwell's equation should be continuous at the index profile boundaries. A useful image to keep in mind is that the phase acquired by a plane wave upon reflection on a flat surface depends on the polarization. In circular fibers, transverse electric and transverse magnetic (TE and TM) modes are supported. In the case of more complex index profiles, the $\nabla_T \ln n^2$ term in 1.71 induces hybrid modes which are referred to as EH and HE modes.

which the shaped beam is composed will "see" the same transformation[4]. Mode-dependent shaping will be introduced in the following section.

1.2.1 Conventional geometrical optics

Conventional geometrical optics are actually a very versatile and convenient tool to change the spatial mode of a beam. Geometrical optics consists of the use of mirrors, lenses, prisms, beam splitters, diaphragms and such combined with free-space propagation. Flat mirrors allow to change the propagation direction of a beam, curved mirrors can collimate or focus a beam. Lenses used in $f - f$ configuration implement a spatial Fourier transform of the field. Telescopes allow to change the waist size of a beam. Iris diaphragms and blades filter out a fraction of the beam. Lenses can also introduce aberrations, which usually represent a hindrance in imaging systems. These aberrations are in fact undesired shaping of the optical beam which can in certain cases prove useful. Cavities and interferometers are in most cases composed exclusively of geometrical optic elements.

A noteworthy example of the large shaping possibilities offered by a smart optical design is the conversion of Hermite-Gaussian modes into Laguerre-Gaussian modes and vice-versa using two spherical lenses and two cylindrical lenses with perpendicular principal planes [Beijersbergen 93, Alekseev 98]. This transformation uses the following equality:

$$LG_{min\{n,m\}}^{|m-n|} = \sum_{k=0}^{N} i^k b(n,m,k) HG_{N-k,k} \tag{1.73}$$

with

$$b(n,m,k) = \left(\frac{(N-k)!k!}{2^N n!m!}\right)^{1/2} \frac{1}{k!}\frac{d^k}{dt^k}\left[(1-t)^n (1+t)^m\right]|_{t=0}. \tag{1.74}$$

The conversion of the beams uses the fact that the Gouy phases acquired by the beams is not the same in both direction (x and y). The distance between lenses is chosen such that the phase difference at the exit from the lens at the time when the wave fronts overlap will be $\pi/2$ for a conversion from LG_{nm} to HG_{kl} modes.

[4]This remark holds true for the case for *co-propagating* modes — modes which are already separated spatially can be shaped independently.

1.2.2 Phase masks

The other main type of spatial shaping tool is phase masks. The physical implementation of a phase mask is called a phase plate. After transmission or reflection on phase mask, the output electric field ($\mathbf{E}_{out}$) can be expressed as:

$$\mathbf{E}_{out}(x, y, z, t) = e^{ip(x,y)} \times \mathbf{E}_{in}(x, y, z, t) \tag{1.75}$$

where p is the phase mask pattern and $\mathbf{E}_{in}$ the incident field. This operation can, for instance, be completed using a piece of glass of varying thickness: the glass having a propagation index different than that of air, different parts of the beam will accumulate different phases after having propagated in the phase plate. If the phase plate is thin enough, refraction effects can be neglected, allowing the phase-plate to be approximated as having a pure phase effect. Another type of plates — which use birefringence effects as well as phase difference engineering — can be used to introduce an angular orbital momentum in an optical beam and are referred to as *q-plates*. Beams with circular polarization have an angular momentum called *spin angular momentum*, while beams with helical wavefronts (such as LG_p^m with $m \neq 0$) carry orbital angular momentum. *q-plates* are generally liquid crystal cells, which display a pattern called a topological charge. These phase plates allow to perform an exact conversion of the spin angular momentum to the orbital angular momentum [Marrucci 06, Marrucci 13].
Another such example is that of glass *axicons* [McLeod 54]. An axicon is a conical piece of glass with the revolution axis of the cone aligned with the direction of propagation of the beam. A plane wave propagating in an axicon will produce a Bessel beam: in A we use reflective axicons — and not transmissive axicons — to produce and characterize Bessel beams with record conical angles. It can be noted that since glass axicons rely on refraction, they do not need to be thin.

1.2.3 Reconfigurable phase masks

Another type of phase masks are *reconfigurable* phase masks, such as spatial light modulators (SLM) [Efron 94, Rosales-Guzmán 17]. SLM can work in transmission or reflection — and require to work with only one polarization of light. The key element of a SLM is a thin layer of, for instance, nematic liquid crystals. This liquid is composed of molecules which orientation can be controlled by an electric field. Their orientation in the liquid will, in turn, modify the optical index of the liquid. By using arrays of independently controlled liquid crystal pixels, one can at will and in a dynamical way control the phase pattern imposed on a beam. SLM are the most

polyvalent tool to produce a phase mask because of their very good resolution and reconfigurability. This technology has had an enormous impact in optics allowing the emergence of many new fields such as optical manipulation [Grier 03], pulse shaping [Weiner 00], microscopy [Maurer 10] and many more. However, they suffer from a number of drawbacks. First of all, their price remains quite high. Another drawback is the fact that in most technologies, the pixels need to be "refreshed" which means the phase pattern they display is not rigorously constant in time. This phenomenon is referred to as phase flickering. The presence of liquid-crystal cells implies that SLM are not compatible with high powers — whether in the pulsed or continuous regimes. The diffraction patterns displayed on the SLM often have more than one order of diffraction which introduces losses. Finally, the fact that they are only compatible with one linear polarization reduces their efficiency.

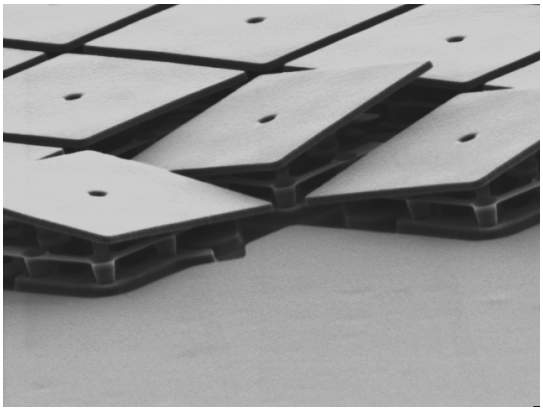

Figure 1.5: Photograph of a DMD. Image taken from [Mignardi 16].

Deformable surfaces are the second type of reconfigurable phase masks. Two main types exist: deformable mirrors and digital micro-mirror devices (DMDs). Deformable mirrors consist of a layer of pixel actuators which is covered by a deformable reflective membrane. DMDs are arrays of small pixel mirrors, which orientation can be very quickly switched between two positions — in practice, the light reaching each pixel can be steered in one of two directions, towards the image one wants to display or towards a dump. Each pixel can thus be defined as "on" or "off". These two technologies have allowed the emergence of the field of adaptive optics, allowing for instance for tremendous progress in astronomical imaging [Beckers 93, Roddier 99, Tyson 10]. Adaptive optics is a technology which allows the improvement of the resolution of imaging systems, by compensating for the turbulence of the medium the light goes through before reaching its observation plane. These

turbulences can be caused by atmospheric distortions in the case of astronomy or free-space communications. In living tissue imaging, these turbulences will originate from the movements of the tissues that light goes through. The adaptive surface will dynamically correct the aberrations caused by the imaging medium by compensating them.

Figure 1.6: Images of Neptune captured by ESO Very Large Telescope with and without the use of adaptive optics. In this implementation, four lasers are shot from the ground to the region of the sky which is to be imaged. The excitation they cause in the atmosphere is used to measure the amount of turbulence in the air. This information is then used to modulate the surface of the ground telescope mirror for it to compensate the aberrations caused by the atmosphere. On the corrected image, one can distinguish much sharper details of the surface of Neptune than on the image obtained without adaptive optics. (18/07/2018 — ESO/P. Weilbacher (AIP)).

1.2.4 Multiplexing techniques

Multiplexing is a method by which several signals are sent through the same medium, using the same physical resource. In optics, this consists in bringing several independent optical modes together so that they become co-propagating. Several degrees of freedom can be exploited to implement multiplexing: wavelength [Keiser 99], time-domain [Kani 06] and polarization [Hill 92] multiplexing are for instance mature technologies. We will present in this section spatial multiplexing [Berdagué 82, Murshid 08, Richardson 13]. The assessment of a multiplexing technique uses the

following figures of merit: the number of modes, the coupling efficiency, the mode selectivity as well as the cross-talk between modes.

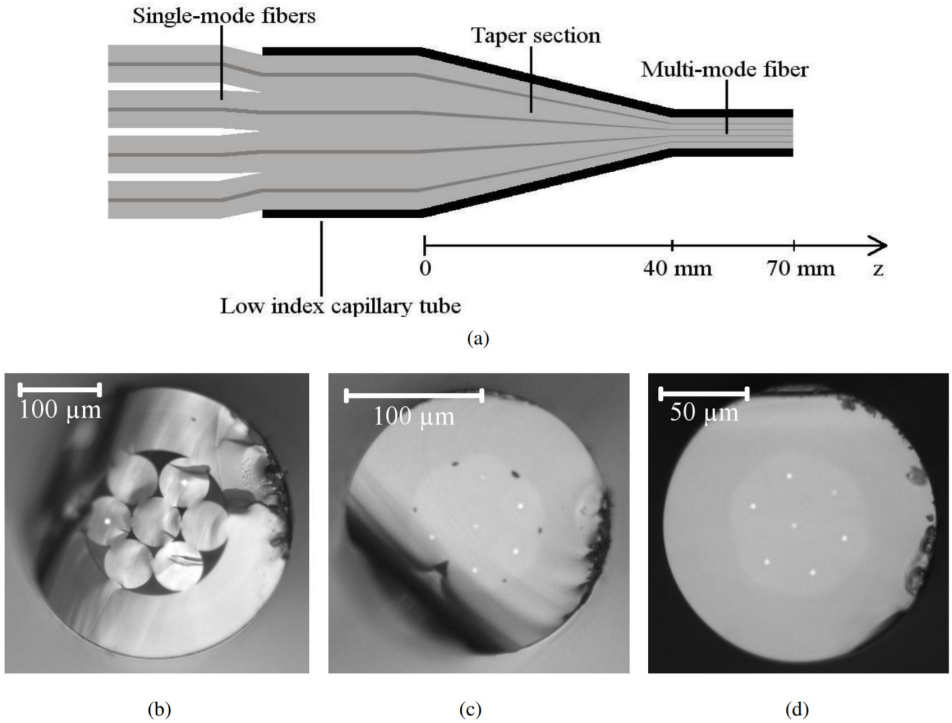

Figure 1.7: Schematic illustration of a photonic lantern. (b)-(d) Microscope images of a fiber bundle cross section at different positions in the taper transition. Image taken from [Noordegraaf 09].

Photonic lanterns are another type of technology which allows the multiplexing of modes in the context of fibered optics [Richardson 13]. Schematically, a photonic lantern is an assembly of the cores of single mode fibers which are fused together to form a multi-mode fiber core. Two main types of photonic lanterns exist. The first technique amounts to the tapering of a preform consisting of a bundle of fibers placed inside a low-index capillary. The fibers arrangement in the preform as well as their type (step-index or graded-index, hollow-core, single or multi-mode fiber, etc.) leads to a variety of performances for these types of lanterns [Velázquez-Benítez 18]. Nonmode selective photonic lanterns can be fabricated using identical single mode fibers. These lanterns require the use of multiple-input multiple-output (MIMO) methods to unscramble the information at the output of

the demultiplexer [Ryf 14, Guan 15]. Mode selectivity can be achieved by using dissimilar fibers and/or using specific geometrical arrangements of fibers in the preform such that each single-mode fiber excites only one spatial mode. When this fusing process respects adiabatic conditions, the modes of the different single-mode fibers can be shaped independently into the different eigenmodes of the multi-mode fiber [Leon-Saval 05, Birks 15]. Fiber-based photonic lanterns offer the advantage of being all-fibered devices which can be directly spliced to transmission fibers. The second type of photonic lantern fabrication uses ultrafast laser inscription which writes 3D waveguides in a bulk of glass [Guan 15]. This fabrication technique can be fully automated which ensures accurate layout control.

To be exhaustive on the question of spatial multiplexing technologies, the first implementation of multiplexing operations was made using binary phase plates and beam splitters [Ryf 12]. We also mention mode couplers [Gross 15], Y-junction couplers [Driscoll 13], micro-ring mode couplers [Luo 14] and fibered mode-selective couplers [Chang 14]. One of the important drawbacks of these approaches is their limited bandwidth of operation as well as the fact that they cannot easily be scaled to many modes.

1.3 The Multi-Plane Light Conversion Technology

The Multi-Plane Light Conversion (MPLC) technology was invented in the Laboratoire Kastler-Brossel and is now developed and commercialized by the company CAILabs. This thesis revolves around the study of its properties as well as the development of novel uses cases. All the results developed here are derived in [Morizur 10, Morizur 11b, Labroille 14] in greater details.

1.3.1 Back to mode basis change

We defined in 1.1.3 the matrix U which gives the mode basis change from $\boldsymbol{u}_i(\boldsymbol{r}, t)$ to $\boldsymbol{w}_i(\boldsymbol{r}, t)$. Since we are here interested in spatial modes only, we will often drop the t dependency and make the $\boldsymbol{r}$ dependency implicit. We can translate this in more conventional mathematical concepts by considering that, in free space, $\mathbf{E}$ lies in an infinite dimensional space. We rewrite equation 1.12 as

$$\mathbf{E}^{(+)}(\boldsymbol{r}, t) = \sum_i \mathscr{E}_i \, \boldsymbol{u}_i(\boldsymbol{r}, t) \,. \tag{1.76}$$

Because $\{\boldsymbol{u}_i\}$ is a complete orthonormal mode basis, $\mathbf{E}$ can uniquely be described by the vector $\boldsymbol{e} = \{\mathscr{E}_i\}$. If we refer to W as the matrix for the mode basis change

from $\boldsymbol{w}_i(\boldsymbol{r},t)$ to $\boldsymbol{u}_i(\boldsymbol{r},t)$ we can write:

$$W_{ij} = \int d^3\boldsymbol{r}\,\boldsymbol{w}_i^*\boldsymbol{u}_j = \left(\int d^3\boldsymbol{r}\,\boldsymbol{w}_i\boldsymbol{u}_j^*\right)^* = \left(\int d^3\boldsymbol{r}\,\boldsymbol{u}_j^*\boldsymbol{w}_i\right)^* = U_{ji}^* \tag{1.77}$$

that is $W = U^\dagger$. Furthermore, we can deduce from this derivation that $U^\dagger U = I$: U is a unitary matrix.

Any element of an optical setup can also be modeled as a matrix M: with a fixed choice of transverse mode basis, the electric field before and after this optical element will be related by the relation $\boldsymbol{e}_{out} = M\boldsymbol{e}_{in}$. If the optical element introduces no losses, this matrix M is unitary as well. Hence, any linear optical element implements a mode basis change. However, the set of the unitary transforms that can be implemented with lossless conventional optics (mirrors, lenses, beam-splitters) is only a subset of the unitary group[5].

1.3.2 Exemples of matrix representation

We will here describe two specific types of unitary transform: the Fourier transform and the transformation implemented by a phase mask. In specific mode bases, these transformation can be written in a simple way. We introduce first the pixel mode basis — it allows to easily model a mode in the usual cartesian coordinate system. The use of this mode basis only makes sense if the fields we want to describe have finite spatial transverse extensions. We make the assumption that the field $E(x,y,z)$ we want to describe has a spatial extent which is smaller than L in both directions at $z = 0$. We then choose N_{pixel} such that the field has no spatial feature smaller than the length $\Delta_{pixel} = L/N_{pixel}$. Using these definitions, we define a pixel basis $S_{i,j}$ with $(i,j) \in [\![1, N_{pixel}]\!]$ of size $\Delta_{pixel} \times \Delta_{pixel}$. The center of pixel $S_{i,j}$ has coordinates (x_i, y_i). The representation of the field in this basis is defined in the following way

$$\begin{aligned} &E_{i,j}(x,y,z=0) = E(x_i,y_i,z=0) && \text{if } (x,y) \in S_{i,j} \\ &E_{i,j}(x,y,z=0) = 0 && \text{if } (x,y) \notin S_{i,j} \end{aligned} \tag{1.78}$$

Hence, $E(x,y,z=0)$ is correctly described in the pixel mode basis by the vector $\mathbf{E}_k$ with $k \in [\![1, N_{pixel}]\!]^2$. In such a basis, the matrix M_{PP} describing the action of a

[5] If we exclude cylindrical lenses from this ensemble, this subset for instance preserves cylindrical symmetry, which means it is not equal to the unitary group.

phase plate on a field can be written as a diagonal matrix:

$$M_{PP} = \begin{pmatrix} e^{i\phi_1} & 0 & \cdots \\ 0 & e^{i\phi_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{1.79}$$

As indicated in 1.2.3, a phase plate can be physically implemented using a SLM. The number and the size of the pixels will be fixed by that of the SLM.

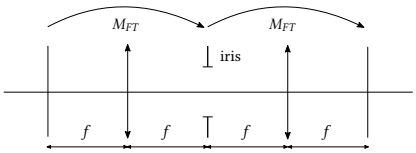

Figure 1.8: Implementation of $f-f$ and $4-f$ transforms and definition of M_{FT}.

While the pixel mode basis in convenient to express the transformation induced by a phase mask, the Hermite-Gauss mode basis is also an interesting mode basis to describe certain optical transformations, in particular the optical Fourier transform. Indeed, the Hermite-Gauss mode basis is, like other mode bases and up to a scaling constant, stable by Fourier transform. A spatial Fourier transform can be implemented using a lens of focal length f in a $f-f$ configuration (see figure 1.8) [Lohmann 93]. In this case, the relation between the input and output waists of a Hermite-Gauss mode basis is $w_{out} = (2\pi\lambda f)/w_{in}$. In the input Hermite-Gauss mode basis $HG_{mn,in}$ defined by w_{in} and output mode basis $HG_{mn,out}$ defined by w_{out}, we have $M_{FT}HG_{mn,in} = i^{m+n}HG_{mn,out}$, where M_{FT} designated the Fourier transform matrix. We note that four Fourier transforms in a row make up for the identity:

$$(M_{FT})^4 = I \tag{1.80}$$

1.3.3 Implementing any unitary matrix

Let U_{PP} be the group of all "phase-plate" matrices and U_{FT} that of the Fourier transform matrices in the pixel basis. It is shown in [Borevich 81] that any unitary matrix M can be constructed by a finite product of elements of U_{PP} and U_{FT}. This means that any mode basis change (unitary matrix) can be realized by a *finite* number of phase plates (M_{PP}) and spatial Fourier transforms (M_{FT}). This result lies at the heart of the working principle of the MPLC [Morizur 11b]. As an illustration, we define the matrix $T_{ij}(\theta)$ as the matrix representing the action of a beamsplitter between

pixels i and j in the pixel mode basis. Using a permutation in the mode basis it is always possible to write $T_{ij}(\theta)$ as:

$$T_{ij}(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 & 0 & \cdots \\ -\sin(\theta) & \cos(\theta) & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & & & \ddots \end{pmatrix} \tag{1.81}$$

A finite sequence of a maximum of 17 successions of M_{PP} and M_{FT} is sufficient to build any $T_{ij}(\theta)$ [Borevich 81].

1.3.4 Physical implementation of the MPLC

The previous section introduces a result which states the existence of a finite series of elements of M_{PP} and M_{FT} that can implement any given mode unitary transform. One can easily see that a physical implementation of such a series is in practice ill-suited, due to the high number of matrices/optical operations to be performed.

The problem the MPLC addresses is slightly different from the problem we addressed theoretically. The MPLC aims at performing unitary transforms on a *limited* number of modes: we only aim at shaping deterministically a subset of the modes of a complete mode basis. This problem has a smaller complexity than the one which was presented in the previous part. Instead of developing theoretical solutions for every specific transform one could want to perform, numerical optimization routines have been derived to derive a more "reasonable" number of operations for a given unitary transform.

We note here that while the derivation we presented in the previous section makes use of a perfect Fourier transform, we will throughout this thesis use free-space propagation. Free-space propagation, on a distance large enough to cover the near-field to far-field range, implements a Fourier transform. This result can be derived using Fraunhofer approximation (see [Goodman 96]) or alternatively using the fractional Fourier transform [Mendlovic 93, Ozaktas 93, Lohmann 93]. In the case of a Gaussian beam, such a distance can estimated using the characteristic spatial extent of a mode w_0 as $z_R(w_0) = \pi w_0^2/\lambda$.

To go further, one needs to first state the characteristics of the mode basis change one wishes to implement. This includes

- the number N_m of modes one wishes to shape simultaneously
- the wavelength λ of the fields

- the input mode basis $\{\boldsymbol{u}_i\}_{i\in[1,N_m]}$ with its characteristic spatial extent $w_{0,i}$
- the output mode basis $\{\boldsymbol{w}_i\}_{i\in[1,N_m]}$ with its characteristic spatial extent $w_{0,o}$.

One also need to specify the physical characteristics and limitations of the system we wish to build:

- the number of phase plates used in the system N_p
- the size of the pixels composing the phase plate dx and dy
- the number of said pixels n_x and n_y
- the distance between each phase plate $2L$
- the regularity constraints that the choice of material or fabrication technique imposes on the phase functions $\phi_j(x, y)$, $j \in [1, N_p]$ of the phase plates.

As stated above, a rule of thumb for the choice of the propagation distance is the following: $2L \geq \left(\pi \max\{w_{0,i}, w_{0,o}\}^2\right)/\lambda$.

An actual MPLC system is physically implemented as a succession of phase plates placed next to each other between which light bounces off of a plane mirror placed parallel to the phase plates. The distance between the phase plates and the mirror is given by L so that the total free-space propagation the beam takes place on a $2L$ distance. Figure 2.9 shows how a MPLC is physically implemented. The phase plates are stacked next to each other on a larger phase plate. This phase plate faces a plane mirror. The angle introduced by these bounces is taken into account in the optimization routines but is small enough so that the paraxial approximation is still valid. Finally, figure 1 shows a photograph of such a physical implementation.

Physical implementations of MPLC have first been developed by CAILabs to perform spatial multiplexing operations: several spatially separated independant monomode beams are "combined" — meaning they share the same propagation axis — using a MPLC. The operation performed by the MPLC is a spatial mode basis change between a basis of displaced gaussian beams to the basis defined by the eigenmodes of a multi-mode fiber. This operation allows different signals carried by independant light beams to travel along the same optical fiber.

Other unitary matrix implementations In a Mach-Zehnder interferometer used with only one beam as an input, the amount of light in each of the output arms can be continuously modulated from 0 to 100 % by adjusting the length of one of the two

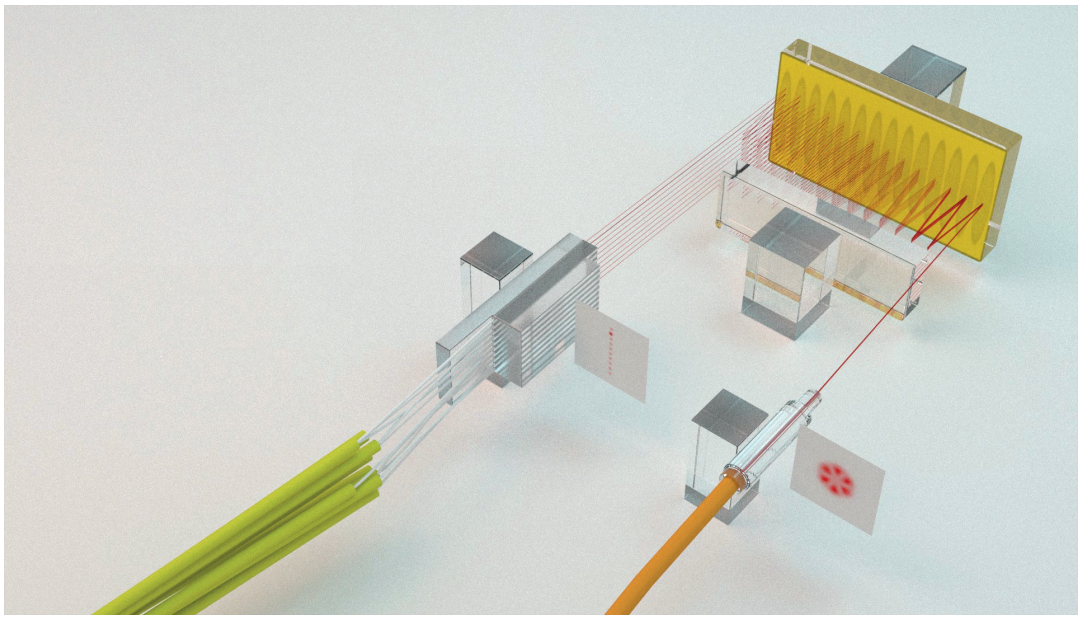

Figure 1.9: Schematic representation of a MPLC system. The unitary transformation which is presented on this image is a mode basis change between spatially separated modes which are outputs of single-mode fibers (in green) and co-propagating modes of a multi-mode fiber (in orange). The phase plate is represented as the golden surface at the back of the picture, while the plane mirror lies in front of it. The modes sizes are matched between the fibers and the MPLC system using a vertical array of micro-lenses for the single-mode fibers and a single micro-lens for the multi-mode fiber. Next to the micro-lenses, the spatial distributions of the modes are plotted. The transformation can be made in both directions: if one injects light into one of the single mode fibers, the output mode will be one of the multi-mode fiber. Inversely, if only one mode of the multi-mode fiber is excited, only the corresponding single mode fiber will receive energy. ©CAILabs.

arms — or equivalently the phase of the light in one arm with respect to the other arm. Since both arms are spatially separated, this implements a simple multiplexing operation. Such interferometers can be cascaded, allowing for more complex multiplexing implementations (in [Reck 94], authors demonstrate that such simple interferometers can be used as the "building blocks" of any unitary matrix). However, a free-space implementation of such an array is not quite scalable. On the contrary, on-chip waveguide arrays of such interferometer meshes produce promising results [Clements 16, Annoni 17].

1.4 Conclusion

In this chapter, the formalism necessary to the description of spatial modes of the electromagnetic field is introduced. Spatial mode bases are a central object throughout this manuscript and we described the different existing techniques which allow to perform mode basis changes on spatial modes of the electromagnetic field. Finally, the MPLC system was introduced.

Chapter 2

Improving optical resolution using modal analysis

"Pourtant s'ils savaient ! Là est le pays des merveilles, le vrai pays des fées: le réel."

– François Cavanna

Contents

Imaging systems often have the goal of providing the observer with more information than what can be obtained with the naked eye. More information can mean more detail or more accuracy. Its function might also be to record the obtained image for later use. The link between the object, that is the information we wish to acquire, and the image, the information we do acquire, is given not only by the physical properties of the imaging system, but also by the properties of the light which is used to perform the imaging operation. Very good introductions to this field can be found in [Mandel 95, Goodman 00]. Most of the derivations of the first part of this chapter are taken from both these references. We will first establish the formalism describing optical imaging and give the classical resolution limits which apply in the case of the imaging of two incoherent point sources. We will look in a second part at the different techniques which were developed to overcome these limitations. We will then make an introduction to parameter estimation theory and using this formalism, derive the limitations in the estimation of the displacement of a point source. Using the tools of estimation theory we show theoretically and experimentally that one can design spatial demultiplexers which are optimal tools

for the measure of the displacement of one single point source. We also show that the problem of the distance measurement between two point sources and the displacement measurement of a unique source can be described using the same tools. Using a MPLC system, we make the experimental demonstration of the measurement of both the displacement of one source and the distance between two incoherent point sources by performing intensity measurements on the different modes of the signal. Finally, we take a step back from parameter estimation by introducing the field of compressed sensing. This discipline opens the path to new types of imaging systems through optimization and careful sensing mode basis design — the experimental realization of which could be made using MPLC systems.

2.1 Optical resolution

Electromagnetic fields have fluctuations associated to them. The statistical properties of light, which we have not discussed in the previous chapter, play an important role in the outcome of optical experiments. The simplest phenomenon illustrating these fluctuations is the interference effect: when two fields which are *coherent* with one another are combined, one can observe interference fringes. On the contrary, if the two fields' fluctuations are independent, no interference pattern is observed. The field studying these effects is called *optical coherence theory* since these interference phenomena depend on the correlations that exist between the light fluctuations of both beams. The problem we are interested in can be schematically described as follows: an electric field illuminates an object and the resulting light is transmitted through an imaging system. What we have access to is the intensity pattern in the *image* plane. From this image, we aim at extracting as much information as possible on the physical system, the *object*. The imaging system, but also the very properties of the light- and in particular its coherence properties — will determine the intensity pattern we measure, and thus the information we can access.

2.1.1 Mathematics of coherence and imaging

In this section, we introduce the formalism which is necessary for the description of the fluctuations of the electric field — regardless of their classical or quantum origin. This formalism of image formation in partially coherent light is necessary to the characterization of the intensity distribution found in the image plane for any experimental situation. It also allows to identify the individual effects of the

imaging system, the object and the illuminating field. We take $E(\boldsymbol{r}, t)$ the positive frequency component of the field, which we defined in subsection 1.1.1 (we drop the (+) superscript to simplify notations). $E(\boldsymbol{r}, t)$ fluctuates in time. We define the cross-correlation function of the field as

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2; t_1, t_2) = \langle E^*(\boldsymbol{r}_1, t_1) E(\boldsymbol{r}_2, t_2) \rangle. \tag{2.1}$$

The average is taken over different realizations of E: it is a statistical ensemble average. Since we are in a stationary regime, the origin of time plays no role, which means we can replace the expression of the cross-correlation function by

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2; \tau) = \langle E^*(\boldsymbol{r}_1, t) E(\boldsymbol{r}_2, t + \tau) \rangle. \tag{2.2}$$

This function is known as the *mutual coherence function*. The intensity at point $\boldsymbol{r}$ is given by $I(\boldsymbol{r}, t) = E^*(\boldsymbol{r}, t) E(\boldsymbol{r}, t)$. Then, the average intensity at point $\boldsymbol{r}$ is:

$$\langle I(\boldsymbol{r}, t) \rangle = \langle E^*(\boldsymbol{r}, t) E(\boldsymbol{r}, t) \rangle = \Gamma(\boldsymbol{r}, \boldsymbol{r}; 0). \tag{2.3}$$

Finally, we define the *mutual intensity* as

$$J(\boldsymbol{r}_1, \boldsymbol{r}_2) = \Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2; 0). \tag{2.4}$$

Coherence time The coherence time $\Delta\tau$ at point $\boldsymbol{r}$ is a measure of the time period during which correlations of the field can be observed:

$$(\Delta\tau)^2 = \frac{\int_{-\infty}^{\infty} \tau^2 |\Gamma(\tau)|^2 \, \mathrm{d}\tau}{\int_{-\infty}^{\infty} |\Gamma(\tau)|^2 \, \mathrm{d}\tau} \tag{2.5}$$

with $\Gamma(\tau) = \Gamma(\boldsymbol{r}, \boldsymbol{r}; \tau)$[1].

The situation we are interested in is described in figure 2.1. The electric field $E_{\text{source}}(\alpha, \beta; t)$ emitted by the light source is transmitted to the object plane through the illumination system. The light $E_o(x, y; t)$ which reaches the object plane goes through the object (which is characterized by the transmittance function $t_{\text{object}}(x, y)$) and leaves the object plane as $E'_o(x, y; t)$. Finally, the field $E_{\text{image}}(u, v; t)$ reaches the image plane after going through the imaging system. We aim at establishing a relation between the mutual intensities in the object and image planes. We first place ourselves in the paraxial approximation, with z as the propagation axis of light as well as the rotational symmetry axis of the imaging system. We will consider the

[1] We also have $\Gamma(-\tau) = \Gamma^*(\tau)$.

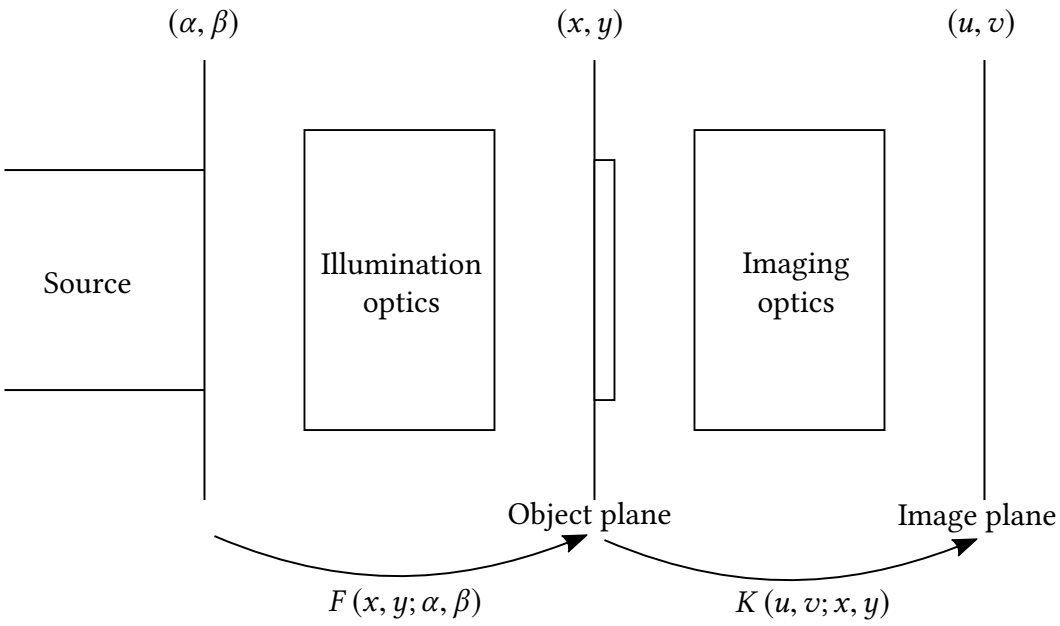

Figure 2.1: Illumination and imaging systems (from [Goodman 00]).

transverse profiles of the field with respect to the z axis. For this reason, we now write $J(x_1, y_1; x_2, y_2)$.

In order to describe imaging setups we introduce the *amplitude spread function* of a system $K(u, v; x, y)$. We assume that under the quasi monochromatic approximation, any imaging system can be described by such a function. The amplitude spread function allows to relate the mutual intensity leaving the object plane $J'_o(x_1, y_1; x_2, y_2)$ to the mutual intensity reaching the image plane $J_{\text{image}}(u_1, v_1; u_2, v_2)$ with the following equation:

$$J_{\text{image}}(u_1, v_1; u_2, v_2) = \iiiint_{-\infty}^{\infty} J'_o(x_1, y_1; x_2, y_2) K(u_1, v_1; x_1, y_1) \times K^*(u_2, v_2; x_2, y_2) \, dx_1 dy_1 dx_2 dy_2 \tag{2.6}$$

The intensity distribution in the image field can be calculated using the previous expression: $I_{\text{image}}(u, v) = J_{\text{image}}(u, v; u, v)$. The quantity $K(u_1, v_1; x_1, y_1) K^*(u_2, v_2; x_2, y_2)$ is the impulse response of the system — it allows to relate the mutual intensity at the image point $(u_1, v_1; u_2, v_2)$ to an object of mutual intensity consisting of an impulse at $(x_1, y_1; x_2, y_2)$. Under conditions

specified in detail in [Goodman 00][2], which amounts to the system being *space invariant*, equation 2.6 can be written as a convolution:

$$J_{\text{image}}(u_1, v_1; u_2, v_2) = \iiiint_{-\infty}^{\infty} J_o(x_1, y_1; x_2, y_2)\, K(u_1 - x_1, v_1 - y_1) \times K^*(u_2 - x_2, v_2 - y_2)\, dx_1 dy_1 dx_2 dy_2. \tag{2.7}$$

In this case, the amplitude spread function is now a function of only two variables. The *point spread function* is the normalized amplitude spread function $K(u, v)$. The formalism we have derived here is also valid to describe the illuminating system, that is the optical system which links the optical source to the object plane. The light emitted by the source at the point (α, β) can be represented by the field $E_{\text{source}}(\alpha, \beta; t)$. The illumination from this point which reaches the object on point (x, y) is expressed as

$$E_o(x, y; t) = \iint_{-\infty}^{\infty} F(x, y; \alpha, \beta)\, E_{\text{source}}(\alpha, \beta; t - \delta_1)\, d\alpha d\beta \tag{2.8}$$

where F is the amplitude spread function of the illuminating system and δ_1 is a delay which depends on (x, y) and (α, β)[3]. The illumination which leaves the object is $E'_o(x, y; t) = E_o(x, y; t) \times t_{\text{object}}(x, y)$ where $t_{\text{object}}(x, y)$ is the transmittance of the (thin) object — which we assume to be independent of the source point. t_{object} can be complex: the object can absorb light and add spatial dephasing.
As an example, we give the transmittance function $t_{\text{lens}}(x, y)$ of a thin lens, which is given by the following equation:

$$t_{\text{lens}}(x, y) = P(x, y) \exp\left[-i \frac{\pi}{\lambda f}\left(x^2 + y^2\right)\right] \tag{2.9}$$

where λ is the wavelength of the light, f is the focal length of the lens, and $P(x, y)$ is the aperture function of the lens ($P = 0$ outside the lens aperture).

[2]These conditions are non trivial and not always satisfied. The object coordinates and coordinate axis must be normalized such that the magnification factor between the coordinates (x, y) and (u, v) is unity. The coordinate axis must be oriented such that image inversion effects are removed. Finally, the amplitude spread function K must not contain any space-variant phase factors. This last condition is the most binding. As an example, a two lens telecentric system satisfies this condition. Systems where the phase factors of K and K^* cancel each other in the interest range of separation can also be considered space-invariant.

[3]In the same way, we can write $E_{\text{image}}(u, v; t) = \iint_{-\infty}^{\infty} K(u, v; x, y)\, E'_o\left(x, y; t - \delta'_1\right) dx dy$, with δ'_1 a delay which depends on (u, v) and (x, y).

We can express the mutual intensity reaching the object as

$$J_o(x_1, y_1; x_2, y_2) = \langle E_o(x_1, y_1; t - \delta_1) E_o^*(x_2, y_2; t - \delta_2) \rangle \tag{2.10}$$

in the monochromatic assumption where $|\delta_1 - \delta_2| \ll \Delta\tau$. The intensity at coordinates (u, v) in the image plane can finally be expressed as

$$I_{\text{image}}(u, v) = \iiiint_{-\infty}^{\infty} K(u - x_1, v - y_1) K^*(u - x_2, v - y_2) t_{\text{object}}(x_1, y_1) t^*(x_2, y_2) \times J_o(x_1, y_1; x_2, y_2) \, dx_1 dy_1 dx_2 dy_2. \tag{2.11}$$

This equation states that the knowledge of K, t_{object} and J_o allows the calculation of I_{image}.

In this section, no assumption was made on the coherence properties of the light source and the formalism we derived is valid for partially coherent sources. However, we will now focus on the simpler cases of totally coherent or incoherent illumination.

2.1.2 Incoherent case

In the case of total incoherence, each point of the source is considered to be totally independent of the other points. We can describe the mutual intensity reaching the object plane as

$$J_o(x_1, y_1; x_2, y_2) = I_0 \, \delta(x_2 - x_1, y_2 - y_1) \tag{2.12}$$

where we suppose the illumination to be spatially homogeneous, that is equal to I_0 at any point. By inserting this expression in equation 2.11, we find the following expression:

$$I_{\text{image}}(u, v) = I_0 \iint_{-\infty}^{\infty} |K(u - x, v - y)|^2 \left| t_{\text{object}}(x, y) \right|^2 dx dy. \tag{2.13}$$

In the case of incoherent homogeneous illumination, the image intensity is proportional to the convolution of the object intensity transmittance $|t|^2$ with the intensity spread function $|K|^2$. An imaging system with incoherent imaging is *linear in intensity*.

2.1.3 Coherent case

In the case of full coherence, the expression of the mutual intensity is

$$J_o(x_1, y_1; x_2, y_2) = I_0 \tag{2.14}$$

where we again make the assumption of spatial homogeneity. This equation is equivalent to assuming that the object is illuminated with a plane wave whose wave-vector is normal to the object. When inserting this equation in 2.11 we now find

$$I_{\text{image}}(u,v) = I_0 \left| \iint_{-\infty}^{\infty} K(u-x, v-y)\, t_{\text{object}}(x,y)\, dxdy \right|^2 \tag{2.15}$$

which can be rewritten in the following form (since $I = |E|^2$)

$$E_{\text{image}}(u,v) = \sqrt{I_0} \iint_{-\infty}^{\infty} K(u-x, v-y)\, t_{\text{object}}(x,y)\, dxdy. \tag{2.16}$$

An imaging system with fully coherent illumination is thus *linear in complex amplitude*.

Coherence of natural light sources To illustrate this notion of spatial coherence of light sources, we derive here the spatial coherence area of two natural light sources: the sun and a star (Betelgeuse). In an observation plane at a given distance from a light source, the fields considered at two different points $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are coherent if $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ both belong to the coherence area associated to this source. We designate by $\Delta\Omega$ the solid angle a source subtends at the point of observation. For the sun, this value is $\Delta\Omega_{sun} \simeq 6.81 \times 10^{-5}$ sr, while for the Betelgeuse star, it is $\Delta\Omega_B \simeq 4.15 \times 10^{-14}$ sr. The coherence area of a source is given by ([Mandel 95]):

$$\Delta A \sim \overline{\lambda}^2 / \Delta\Omega \tag{2.17}$$

where $\overline{\lambda}^2$ is the wavelength of observation. For $\overline{\lambda}^2 = 1.55\mu$m, $\Delta\Omega_{sun} \simeq 3.53 \times 10^{-2}$ mm^2 and $\Delta\Omega_B \simeq 57.9$ m^2.

2.1.4 Imaging two point sources

One of the two imaging problems which are at the heart of of this chapter consists in imaging two point sources. In the terms of the formalism introduced above, this simply writes

$$t(x,y) = a\,\delta\left(x - \frac{d}{2}, y\right) + b\,\delta\left(x + \frac{d}{2}, y\right) \tag{2.18}$$

where $(a,b) \in \mathbb{C}$ and we assume that the point sources lie on the x axis and are separated by a distance d. Using the expression of the image plane intensity in the

coherent case we can write

$$I_i(u,v) = I_0 \left| a\,K\left(u - \frac{d}{2}, v\right) + b\,K\left(u + \frac{d}{2}, v\right)\right|^2 . \tag{2.19}$$

For incoherent illumination, we have

$$I_i(u,v) = I_0 \left(|a|^2 \left|K\left(u - \frac{d}{2}, v\right)\right|^2 + |b|^2 \left|K\left(u + \frac{d}{2}, v\right)\right|^2 \right). \tag{2.20}$$

For further analysis, we need to have access to the amplitude spread function of the system — hence to define an imaging system. We here choose the simple and classic situation of a $4f$ imaging system (see figure 1.8). Two identical lenses of focal length f are placed at distances f of the object and image plane. They are separated by a distance $2f$. At a distance $2f$ of the object plane lies a pupil which is the aperture stop of the whole system. In the case of a simple circular pupil, we can give an analytical expression of K [Goodman 00]:

$$K(\rho) = \frac{2\pi R^2}{(\lambda f)^2} \left[\frac{J_1\left(\frac{2\pi R\rho}{\lambda f}\right)}{\frac{2\pi R\rho}{\lambda f}} \right] \tag{2.21}$$

where R is the radius of the pupil, ρ is the radial coordinate and J_1 is the first Bessel function of the first kind. The intensity pattern of a single point source is called an Airy pattern (see figure 2.2).

In the remaining of this thesis, we will only consider the case of incoherent illumination. While the expression of the Airy pattern is analytical, it is in most imaging contexts approximated by a Gaussian function since this form often allow simplifications in calculations.

The Airy pattern is responsible for the limits in resolution of an imaging system: indeed, a point source is imaged as a pattern with spatial extent — it is blurred. The size of the Airy pattern is a function of the parameters of the imaging system. It also gives the smallest size of detail which can be present in an image: it quantifies the resolution power of the imaging system.

2.1.5 Historical resolution criteria

The question of telling apart two distinct light sources is as old as optical imaging. Several criteria have been derived to estimate the smallest distance at which two

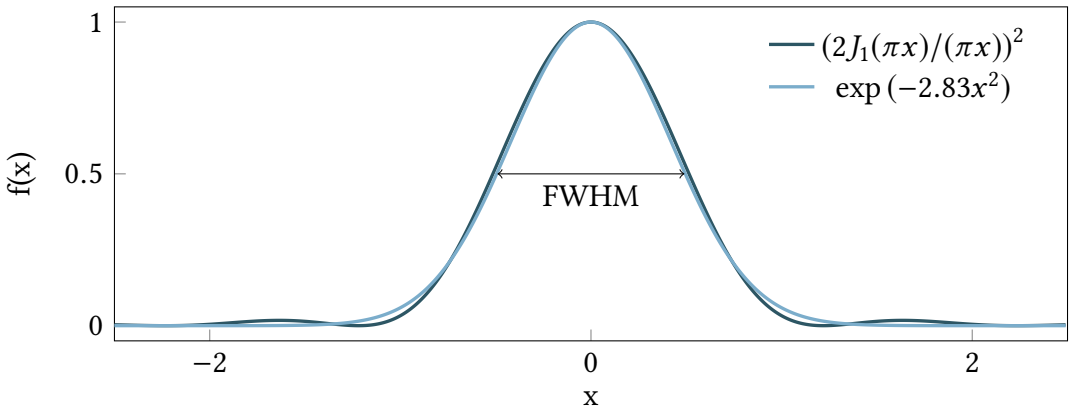

Figure 2.2: Airy pattern and its Gaussian approximation (FWHM $\simeq 0.99$).

points could be distinguished by their images, that is the smallest distance at which one can still obtain information on the actual distance between two light sources. We present here some of the most well-known resolution criteria.

The numerical aperture (NA) of an imaging system is dimensionless number which characterizes the range of angles (in a geometrical optics picture) over which a ray can enter or exit the system: $NA = n \sin i_{max}$ with n the index of refraction of the medium and i_{max} the half-angle of the cone of light that can enter or exit the system.

Abbe criterion Ernst Abbe [Abbe 73] devised the first resolution criterion. He stated that the imaging resolution was limited to half a wavelength normalized by the numerical aperture of the system, which amounts to

$$d_{Abbe} = 0.5 \frac{\lambda}{NA}. \tag{2.22}$$

Rayleigh criterion The most famous of these criteria is attributed to Lord Rayleigh [Lord Rayleigh 79]. It states that two points can be resolved if the first minimum of the Airy pattern formed by one source coincides with the maximum of the pattern formed by the second source:

$$d_{Rayleigh} \simeq 0.61 \frac{\lambda}{NA}. \tag{2.23}$$

Sparrow criterion In an astronomical observation context, another criterion was derived in [Sparrow 16] stating that as long as the derivative of the total intensity

around zero went negative, two sources could still be distinguished:

$$d_{Sparrow} = 0.47\frac{\lambda}{NA}. \tag{2.24}$$

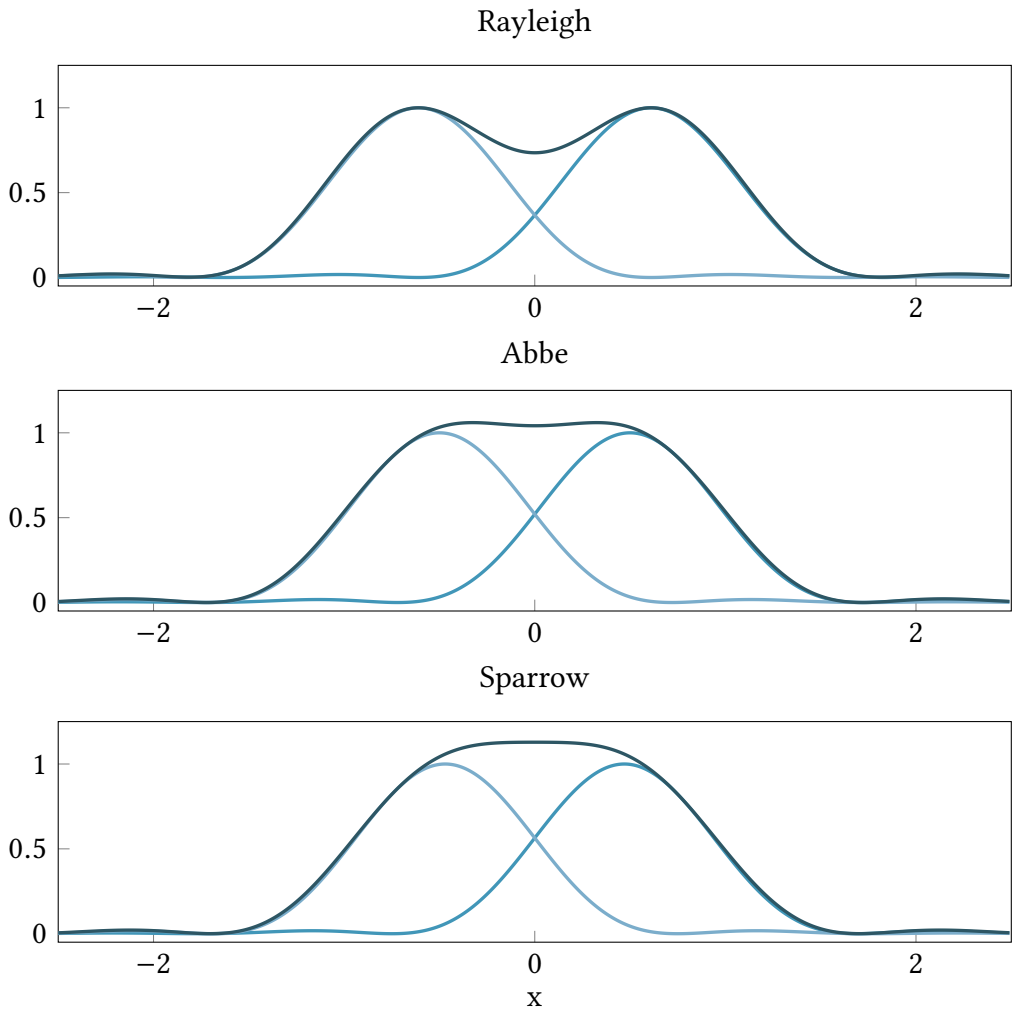

Figure 2.3: Illustration of different resolution criteria.

Generally speaking, imaging techniques which allow one to distinguish sources separated by a distance smaller than the Rayleigh criterion are called *super-resolution* techniques. To conclude this part, we wish to include here a quote from [Goodman 00] which we find particularly insightful:

> We note [...] that the question of when two closely spaced point sources are barely resolved is a complex one and lends itself to a variety of rather subjective answers. [...] In fact, the ability to resolve two point sources depends fundamentally on the signal-to-noise ratio associated with the detected image intensity pattern, and for this reason, criteria that do not take account of noise are subjective.

2.1.6 A different take on the problem

Compared to the approach used in the definition of the above resolution criteria, the problem of the smallest distinguishable detail in an image has been addressed in a slightly different way in [di Francia 69, Bertero 82b, Bertero 82a]. Instead of considering an image as being a collection of points, it is modeled as a continuous function. In section 2.1.1, we have derived that the image and the object are linked by a convolution. This convolution operation can be modeled as an operator, whose properties (singular values, etc..) can be linked to the imaging performances of the system. We give here the case of an incoherent illumination in one dimension with a rectangular aperture stop, details of which can be found in [Bertero 82a]. Derivation for the case of coherent illumination can be found in [Bertero 82b]. We define X as the spatial extent of the object and I'_o as the intensity transmittance function of the object. The bandwidth of the system is defined as $[-2\Omega, 2\Omega]$, which can be calculated from the parameters of the imaging system. We can define the operator A in one dimension with the following equation [Bertero 82a]:

$$I_{\text{image}}(x) = \left(AI'_o\right)(x) = \int_{-X/2}^{X/2} \frac{\sin^2[\Omega(x-y)]}{\pi\Omega(x-y)^2} I'_o(y)dy,\ |x| \leq X/2 \quad [4] \tag{2.26}$$

The sampling theorem states that the smallest useful sampling length is $R = \pi/(2\Omega)$. This is a rephrasing of the fact that details smaller than this length will not be resolved by the imaging system — which is coherent with the results we presented in the previous part.

[4] In the coherent case, the operator is defined by [Bertero 82b]:

$$\left(AI'_o\right)(x) = \int_{-X/2}^{X/2} \frac{\sin[\Omega(x-y)]}{\pi(x-y)} I'_o(y)\,dy,\ |x| \leq X/2. \tag{2.26}$$

2.2 Overcoming the classical resolution limitations in imaging

The resolution limits we presented in the previous part have been derived early on in the study of imaging. A huge body of work has been dedicated to overcoming these limitations. We present in this section some of the most well known and/or noteworthy works aiming at improving the resolving power of imaging systems — without claiming to establish an exhaustive list on the subject.

Analytical continuation We build on the formalism introduced in 2.1.6 and define the reduced variables $t = 2x/X$ and $c = X\Omega/2$. We rewrite 2.25 as

$$\left(AI'_o\right)(t) = \int_{-1}^{1} \frac{sin^2\left[c\left(t-s\right)\right]}{\pi c\left(t-s\right)^2} I'_o(X/2s)\frac{2ds}{X}, |t| \leq 1. \tag{2.27}$$

A is an operator, which is self-adjoint, non-negative, injective and compact in $L^2(-1,1)$ [Bertero 82a]. The trace of this operator is the Shannon number $S = X\Omega/\pi$. S is also the number of coefficients, according the sampling theorem, which is necessary to describe the image. In [di Francia 69], using information theory, the Shannon number is shown to be the number of degrees of freedom of the image.

As an operator, A admits eigenvalues λ_k — which are all strictly positive — and normalized eigenfunctions ϕ_k which form a basis in $L^2(-1,1)$. These functions are the spheroidal prolate functions we introduced in 1.1.5. When ordered in a decreasing sequence, the eigenvalues of A have the following property: for $k > 4c/\pi$, they go to zero exponentially fast. To simplify notations, we write $I'_o = f$ and $I_{\text{image}} = g$. We define $\bar{g} = A\bar{f}$ as the noise-free image in the region $|t| \leq 1$, we can derive $\bar{f}$ as

$$\bar{f}(t) = \sum_{k=0}^{+\infty} \frac{\bar{g}_k}{\lambda_k}\phi_k(t) \tag{2.28}$$

with

$$\bar{g}_k = \int_{-1}^{1} \bar{g}(t)\,\phi_k(t)\,dt. \tag{2.29}$$

This results is referred to as analytical continuation. It can be rephrased in the following way: if the object has a finite size, its Fourier transform is an analytical function.
Hence, the knowledge of the Fourier transform of the object over a finite interval is sufficient to reconstruct the function unambiguously. In an optical imaging

context, this means that if we were able to perfectly record the intensity profile (without any kind of noise), an analytical continuation of the image intensity profile combined with the knowledge of the imaging system properties would allow to reconstruct perfectly the object profile. This result is not only valid for optical imaging [di Francia 69] and was derived beforehand in the context of radio astronomy [Lo 61] and electronic information [Wolter 61]: indeed, an electronic signal transmitted over a communication channel of limited bandwidth can be modeled with the same mathematical tools. However, as we will develop in the next part, if this mathematical argument is indeed true, the notion of a noise-free image is of little use. Although classical noise may be almost cancelled in the best imaging systems, quantum noise is always present. For $k > S$, $\lambda_k \simeq 0$ and the information carried by ϕ_k is virtually lost. This analysis in developed in more details in [Kolobov 00]. Finally, in this formalism, another measure of the resolution power can be defined as the distance between consecutive zeros of the highest order spheroidal prolate function transmitted by the system.

Aperture synthesis We have seen in 2.1.4 that the image of a point is given by the following expression:

$$K(\rho) = \frac{2\pi R^2}{(\lambda f)^2}\left[\frac{J_1\left(\frac{2\pi R\rho}{\lambda f}\right)}{\frac{2\pi R\rho}{\lambda f}}\right]. \tag{2.30}$$

The size of this pattern is scaled by the size R of the limiting pupil or aperture stop. Aperture synthesis is a process used in interferometry: it combines the signals from several different telescopes, allowing to reach a resolution which is equivalent to that of a telescope which would have the size of the surface defined by all the telescopes (see 2.4). This technique was first developed in the context of radio astronomy by Martin Ryle and his coworkers[5].

Super-resolved fluorescence microscopy Another field which has greatly driven the quest of super-resolution is microscopy. We presented in part 2.1.5 the fact that the resolution limitation is proportional to the wavelength of the field used. Making use of the fact that the De Broglie wavelength of an electron is much smaller than the wavelength of visible light, electron microscopy has allowed for tremendous improvement, for instance in the imaging of surfaces. However, this technique

[5]Martin Ryle was awarded the physics Nobel prize in 1974 for his work along with Antony Hewish.

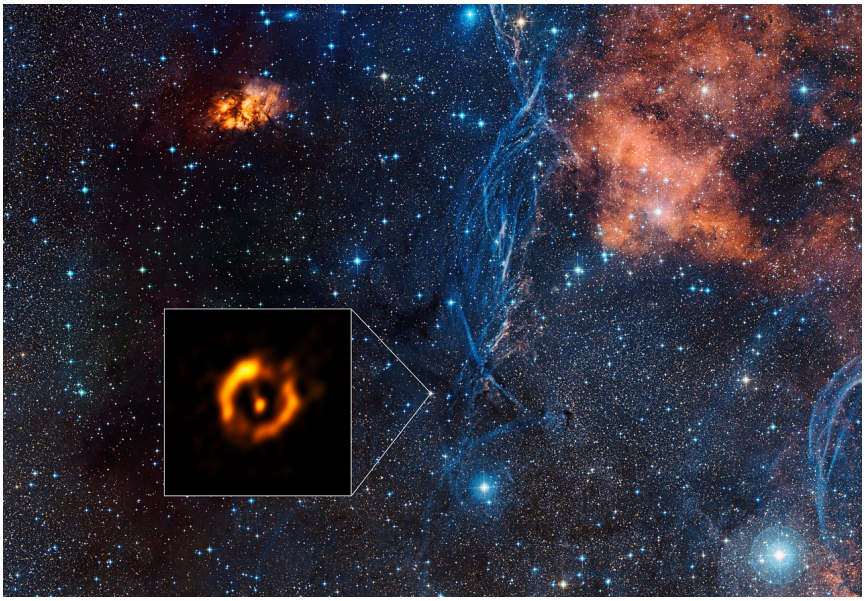

Figure 2.4: Images taken by the Very Large Telescope Interferometer (VLTI) of the dusty ring around the aging double star IRAS 08544-4431. By combining light from four different telescopes, the VLTI reaches the optical resolution of a telescope of 150m radius. Aperture synthesis was here combined with a technique allowing to block the light of the star in order to reveal the star's surroundings (ESO/Digitized Sky Survey 2).

is not viable in the context of live imaging and/or biological imaging since it requires the samples to be prepared — for instance coated with a reflective material. In the

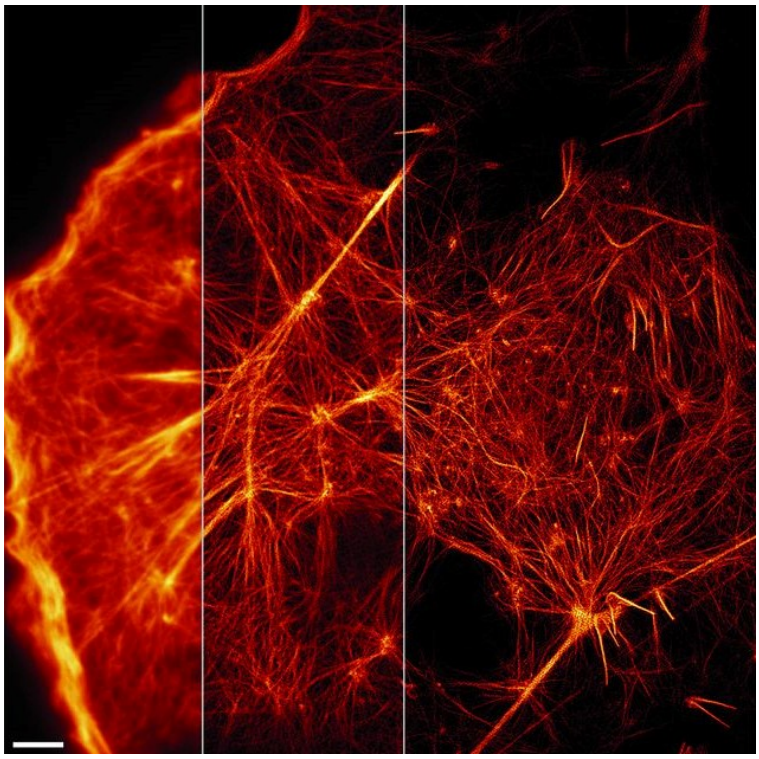

Figure 2.5: Image of the cytoskeleton of a cell taken with three different microscopy techniques (the scale bar indicates 3μm). The leftmost part is taken with conventional diffraction limited total internal reflection fluorescence (TIRF) and has a 220 nm resolution. The second strip illustrated structured illumination (SIM) TIRF and shows a 97 nm resolution. The last strip shows non linear SIM-TIRF, based on the patterned excitation of the photoswitchable fluorescent proteins (62 nm resolution). This image is taken from [Li 15].

context of biological imaging, the need for techniques allowing to image samples without damaging them has led to many important developments. The chemistry Nobel prize was awarded in 2014 to Eric Betzig, Stefan Hell and William Moerner for the development of super-resolved fluorescence microscopy. The use of fluorescent molecules allows to avoid using an illumination system, which means that the complete characterization of the illumination and its coherence properties is in this case not necessary. However, when a high number of fluorophores are present

in a sample, their diffraction patterns still overlap. The work of Betzig, Hell and Moerner [Betzig 06, Klar 00, Hell 94, Moerner 89] has been to make use of the fluorescence properties of different fluorophores to overcome this limitation: using structured illumination and different illumination wavelengths, fluorophores can be made to emit individually or in very small numbers. In this situation, a post treatment on the image can allow to determine their position knowing the properties of the imaging system. By combining many observation cycles, or scanning the samples, the location of the fluorophores can thus be determined with a precision overcoming Abbe's criterion. Knowledge of the imaging system (shape of the diffraction pattern in order to extract a precise position) is combined with sparsity of the signal (no two diffraction patterns overlap) and multiple acquisition frames to achieve extraordinary[6] resolution powers (see figure 2.5).

2.3 Elements of parameter estimation theory

We stated at the beginning of this chapter that electromagnetic fields have fluctuations. These fluctuations have various origins. Thermal light is constituted of many different contributions. These contributions are independent, such that their sum must be described in statistical terms. In the case of laser light, the electromagnetic field originates from a stable source. Even so, spontaneous emission cannot be neglected in the emission process. Moreover, vibration in the laser cavity, from the mirrors for instance, will introduce additional fluctuations. Finally, the complete quantum description of light emission introduces fundamental fluctuations which are inherent to the quantum nature of light. The most fundamental limitation is given by the Heisenberg uncertainty principle[7]. All these fluctuations are modeled through the concept of noise.

2.3.1 Noise

The study of noise in physics lies at the confluence of several different fields which are information theory, statistics, statistical physics, stochastic processes and probability theory [Réfrégier 02]. Noise has two origins: it can be inherent to the signal itself, or be introduced by the detection process. No measurement is a priori free from noise. Probability theory studies and characterizes noise. We are generally interested in the second order properties of noise, which are called correlation —

[6] A resolution power of the order of the Angström is reported in [Weisenburger 17].

[7] For two conjugate variables x and p, the product of their variance is bounded: $\sigma_x \sigma_p \geq \hbar/2$.

for example as in the definition of the cross-correlation function (defined in 2.1). In the context of an experiment, we study the random variable X. By making M different measurements, we have access to the realizations of this random variable $\{m_i\}$, $i \in [\![1, M]\!]$. All the measurements m_i belong to the ensemble of the possible outcome values for the variable X which is $\{x_i\}$ (in the discrete case — in the continuous case, the possible outcome values belong to a continuous interval). The frequency of outcome x_i is given by the probability p_i (with $\sum_i p_i = 1$). The expected value of X and its variance are defined respectively by:

$$\mathbb{E}[X] = \sum_{i_1}^{M} x_i p_i \tag{2.31}$$

$$\sigma^2(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]. \tag{2.32}$$

$\sigma(X)$ is the *standard deviation*. The n^{th} order momentum of X is defined as $\mathbb{E}(X^n)$. From the M measurements we can for example extract the mean of this random variable realization $\overline{X} = \left(\sum_i^M m_i\right)/M$.

Signal to noise ratio A central tool in the study of noise is the *signal to noise ratio* (SNR). It is defined by

$$SNR = \frac{P_{signal}}{P_{noise}} \tag{2.33}$$

which is the ratio of the meaningful information — the signal — to the superfluous and unwanted information — the noise. P is the average power and the signal to noise ratio expresses the ratio of the power of the signal without noise on the power of the noise without signal.

Noise and probability distributions We present here some important noise probability distributions and the contexts in which they are used.

Gaussian distribution The well known central limit theorem states that the addition of independent random variables of finite second order momentum converges to a normal distribution. This is why it is common to make the hypothesis that the noise which is *added* to a signal is gaussian. The normal distribution is given by

$$P(X = x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2.34}$$

where μ is equal to the expected value and σ to the standard deviation.

Poissonian distribution The noise distribution associated with the probability of a given number $k \in \mathbb{N}$ of events occurring in a fixed interval of time is the Poisson distribution:

$$P_\theta(X = k) = \frac{\theta^k e^{-k}}{k!}. \tag{2.35}$$

For instance, $P(X = k)$ gives the probability to count k particles in a fixed period of time. θ is the characteristic parameter of the distribution — it is also equal to its expected value and its variance. *Shot noise* is a type of electronic noise, which originates in the discrete nature of electrical charges — or of photons in optical devices — and is modeled by a Poisson process.

Other types of noise *White noise* refers to a random process in which the spectral power density[8] of noise is constant for all frequencies. *Quantization noise* appears in the context of numerical signal processing. It originates in the mapping of the signal value to a smaller set of values, given by the number of bits this value can be represented on.

Limits All these types of noise introduce fluctuations in the signal. Some of these noises can be completely suppressed by a careful engineering of the measurement apparatus, or be made sufficiently small that they play a negligible role. However, some of them are *inherent* to the nature of the signal one wants to measure, such as shot noise.

Quantum mechanics, through the Heisenberg uncertainty principle, states that some quantities cannot be known simultaneously with arbitrary precision. This principle limits the precision with which one can estimate certain parameters and is referred to as quantum noise. The *standard quantum limit* designates the theoretical limit reachable with coherent states, while the *Heisenberg limit* gives the absolute limit one can achieve, even with specifically engineered quantum states [Caves 81].

2.3.2 Limits on parameter estimation

Estimation theory is a branch of signal processing which is concerned with the estimation of parameters in the presence of noise. Its aim is to derive the ultimate sen-

[8] The spectral power density describes how the power of a signal is distributed over frequencies. It is defined as the Fourier transform of the autocorrelation function of the signal. It is commonly expressed in watts per hertz (W/Hz).

sitivity which can be attained in a given context — that is with a given type of measurement. We give here an introduction to the main concepts of estimation theory. Fuller descriptions and derivations can be found in [Réfrégier 02, Pinel 10, Jian 14]. Figure 2.6 illustrates the process by which one gains information on a parameter

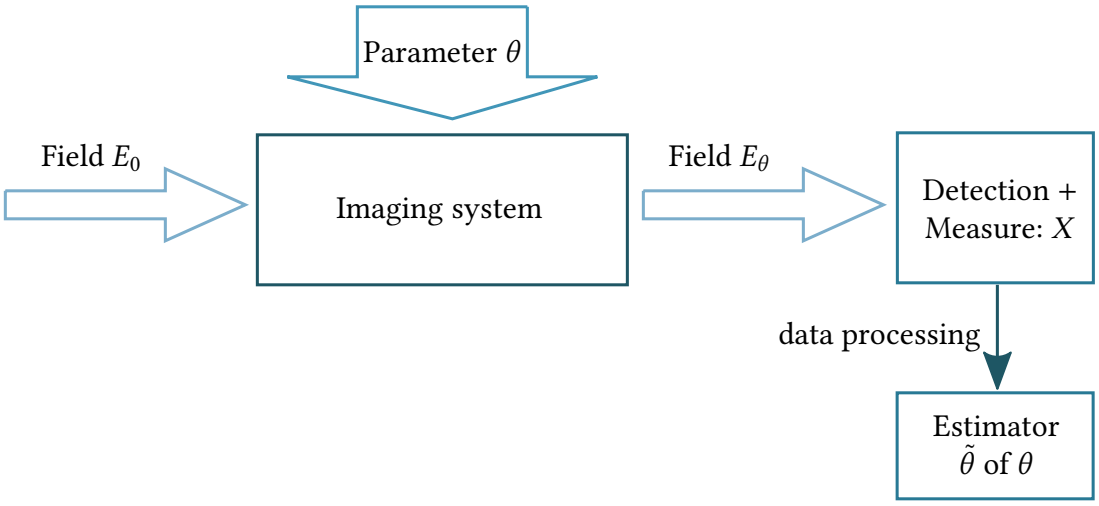

Figure 2.6: Parameter estimation.

through optical measurement. An input field E_0 is altered in a way that depends on a parameter θ. At the output of the imaging system, the field is E_θ. The measurement process gives access to a value x, a realization of the random variable X. From this value — or a collection of these values corresponding the outcome of different measurements — one calculates, possibly with data processing, an estimation of θ: $\tilde{\theta}$. As an example, if X admits the Poisson distribution P_θ, the mean of the measure $\boldsymbol{x} = \{x_1, x_2, ..., x_M\}$ is an estimator of θ. For the purposes of notation, θ will designate both the *parameter* and the *true parameter value*. As an example, further in this chapter, we will try to estimate the displacement of a point source with respect to a reference position or the separation between two incoherent point sources. The displacement will in both these cases be the parameter, and will be expressed in meters.

Bias of an estimator The *bias* of an estimator is defined as

$$b_{\tilde{\theta}} = \mathbb{E}[\tilde{\theta}] - \theta. \tag{2.36}$$

This quantity quantifies the *accuracy* of this estimator by measuring how close to the true parameter value θ the estimation we make is. If $b_{\tilde{\theta}} = 0$, the estimator is

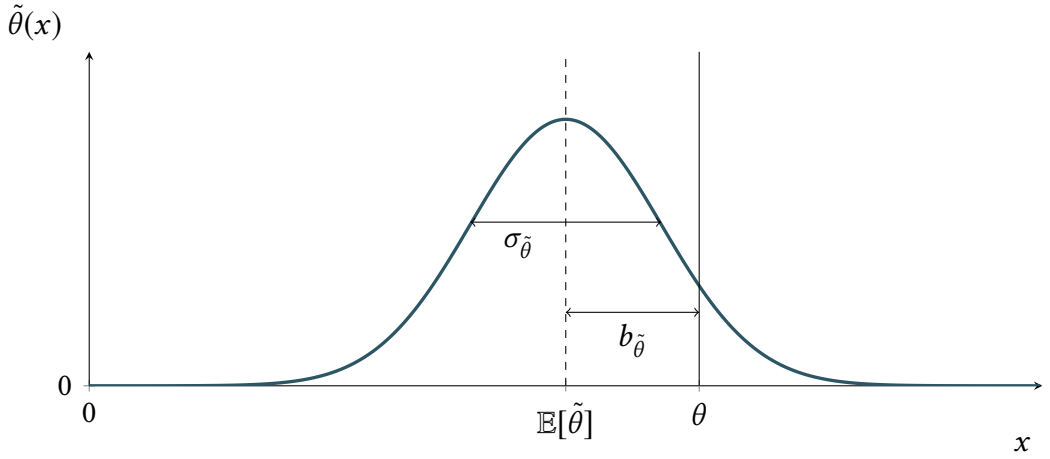

Figure 2.7: Estimator function $\tilde{\theta}$. $b_{\tilde{\theta}}$ is the bias of the estimator and $\sigma_{\tilde{\theta}}$ the standard deviation.

said to be *unbiased*.

Variance of the estimator Another important quantity to consider is the *sensitivity* with which we estimate a parameter — or the *precision*. The sensitivity is characterized by the variance of the estimator:

$$\sigma_{\tilde{\theta}}^2 = \mathbb{E}[\tilde{\theta}^2] - \left(\mathbb{E}[\tilde{\theta}]\right)^2 \tag{2.37}$$

and the standard deviation $\sigma_{\tilde{\theta}}$. The *mean-squared error* of the estimator is $MSE = \sigma_{\tilde{\theta}}^2 + b_{\tilde{\theta}}^2$.

Likelihood and log-likelihood Our goal is to be able to deduce from a measurement sample $\boldsymbol{x}$ the parameter θ of a probability distribution. The *likelihood* L_θ is the probability of measuring $\boldsymbol{x}$:

$$L_\theta(\boldsymbol{x}) = \prod_{i=1}^{M} P_\theta(x_i)\,. \tag{2.38}$$

The *log-likelihood* is defined as

$$l_\theta(\boldsymbol{x}) = \ln L_\theta(\boldsymbol{x}) = \sum_{i=1}^{M} \ln P_\theta(x_i). \tag{2.39}$$

The Shannon information brought by the set of measurements $\boldsymbol{x}$ is $-\sum_i \ln P_\theta(x_i)$: the log-likelihood is the opposite of the Shannon information. With the definition of the likelihood, we can rewrite the definitions of the estimator's expectation value and sensitivity:

$$\mathbb{E}[\tilde{\theta}] = \int L_\theta(\boldsymbol{x})\,\tilde{\theta}(\boldsymbol{x})\,d\boldsymbol{x} \tag{2.40}$$

$$\sigma_{\tilde{\theta}}^2 = \int L_\theta(\boldsymbol{x})\left(\tilde{\theta}(\boldsymbol{x}) - \mathbb{E}[\tilde{\theta}]\right)^2 d\boldsymbol{x}. \tag{2.41}$$

Fisher information and Cramér-Rao bound We wish to know how precise a given measurement can be regardless of the processing or post-treatment one might apply on it: we thus aim at establishing a lower bound to the variance $\sigma_{\tilde{\theta}}$ regardless of the choice of the estimator $\tilde{\theta}$. We note that the estimator $\tilde{\theta}(x)$ is independent of θ. Using the definition of the expectation value of an estimator we can write

$$\int L_\theta(\boldsymbol{x})\left(\tilde{\theta}(\boldsymbol{x}) - \mathbb{E}[\tilde{\theta}]\right) d\boldsymbol{x} = 0 \tag{2.42}$$

and derive this expression with respect to θ (all the derivatives are taken at $\theta = \theta$, that is the parameter and its value are referred to by the same letter):

$$-\int L_\theta(\boldsymbol{x})\left.\frac{d\mathbb{E}[\tilde{\theta}]}{d\theta}\right|_\theta d\boldsymbol{x} + \int \left.\frac{\partial L_\theta(\boldsymbol{x})}{\partial\theta}\right|_\theta \left(\tilde{\theta}(\boldsymbol{x}) - \mathbb{E}[\tilde{\theta}]\right) d\boldsymbol{x} = 0. \tag{2.43}$$

We can simplify this expression by integrating the first term and noticing that $\left.\frac{\partial l_\theta}{\partial\theta}\right|_\theta = \frac{1}{L_\theta}\left.\frac{\partial L_\theta}{\partial\theta}\right|_\theta$:

$$\left.\frac{d\mathbb{E}[\tilde{\theta}]}{d\theta}\right|_\theta = \int L_\theta(\boldsymbol{x})\left.\frac{\partial l_\theta(\boldsymbol{x})}{\partial\theta}\right|_\theta \left(\tilde{\theta}(\boldsymbol{x}) - \mathbb{E}[\tilde{\theta}]\right) d\boldsymbol{x}. \tag{2.44}$$

We use the Cauchy-Schwarz inequality and write

$$\left(\left.\frac{d\mathbb{E}[\tilde{\theta}]}{d\theta}\right|_\theta\right)^2 \leq \left(\int L_\theta(\boldsymbol{x})\left(\tilde{\theta}(\boldsymbol{x}) - \mathbb{E}[\tilde{\theta}]\right)^2 d\boldsymbol{x}\right) \times \left(\int L_\theta(\boldsymbol{x})\left(\left.\frac{\partial l_\theta(\boldsymbol{x})}{\partial\theta}\right|_\theta\right)^2 d\boldsymbol{x}\right). \tag{2.45}$$

The last term of the right-hand part of the inequality can be rewritten as $M\int L_\theta(x)\left(\left.\frac{\partial \ln(L_\theta(x))}{\partial\theta}\right|_\theta\right)^2 dx$ where M is again the number of measurements per-

formed. The first term corresponds to the variance of $\tilde{\theta}$. We define the *Fisher information* as

$$I_F(\theta) = \int L_\theta(x) \left(\left. \frac{\partial \ln(L_\theta(x))}{\partial \theta} \right|_\theta \right)^2 dx = \int \frac{1}{L_\theta(x)} \left(\left. \frac{\partial L_\theta(x)}{\partial \theta} \right|_\theta \right)^2 dx. \tag{2.46}$$

We can thus rewrite inequality 2.45 as:

$$\left(\left. \frac{d\mathbb{E}[\tilde{\theta}]}{d\theta} \right|_\theta \right)^2 \leq \sigma_{\tilde{\theta}}^2 \times MI_F(\theta) \tag{2.47}$$

The Fisher information only depends on the measurement noise: it does not depend on the estimator $\tilde{\theta}$. This means that the Fisher information does not depend on how we use the measurement to calculate the estimator $\tilde{\theta}$. However, it does depend on the *value* of θ. The Fisher information measures the amount of information that one measurement reveals about the value of θ. Moreover, we note that if the estimator is unbiased, $\mathbb{E}[\tilde{\theta}] = \theta$ which implies that $\left. \frac{d\mathbb{E}[\tilde{\theta}]}{d\theta} \right|_\theta = 1$. We can now rewrite equation 2.47 as

$$\sigma_{\tilde{\theta}}^2 \geq \frac{1}{MI_F(\theta)}. \tag{2.48}$$

This inequality is known as the *Cramér-Rao bound* [Cramér 99, Rao 47]. It states that no estimator of θ can have a variance smaller than that given by the Cramér-Rao bound. The estimators which saturate the bound are said to be efficient estimators. We define

$$\delta\theta_{min} = \frac{1}{\sqrt{MI_F(\theta)}}. \tag{2.49}$$

This equality states that the smallest variation of θ we can discriminate from noise in M measurements (that is with a SNR of 1) is given by $\delta\theta_{min}$.

The quantum Cramér-Rao bound can be derived using the tools of quantum mechanics [Caves 81, Helstrom 69, Pinel 10, Jian 14]. The bound we derived here can be referred to as the semi-classical Cramér-Rao bound. The quantum Cramér-Rao bound takes into account all the positive operator-valued measures[9] — which means that this bound, as opposed to the semi-classical one, is calculated for all possible measurements. The quantum Cramér-Rao bound reaches a similar expression — the quantum Fisher information is expressed using the density operator of the state. The quantum Cramér-Rao bound of coherent states is equal to the standard quantum limit introduced earlier in this chapter.

[9] A measure assigns numbers — here probabilities — to the subsets of a given set. The elements of a measure are probability operators. In quantum mechanics they are positive operators, hence the positive operator-value measure name [Barnett 09].

Statistical distance The problem of parameter estimation can be apprehended in a different manner: with a given experimental outcome $\{x_1, x_2, ...\}$, we want to be able to know with certainty whether the parameter of the distribution is equal to θ or θ'. We want to be able to distinguish the probability distributions P_θ and $P_{\theta'}$. In order to tackle this problem, one needs to use the concept of *statistical distance* [Jian 14][10]. This study is however outside the scope of this thesis.

2.3.3 Parameter estimation for optical resolution

The experimental situation we are interested in is the following: we want to evaluate the separation between two point sources with the best possible accuracy and sensitivity. This statement illustrates why parameter estimation theory is useful for us: its gives a tool to find the best sensitivity we may aim to reach — and can also allow us to evaluate whether a certain type of measurement saturates the Cramér-Rao bound. In order to assess that problem, we will first consider another experimental situation: the estimation of the displacement of *one* beam with respect to a reference position. We make here the assumption that the only source of noise present in our measurement is shot noise.

2.3.3.1 Intensity measurement

We consider intensity measurements, which are the most common (and easiest) measurements one can make. For instance, a camera records intensity on each one of its pixels. The electric field we consider in this part is a coherent, single-mode field. We introduce the total number of photons of the field N and give the expression of the mean value of the electric field in the image plane during one coherence time $\Delta\tau$:

$$\bar{E}(\boldsymbol{r}, \theta) = 2\sqrt{N}\mathscr{E} u_0(\boldsymbol{r}, \theta) \tag{2.50}$$

where $u_0(\boldsymbol{r}, \theta)$ is the normalized spatial distribution of the field. We assume that the total intensity of the beam is independent of the parameter θ. In the case of the displacement of a beam, one can easily see that it is indeed the case. The mean photon number during the same time interval is given by

$$\bar{n}(\boldsymbol{r}, \theta) = \left|\bar{E}(\boldsymbol{r}, \theta)\right|^2 / (4\mathscr{E}) = N \left|u_0(\boldsymbol{r}, \theta)\right|^2 . \tag{2.51}$$

We wish to establish the Cramér-Rao bound for the evaluation of θ in the sole presence of shot noise. We make the assumption that the light can be described as

[10] The relative entropy, or Kullback-Leibler divergence is a measure of the discrepancy between two probability distributions.

being composed of independent photons, meaning its intensity fluctuations follow a Poisson distribution: the probability of measuring n_k photons on pixel k during one coherence time[11], given the parameter θ, is given by

$$P_\theta\left(k = n_k\right) = \frac{\bar{n}_k\left(\theta\right)^{n_k} e^{-\bar{n}_k(\theta)}}{n_k!} \tag{2.52}$$

with $\bar{n}_k\left(\theta\right)$ the real number of photons. The likelihood of measuring a given intensity image $\boldsymbol{n} = \{n_k\}$ is given by

$$L_\theta\left(\boldsymbol{n}\right) = \prod_k P_\theta^k\left(k = n_k\right). \tag{2.53}$$

We also can write

$$\int \frac{\partial \bar{n}}{\partial \theta}\left(\boldsymbol{r}, \theta\right) d\boldsymbol{r} = 0 \tag{2.54}$$

since we specified that the total intensity is independent of θ (the quantity $\int \bar{n}\left(\boldsymbol{r}, \theta\right) d\boldsymbol{r}$ is a constant). We can derive the Fisher information from these definitions (the complete derivation can be found in [Delaubert 07]):

$$I_F = \int \left[\frac{1}{\bar{n}\left(\boldsymbol{r}, \theta\right)}\left(\frac{\partial \bar{n}\left(\boldsymbol{r}, \theta\right)}{\partial \theta}\right)^2 - \frac{\partial^2 \bar{n}\left(\boldsymbol{r}, \theta\right)}{\partial \theta^2}\right]\Bigg|_{\theta=0} d\boldsymbol{r}. \tag{2.55}$$

Using equation 2.54 we can show that the integral of the left part is equal to zero. By interverting the order of the norm and the derivative, one can further show that

$$I_F = 4N \int \left[\frac{\partial \left|u_0\left(\boldsymbol{r}, \theta\right)\right|}{\partial \theta}\Bigg|_{\theta=0}\right]^2 d\boldsymbol{r}. \tag{2.56}$$

This results tells us that a displacement can be measured with precision if the field has a spatial profile such that the overlap between its spatial mode u_0 and its displaced version is small. The measurement of the displacement of fields with rapid spatial variations is more precise than that of fields with uniform spatial profiles.

2.3.3.2 Case of a gaussian transverse profile

To go further, we need to specify the spatial profile of the field in equation 2.56 and its dependence on the parameter we want to probe. We make the hypothesis that

[11] We take the measurement time equal to the coherence time $\Delta\tau$.

the spatial profile of the field is Gaussian: we describe the situation of the image of a point source displaced from the center of the imaging system. This assumption is coherent with the fact that we want to consider a point source — we showed in part 2.1.4 that the intensity distribution of a point source in a diffraction limited imaging system could be well approximated by a Gaussian function. To make use of the notations we introduced in the previous chapter, we note that the field of a gaussian beam at $z = 0$ possesses a spatially uniform phase profile. The parameter we want to probe is the displacement in the transverse plane. It can be characterized by two parameters d and β. In cartesian coordinates, with respect to a reference position that we take to be $(x_0, y_0) = (0, 0)$, the displacement is defined as $(d \cos\beta, d \sin\beta)$. In this paragraph we take $\beta = 0$ to keep notations simple. The generic parameter θ introduced in the previous notations is thus replaced by the displacement amplitude d. We will thus, without loss of generality, write $u_0(\boldsymbol{r}, \theta) = HG_{00}(x - d, y, z = 0)$[12]. A simple derivation gives

$$\frac{\partial HG_{00}(x - d, y, 0)}{\partial d} = \frac{1}{w_0} HG_{10}(x - d, y, 0). \tag{2.57}$$

Using this equality and equation 2.56 we can express the Fisher information as

$$I_F = \frac{4N}{w_0^2}. \tag{2.58}$$

Finally, the Cramér-Rao bound states that with any intensity measurement, the smallest displacement d_{min} that can be distinguished from noise is

$$d_{min} = \frac{w_0}{2\sqrt{N}}. \tag{2.59}$$

This result is well-known [Fabre 00, Treps 04, Hsu 04]: parameter estimation theory indicates that, in the case of a centered beam ($d = 0$), the sensibility with which one can measure the displacement of a field around does not depend on said displacement. Since w_0 can be made as small as the order of magnitude of λ ($w_0 \geq \lambda$), using an intense field ($N \gg 1$), d_{min} can be made quite small compared to λ.

[12]The choice of the Hermite-Gaussian beam is justified by the fact that we choose $z = 0$: this ensures that the phase term plays no role. The formalism of the Hermite-Gaussian beam also ensures that the modes are orthogonal and are normed to unity. The validity of this choice for $z \neq 0$ is not, to our knowledge, demonstrated.

Sensitivity enhancement strategies In the case where one can control the parameters of the field, we see that in this measurement scheme, the measurement sensitivity can be enhanced using classical means by:

- using a beam with a smaller waist w_0 (ie. by increasing the aperture stop size, see 2.2).
- increasing the number of photons N (that is the intensity of the beam).

With non-classical illumination, one can engineer the field to have sub-Poissonian quantum intensity fluctuations described by a noise standard deviation $\sigma_I < 1/\sqrt{N}$[13] [Caves 81, Barnett 03]. In this case, the Cramér-Rao bound gives

$$d_{min} = \frac{w_0 \sigma_I}{2\sqrt{N}}. \tag{2.60}$$

2.3.3.3 Phase and field measurements

In [Delaubert 07], the Cramér-Rao bound for the transverse displacement of a beam is also derived in the case of phase and field measurements. We make here a brief summary of the main results presented in this reference. We first write the transverse profile of the field as

$$u_0(\boldsymbol{r}, \theta) = |u_0(\boldsymbol{r}, \theta)| \, e^{i\phi_0(\boldsymbol{r},\theta)} \tag{2.61}$$

where $\phi_0(\boldsymbol{r}, \theta)$ is the phase distribution of the field. The description we use here is only semi-classical and not a rigorous quantum mechanical derivation since we make the assumption that the intensity and the phase of the field can be measured independently. We introduce the following quantities:

$$\frac{1}{a^2} = \int \left[\left. \frac{\partial |u_0(\boldsymbol{r}, \theta)|}{\partial \theta} \right|_{\theta=0} \right]^2 d\boldsymbol{r} \tag{2.62}$$

$$\frac{1}{b^2} = \int \left[\left. \frac{\partial u_0(\boldsymbol{r}, \theta)}{\partial \theta} \right|_{\theta=0} \right]^2 d\boldsymbol{r} \tag{2.63}$$

$$\frac{1}{c^2} = \int |u_0(\boldsymbol{r}, 0)|^2 \left[\left. \frac{\partial |\phi_0(\boldsymbol{r}, 0)|}{\partial \theta} \right|_{\theta=0} \right]^2 d\boldsymbol{r} \tag{2.64}$$

[13] The field is said to be *squeezed* [Grynberg 10].

which are linked by the following equation: $\frac{1}{b^2} = \frac{1}{a^2} + \frac{1}{c^2}$. a, b and c are respectively associated with field, intensity and phase measurements. The Fisher informations associated with each type of measurement can be expressed as $I_F^{field} = 4N/b^2$, $I_F^{intensity} = 4N/a^2$ and $I_F^{phase} = 4N/c^2$. As a consequence, we can write:

$$I_F^{field} = I_F^{intensity} + I_F^{phase}. \tag{2.65}$$

This result, while elegant, is not surprising: the field having phase and intensity components, it is only natural that the information contained in the field should be the sum of that contained in the phase and intensity parts. However, in the case we address here, that is of a gaussian intensity profile $u_0(\boldsymbol{r}, \theta)$ (again with $z = 0$) we have $\frac{\partial \phi_0(\boldsymbol{r},0)}{\partial \theta} = 0$ which means that $I_F^{phase} = 0$ and

$$I_F^{field} = I_F^{intensity}. \tag{2.66}$$

Hence, the Cramér-Rao bound derived in 2.3.3.2 is valid for any measurement scheme.

2.3.3.4 Split detector sensitivity

A common and simple way to measure the displacement of a beam is to use a split-detector. A split-detector[14] is the simplest version of a difference measurement [Treps 05]. In the following derivation, the time dependency will not be written explicitly, and the mean values will be averaged over time. We define D_- as the left half of the detector and D_+ as the right half. The intensity which is measured by either halves of the detector is given by

$$\begin{aligned} I_- &= \iint_{D_-} E^*(x,y)\, E(x,y)\, dxdy \\ I_+ &= \iint_{D_+} E^*(x,y)\, E(x,y)\, dxdy \end{aligned} \tag{2.67}$$

and the signal delivered by the split detector is $s \propto (I_+ - I_-)$. We decompose the field on a mode basis such that the first mode of the basis u_0 is the *mean field* mode:

$$E(x,y) = \sum_i X_i\, u_i(x,y). \tag{2.68}$$

[14] A split-detector consists in two photodiodes which have the shapes of two half-disks. The intensity signals of the two halves are subtracted and normalized by the total intensity: if the beam is symmetrical, this signal gives the distance of the barycentre of the beam with respect to the center of the detector.

u_0 is the mean field mode, which means we have $\langle X_0 \rangle = E_0$ (we can choose $E_0 \in \mathbb{R}$) and $\langle X_{i>0} \rangle = 0$. Since we are interested in the fluctuations of these quantities, we write $X_i = \langle X_i \rangle + \delta X_i$. We can now write

$$I_+ - I_- = \sum_{i,j} X_i^* X_j \left[\iint_{D_+} u_i^* u_j - \iint_{D_-} u_i^* u_j \right] \tag{2.69}$$

and from this

$$\begin{aligned} \delta\left(I_+ - I_-\right) &= \sum_{i,j} \left(\langle X_i^* \rangle \delta X_j + \langle X_j \rangle \delta X_i^* \right) \left[\iint_{D_+} u_i^* u_j - \iint_{D_-} u_i^* u_j \right] \\ &= \langle X_0 \rangle \sum_j \left(\delta X_j + \delta X_j^* \right) \times \left[\iint_{D_+} u_0^* u_j - \iint_{D_-} u_0^* u_j \right]. \end{aligned} \tag{2.70}$$

We now introduce the *flipped* transverse mode $u_1(x, y)$ as

$$u_1(x,y) = \begin{cases} HG_{00}(x,y) & \text{if } x > 0 \\ -HG_{00}(x,y) & \text{if } x < 0 \end{cases} \tag{2.71}$$

and we complete the basis for $i > 1$ so as to form an orthonormal basis. Using the definition of u_1, we can rewrite equation 2.70 as

$$\delta\left(I_+ - I_-\right) = \langle X_0 \rangle \sum_j \left(\delta X_j + \delta X_j^* \right) \times \iint_D u_1^* u_j. \tag{2.72}$$

Since the mode basis we chose is orthonormal, we finally have

$$\delta\left(I_+ - I_-\right) = \langle X_0 \rangle \left(\delta X_1 + \delta X_1^* \right) \tag{2.73}$$

The fluctuation of the signal we measure with a split detector is thus proportional to the projection of the field on the flipped mode u_1.
In the case of a displacement measurement, in the limit of small displacement, and using a Taylor series, we can write (see 2.57):

$$HG_{00}\left(d, 0\right) = HG_{00}\left(0, 0\right) + \frac{d}{w_0} HG_{10}\left(0, 0\right) + o\left(d\right). \tag{2.74}$$

Using this equation and noting that $\int u_1^* \, HG_{00} = 0$ and $\int u_1^* \, HG_{10} = \sqrt{2/\pi}$, the signal we measure with a split detector is proportional to $\left(\sqrt{\frac{2}{\pi}} \frac{d}{w_0} \right)^2$. Hence, since the noise we consider in both cases are equal, the split detection has a $2/\pi \simeq 0.64$ (in)efficiency relative to an ideal intensity measurement. This widespread displacement measurement technique is therefore suboptimal.

2.4 Distance measurement using a spatial mode demultiplexer

As we stated beforehand, the problem we aim to tackle is that of the resolution of two incoherent point sources. While old and very well described since the 19$^{\text{th}}$ century, this problem is still very much current, as evidenced by the activity in this research domain. We presented in section 2.2 two domains in which this problem is particularly relevant: optical microscopy and astronomy. In this section we show that spatial mode demultiplexing is, for this particular problem, an optimal measurement scheme.

2.4.1 Hypothesis testing

We first make a brief mention of the use of spatial demultiplexing in a closely related context. A similar yet different problem than the separation estimation for two point sources was treated in [Helstrom 73]. The physical situation is the same as the one we consider in this chapter. The question which is addressed is however slightly different: how can an observer know for sure he is seeing two close point sources as indeed two rather than one? Using quantum detection theory to find the optimum processing of the aperture field, the author seeks the ultimate resolvability of the sources[15].

The optimum strategy found in this work is to decompose the aperture field into a countable set of spatial modes. This is, to our knowledge, the first time that this type of projective measurement was put forward. This idea will be at the heart of the next section. In [Helstrom 73] however, the spatial modes profiles depend on the separation between the two sources, thus hindering the usefulness of this method.

2.4.2 Detection modes

We introduce in this part the formalism of *noise* or *detection* modes developed in the group in the last decade(s). We first define the *normalized mean photon field* at $\theta = 0$ as

$$v_0(\boldsymbol{r}, \theta) = \frac{u_0(\boldsymbol{r}, \theta)}{\|u_0(\boldsymbol{r}, \theta)\|} \tag{2.75}$$

[15] An important hypothesis which is made (and which we also have made) is that no background light is present, such that only the quantum nature of the light introduces noise in the measurement.

with $\mathbf{E}^{(+)}(\boldsymbol{r}, t; \theta) = u_0(\boldsymbol{r}, \theta) e^{i(kz-\omega_0 t)}$ (see equation 1.25). We define the *detection mode* as

$$v_1(\boldsymbol{r}) = \left. \frac{\partial u_0(\boldsymbol{r}, \theta)/\partial\theta}{\|\partial u_0(\boldsymbol{r}, \theta)/\partial\theta\|} \right|_{\theta=0}. \tag{2.76}$$

Finally, we complete this mode basis by other orthonormal modes $v_{i>1}$. The mode basis $\{v_{i\geq 1}\}$ does not depend on θ since v_1 was defined at $\theta = 0$. It can be shown (see [Pinel 12]) that the quantum Cramér-Rao bound for a pure Gaussian state in this mode basis has a very simple expression:

$$I_F = 4\Gamma^{-1}_{\theta=0,[1,1]} \left\| \frac{\partial u_0(\boldsymbol{r}, \theta)}{\partial\theta} \right\|_{\theta=0} \tag{2.77}$$

where $\Gamma^{-1}_{\theta,[1,1]}$ is the first left top element of the symmetrized covariance matrix of the field calculated in the basis $\{v_{i\geq 1}\}$[16]. In the case of a coherent state, we have $\Gamma^{-1}_{\theta,[1,1]} = 1$. We can thus rewrite the Fisher information as

$$I_F = N_\theta \left[4 \left\| \frac{\partial v_0}{\partial\theta} \right\|^2_{\theta=0} + \left(\left. \frac{\partial N_\theta/\partial\theta}{N_\theta} \right|_{\theta=0} \right)^2 \right] \tag{2.78}$$

where $N_\theta = \|u_0(\boldsymbol{r}, \theta)\|^2$. If the total number of photons does not depend on the parameter of interest θ, the Fisher information only depends on the normalized first derivative of the normalized mean field. Therefore, one needs only to measure this specific mode to reach the best sensitivity. This result has quite a broad scope and was used in a number of works [Treps 05, Lassen 07, Delaubert 08, Pinel 12, Pinel 10, Jian 14]. Many of the results we will present later in this section can be derived using equation 2.77.

2.4.2.1 Case of displacement measurement

Measurement of a single precise mode of the field is possible, for instance through the use of homodyne detection. Homodyne detection uses a second intense field, the reference field, which has a well controlled spatial mode. It allows to very precisely measure the fluctuations of the signal field on the spatial mode defined by the reference field. This measurement scheme was used in [Hsu 04, Delaubert 06a, Delaubert 06b] but also in [Sun 14].

[16] See [Pinel 12, Jian 14] for more details.

2.4.3 Spatial demultiplexing for distance and separation measurement

The idea we present in this section is the following. In the case of separation or displacement measurement of Gaussian beams, the measurement scheme which consists in measuring the intensities of the modal components of the electrical field in the Hermite-Gauss mode basis saturates the Cramér-Rao bound. The information this measurement gives on the modal decomposition of the field is sufficient to determine the displacement or separation. The information which is lost through the use of this scheme could not allow for any precision improvement for transverse distance measurement.

2.4.3.1 Derivation of the Cramér-Rao bound for separation measurement

While the concepts introduced in this chapter date back to more than ten years, a recent surge of interest was sparked by a series of papers [Tsang 16b, Nair 16, Lupo 16]. In these papers, two different groups make the complete derivation of the quantum Cramér-Rao bound on the separation measurement precision in the case of two incoherent, diffraction-limited, low-intensity sources:

$$I_F \simeq N\Delta k^2 \tag{2.79}$$

where N is the total number of photons reaching the detection plane and Δk^2 is the spatial variance of the derivative of the point-spread function $K(x,y)$ in the displacement direction (we again take x as the displacement direction):

$$\Delta k^2 = \iint_{-\infty}^{+\infty} \left| \frac{\partial K}{\partial x}(x,y) \right|^2 dxdy. \tag{2.80}$$

The Cramér-Rao bound for the two-dimensional problem is derived in [Ang 17].

The situation described here (distance between two point sources) and the one we treated earlier in this chapter (displacement of one point source) are not the same, and the expression of the Cramér-Rao bound of equation 2.58 cannot be compared to that of equation 2.79. We will first rephrase equation 2.79 for it to match our notations. Using the formalism we introduced in the previous section (the point-spread function is $K(x,y) = HG_{00}(x,y,z=0)$) and using equation 2.57, we have

$$\Delta k^2 = \iint_{-\infty}^{+\infty} \left[\frac{1}{w_0} |HG_{10}(x,y)| \right]^2 dxdy = \frac{1}{w_0^2}. \tag{2.81}$$

We can now rewrite the Fisher information of equation 2.79 as

$$I_F^q = \frac{N}{w_0^2}. \tag{2.82}$$

where the superscript q stresses the fact that this is the *quantum* Fisher information.

In order to compare this result with the derivation presented in the previous section we must make the following precision: in [Tsang 16b, Lupo 16], the distance which is to be measured is the separation between two point sources. The total intensity can be written as $|HG_{00}(-d/2, 0)|^2 + |HG_{00}(d/2, 0)|^2$. The point sources are located at a distance $d/2$ from the origin for the total separation to be equal to d. We suppose that the centroid of the system is known. We can rewrite equation 2.57 as

$$\frac{\partial HG_{00}(x - d/2, y, 0)}{\partial d} = \frac{1}{2w_0} HG_{10}(x - d/2, y, 0). \tag{2.83}$$

We suppose that the total number of photons is again equal to N. We thus say that the number of photons emitted by one of the two sources is $\alpha_i N$ with $\alpha_1 + \alpha_2 = 1$. We use the same derivation as in 2.3.3.2 and find that the Fisher information associated with one point source is

$$I_{F,i}^c = \frac{\alpha_i N}{w_0^2},\ i \in \{1, 2\} \tag{2.84}$$

where c expresses the fact that this is the (semi-)classical Fisher information. Since the two sources are incoherent, they are independent. Hence, we can sum both their Fisher informations to find the Cramér-Rao bound of the whole system. We find

$$I_F^c = \frac{N}{w_0^2}. \tag{2.85}$$

The quantum treatment and our semi-classical approach find, for this specific problem, the same bound: $I_F^c = I_F^q$. Another derivation of this result can be found in [Tsang 16a]. The link between classical and quantum resolution limits is also investigated in [Tsang 18] using a mean-squared error approach, that the author relates to compressed sensing derivations (see section 2.5). In [Rehacek 17a], the Cramér-Rao bound for the estimation of the centroid, separation and relative intensities of two incoherent point sources is derived.

2.4.3.2 Measurement schemes

The Cramér-Rao bound gives an information on the maximum sensitivity one can achieve. In [Tsang 16b, Lu 16, Tsang 17, Ang 17], Tsang and coworkers introduce a

measurement scheme which saturates the Cramér-Rao bound and consists precisely in spatial mode demultiplexing, whether in one or two dimensions. If the point-spread function of the imaging system is taken to be a Gaussian, the optimum demultiplexing basis is exactly the Hermite-Gaussian mode basis. In [Tsang 16b], it is shown that in the limit of an infinite number of modes, the information gathered on all the modes of this basis is equal to the Fisher information. The Fisher information acquired by demultiplexing only HG_{10} is derived in [Yang 17]. In [Rehacek 17b], the possible demultiplexing mode bases are examined. It is again showed that in the limit of small displacements, a mode basis generated using the derivatives of the point-spread function allows to access all the available information with one measurement. In [Lu 16], it is shown that demultiplexing can also be used to address the problem of hypothesis testing since both problems are related (see 2.4.1). In [Tsang 17], the author shows that a demultiplexing scheme in which the outputs can be made to interfere allows to gain access to any higher order momentum of the field. The validity of a simplified version of spatial mode demultiplexing for dis-

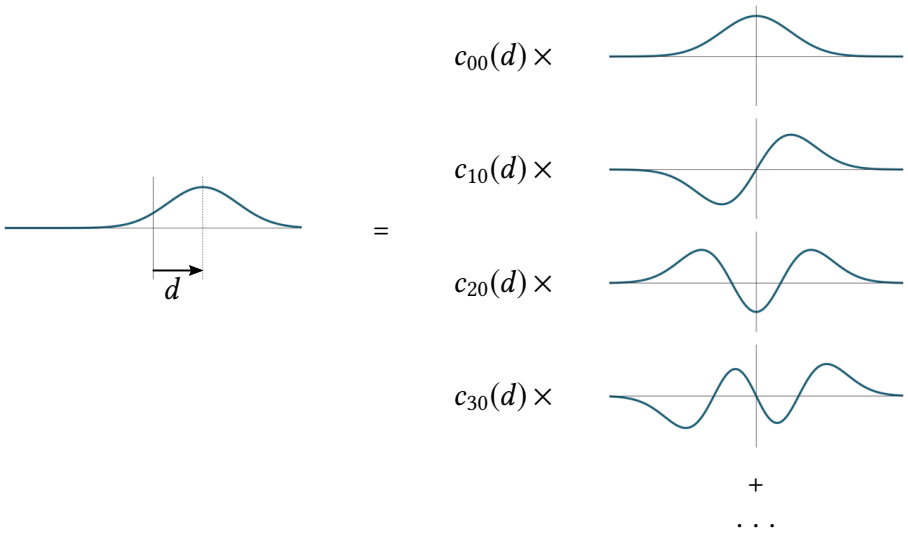

Figure 2.8: Distance measurement through demultiplexing. The $c_i\,(d)$ coefficients are the projection of the displaced Gaussian on the different modes of the Hermite-Gauss mode basis.

tance measurement was tested in recent years [Paúr 16, Tang 16, Yang 16, Tham 17] — the simplification coming from the fact that only one spatial mode is measured. In order to access the HG_{10} detection mode, several strategies are employed: a cav-

ity in [Yang 16] (as in [Sun 14]), a phase plate producing a flip mode in the case of [Tham 17] or a spatial light modulator in the case of [Paúr 16]. Interferences are used in [Tang 16].

2.4.4 Experimental evaluation

We developed an experimental setup to assess the performances of a demultiplexing system on the measurement of the distance between two incoherent sources. The originality in our approach resides in the use of a MPLC system. Since the MPLC can simultaneously demultiplex multiple spatial modes, we can have access to the intensity in several modes at the same time. In this way, we increase the amount of information about the separation that we can access. This also allows us to probe larger displacements (meaning we are not limited to the regime of $d/w_0 \ll 1$), as well as other parameters such as the direction of the displacement (the angle β in figure 2.10).

2.4.4.1 Presentation of the experiment

Optical setup As we indicated earlier in this chapter, we are interested in the measurement of the distance between two incoherent point sources. We have also showed that the intensity profile of a point-like source could be approximated by the point spread function of the imaging system — and that the point spread function of an imaging system with a circular aperture stop could be well approximated by a Gaussian function. For this reason, instead of point sources, we use two Gaussian beams and will study them in the plane where their waist is minimum. In our analysis, we make the assumption that the information about the centroid of the two beams is already known. For this reason, we need the two beams to have symmetrical and opposite displacements with respect to a reference axis — the optical axis of our setup. Furthermore, to probe different separation regimes, the separation between the two sources should be modular.

In order to produce two symmetrical Gaussian beams, we used the setup described in 2.9. The output of a collimated single-mode fiber can be well approximated by a Gaussian beam. This beam is injected in a Mach-Zehnder-like configuration: a beam splitter splits the incoming beam in two, and the two beams are recombined on a second beam splitter. Before being recombined, the two beams travel in two different arms, where two Dove prisms have been inserted. A Dove prism is a reflective prism which is used to invert an image using total internal reflection. The two Dove prisms are oriented at 90° from each other. In this configuration,

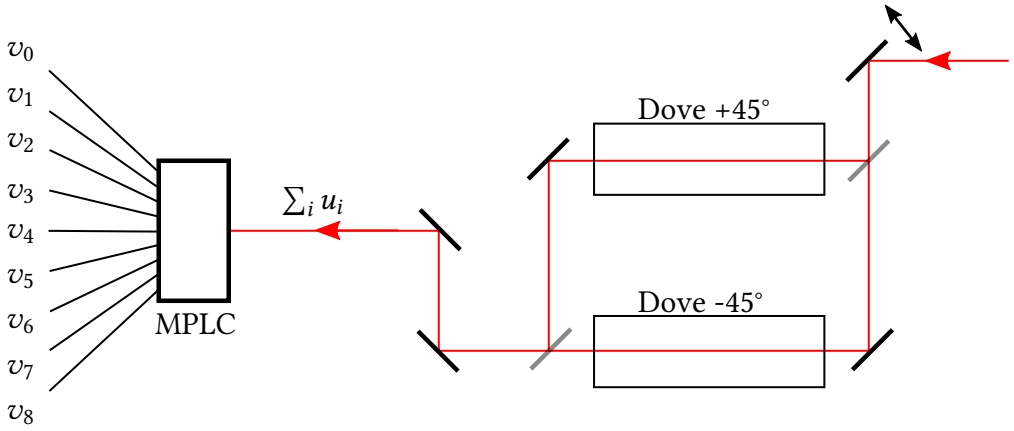

Figure 2.9: Experimental setup. The input beam goes through a Mach-Zehnder interferometer configuration. In both arms of the setup, Dove prisms, orientated respectively at +45° and -45°, allow to make the displacement of the beams in both arms symmetrical with respect to the optical axis. The recombined beams are then injected using two adjustment mirrors into the MPLC system. The sum of the two beams which enters the MPLC system can be described as a superposition of modes $u_{0\leq i\leq 8}$. Each input mode is demultiplexed and sent to a spatially separated spatial mode $v_{0\leq i\leq 8}$. Each of these modes is coupled into a separate single-mode fiber.

if the input beam is perfectly aligned with the optical axis, the two output beams overlap completely at the output. If the input beam is displaced in the transverse plane, the two beams which recombine on the output beam splitter have opposite displacements (see figure 2.10).

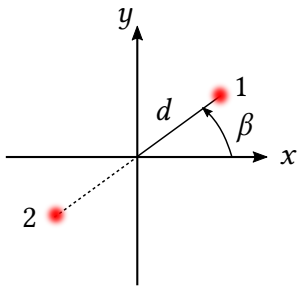

Figure 2.10: The beams 1 and 2 have opposite displacements in the x-y plane. Their position is given by the absolute displacement d and the angle β of the displacement with respect to the x axis.

This configuration is described in figures 2.9 and 2.11. The displacement of the input beam is realized using a translation stage. The mirror which is used to inject the input beam in the first beam splitter is mounted on a micro-meter precision translation stage. The two beams which are recombined thus have Gaussian profiles and symmetrical tunable displacements with respect to the optical axis. We used a fibered SLD (Thorlabs S5FC1005S, 50 nm bandwidth) at a wavelength of $\lambda = 1500$ nm as a source. The coherence length of the light is $L = c/(\pi\Delta\nu) = 15.3\mu$m. The length of each arm of the Mach-Zehnder-like setup is of the order of 30 cm. We can thus consider the two recombining beams to be incoherent. The waist of the beam exiting the Mach-Zehnder-like setup is situated shortly before the output beam splitter. In order to be able to access that plane, we constructed a telescope after the output beam splitter (which is not represented in figure 2.9). This telescope also allows for the scaling of the waist size. Indeed, the waist size of the input beam is $w_0 = 1\,135\mu$m, while the MPLC system we use is designed for a waist of $w_0 = 227\mu$m. The magnification factor of the telescope is equal to 1/5.

This configuration allows not only to study the case of the separation of two incoherent point sources, but also that of one displaced point source. Indeed, by blocking the light in one of the arms of the setup, we recover a single displaced beam. We will use this feature to compare the results of distance measurement for one or two point sources.

The output beamsplitter possesses a second port, indicated by the letter B in figure 2.11. The beams which exit both output ports are identical. We inserted a

quadrant detector (Thorlabs PDQ30C) after this second port. This instrument allows to assess the displacement of a single beam with very good precision (even though we proved earlier in this chapter this it did not perform an optimal measurement). This second measurement will provide us with a comparison for the displacement measurement we perform with through modal analysis.

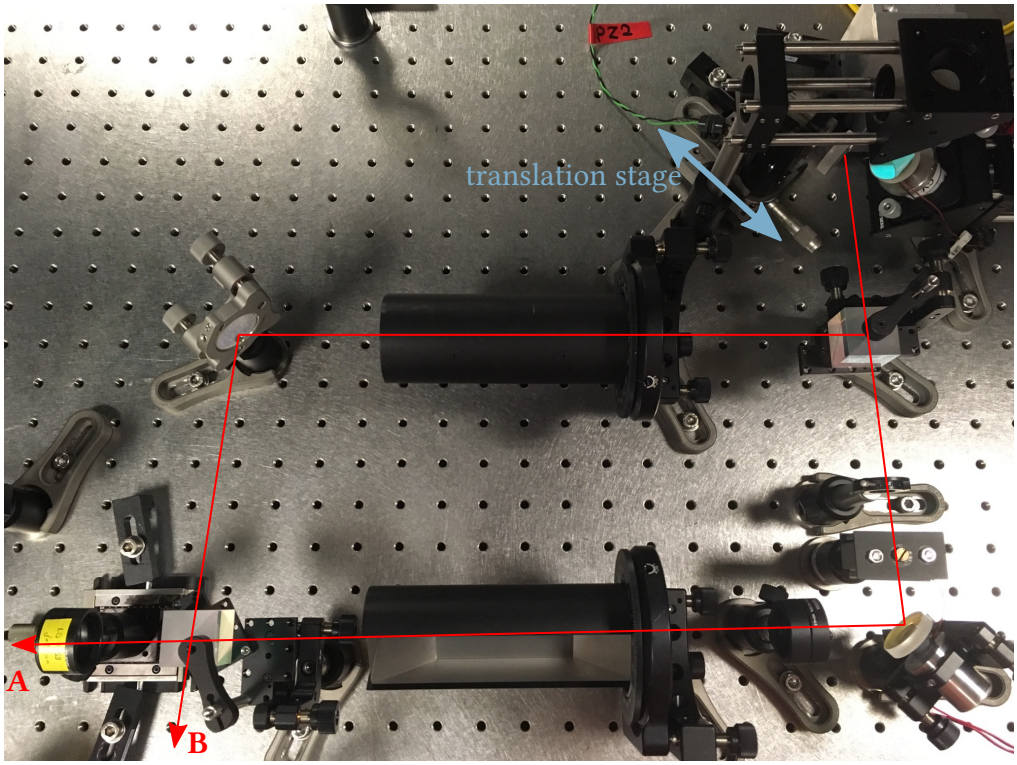

Figure 2.11: Experimental setup. Light at output A goes to the MPLC demultiplexer. Light at output B goes to the 4-quadrant detector.

MPLC system We have presented in the previous sections that modal analysis is an optimal measurement scheme for the estimation of the distance between two incoherent point sources. This modal analysis will be experimentally realized through the use of a MPLC system as a demultiplexing system. The left part of figure 2.9 illustrates how the decomposition of a displaced mode over a Hermite-Gauss mode basis is shaped by the MPLC system into the modes of a collection of single-mode fibers. Using a Taylor series, we have shown that up to the first order in displacement a beam of gaussian profile could be described as a sum of the HG_{00} and HG_{10}

modes. An expansion at higher orders brings into play the higher-order elements of the Hermite-Gauss mode basis. Since this mode basis is complete, a displaced beam can always be described as a sum of Hermite-Gauss modes, whatever the displacement value. This decomposition is described in figure 2.8 and is also valid for the case of two incoherent point sources. Although this analysis is valid even for large displacements or separations, we will stay in a small displacement regime so as to limit the number of modes necessary to describe the displaced beam. Each of these Hermite-Gauss mode is independently shaped into spatially separated Gaussian modes which are coupled to single-mode fibers.

The measurement we perform in this experiment is the intensity measurement of the output of each of these single-mode fibers. For a perfect MPLC system, this amounts to the measurement of the intensity of the decomposition of a beam onto the Hermite-Gauss mode basis. In figure 2.8 the $c_i(d)$ coefficients are amplitude coefficients. The quantity we have access to in the experimental configuration is $|c_i(d)|^2$.

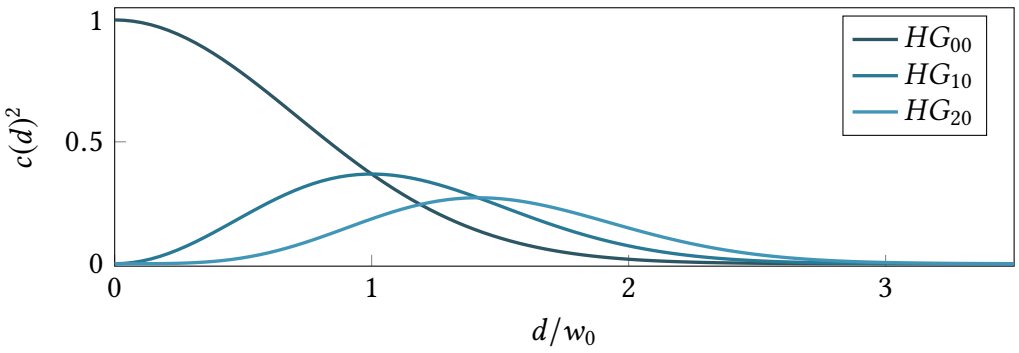

Figure 2.12: Theoretical projections of a displaced beam with $\beta = 0$.

Figures 2.12 and 2.13 give examples of the intensity projections of a displaced beam on the first 9 modes of the Hermite-Gaussian mode basis as a function of the displacement for two displacement directions indicated by the angle β which the displacement makes with the x axis.

2.4.4.2 MPLC system

The MPLC system we designed was a 9 mode demultiplexer (see figure 2.9). The input modes are free-space Hermite-Gauss modes ($u_{0\leq k\leq 8} = HG_{0\leq i,j\leq 2}$). The output modes are spatially separated and coupled into single mode fibers ($v_{0\leq k\leq 8}$). The

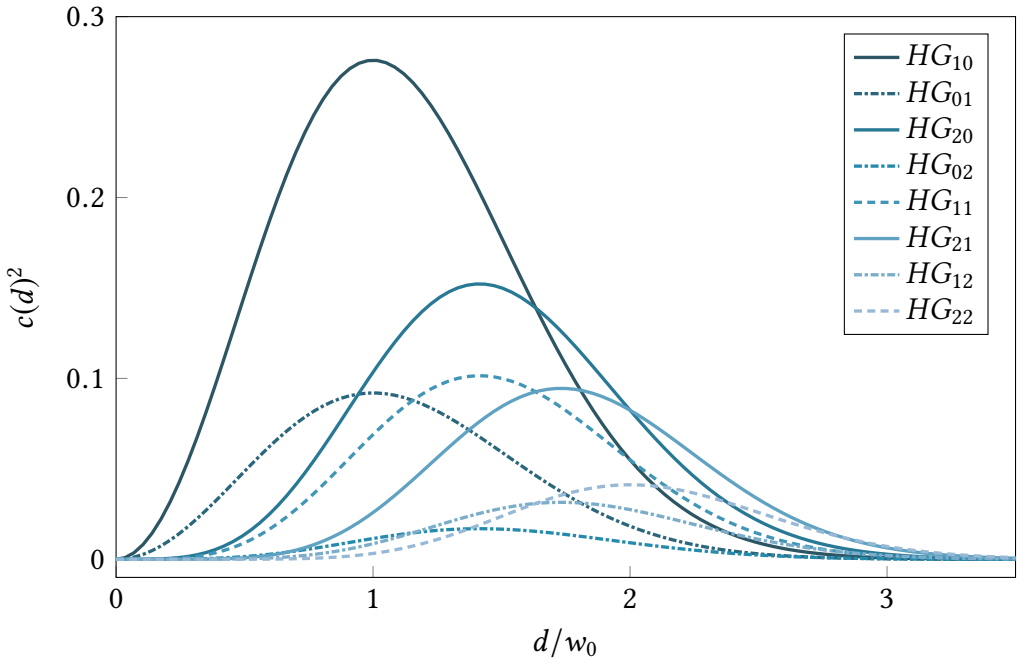

Figure 2.13: Theoretical projections for a displaced beam with $\beta = \pi/6$.

waist of the input modes is $w_0 = 227\,\mu$m. The convergence of the MPLC construction algorithm is not perfect: table 2.1 gives the overlap of the modes produced by the calculated phase plate with respect to the theoretical modes which we aim at.

Table 2.1: Overlap of the modes produced by the algorithm with the theoretical modes.

HG_{00}	HG_{10}	HG_{01}	HG_{20}	HG_{02}	HG_{11}	HG_{21}	HG_{12}	HG_{22}
0.933	0.949	0.935	0.955	0.954	0.928	0.932	0.925	0.937

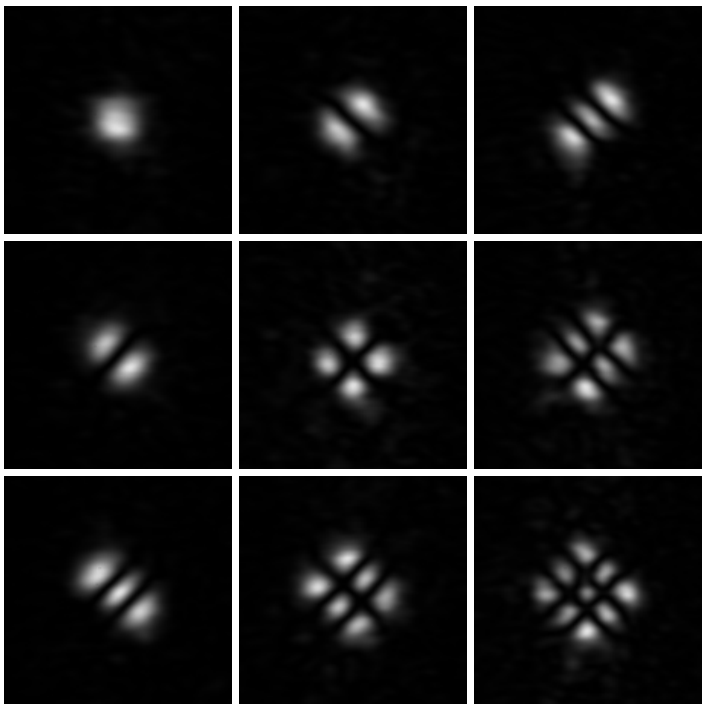

Figure 2.14: Capture of the input modes of the MPLC system.

During the fabrication of the MPLC system, the spatial profiles of the produced modes are recorded. Camera images of such profiles are plotted in figure 2.14. We can assess the quality of the fabricated system by assessing qualitatively that the produced profiles are very close to the theoretical Hermite-Gauss modes.

Table 2.2: Theoretical values of the crosstalk values from mode HG_{00}.

Mode	HG_{10}	HG_{01}	HG_{20}	HG_{02}	HG_{11}	HG_{21}	HG_{12}	HG_{22}
dB	-26.29	-27.45	-26.86	-26.40	-45	-31.06	-38.39	-24.33

We also measure the crosstalk between spatial modes introduced by the system. To do so, light is injected in one input mode, and the intensity in all of the output modes is recorded separately. These values are presented in table 2.2. These measurements are taken during the construction of the MPLC system. In this case, we are only interested in the crosstalk from mode HG_{00}: indeed, when the system is perfectly aligned, we want to assess how much energy leaks to the higher-order modes — since the intensity on the higher-order modes is what we measure to access the displacement. This crosstalk will play an important role in the minimum displacement which can be measured. The optical power measured in the mode HG_{nm} is P_{nm} and the crosstalk of mode (n, m) is given by $10 * \log_{10}\left(P_{n,m}/P_{0,0}\right)$.

2.4.4.3 Characterization of the system

The fabricated MPLC systems' performances do not exactly match with the theoretical ones. The differences find their origins in the imperfections of the optical components, the errors in the fiber arrays' tolerancing, the fabrication errors or the incorrect positioning and alignment of the optical components. For this reason, the numbers presented in tables 2.1 and 2.2 do not give a sufficient account of a MPLC system performances.

A first important point is that the different channels of the MPLC do not have identical losses. This discrepancy in losses between channels is usually probed using two identical MPLC systems in a "back to back" configuration: the output of one system is injected into the second one, and the measured losses correspond to the sum of the losses associated with the multiplexer and the demultiplexer for one channel. However, in this case we only possessed one system of its kind. The procedure we used was the following one: instead of injecting light in the free-space modes, we injected light in the different fibered outputs and measured the intensity exiting the MPLC. Table 2.3 gives the measured *intrinsic* efficiency coefficients of the MPLC system used in the experiment. These coefficients are calculated with mode HG_{00} as a reference. It should be noted that this measurement procedure is not perfect because we cannot discriminate the crosstalk signal from the signal we want to measure.

Table 2.3: Intrinsic efficiency coefficients of the MPLC system relative to mode HG_{00}.

g_{00}	g_{10}	g_{01}	g_{20}	g_{02}	g_{11}	g_{21}	g_{12}	g_{22}
1.00	0.89	1.01	0.94	1.09	0.74	0.79	1.00	0.71

The alignment of the MPLC system with the rest of the experimental setup is another critical point. Indeed, we wish to measure a displacement with respect to a reference position: this reference position is defined by the modes for which the MPLC is designed. We need to match precisely the waist size and position of the beam (in the z or propagation direction), as well as its direction (position in the (x, y) plane and angle). The system we use is precisely designed to be sensitive to any deviation from the design parameters. We make the assumption that the crosstalk values measured during the system fabrication (see table 2.2) are the best ones achievable and take them as our reference. The coarse alignment of the system can be made by maximizing the intensity measured on the HG_{00} mode. When a global maximum is found, it is necessary to monitor and minimize the intensities in the higher-order modes. The minimization of the measured intensity in the higher order modes should coincide with an increase in the intensity measured in HG_{00}, the most important modes to consider being the lowest order ones. The HG_{01} and HG_{10} modes give a measure of the misalignment of the input beam with respect to the reference axis. The HG_{02} and HG_{20} modes give a measure of the waist size and position discrepancies with respect to the reference. The quality of the alignment is thus assessed by how close to the theoretical crosstalk values the measured crosstalk values are.

In the context of our experiment, the reference crosstalk values of table 2.2 will give a bound to the minimum displacement we can access. Indeed, the Cramér-Rao bound gives a minimum displacement d^{CR} and the crosstalk values give another value d^{XT}. At high power, d^{CR} can be made to be arbitrary small (since it is inversely proportional to the square root of the number of photons). In this experiment, $d^{CR} < d^{XT}$. In a stationary regime, the crosstalk values give the minimum displacement we can measure. We suppose, as is the case in this experiment, that the minimum crosstalk intensities are greater than the minimum sensibility of the powermeter we use. The minimum displacement we can measure with the reference theoretical

crosstalk values are the following:

$$d^{XT}_{min,10}/w_0 = \sqrt{10^{-26.29/10}} = 4.85\,10^{-2}$$
$$d^{XT}_{min,01}/w_0 = \sqrt{10^{-27.45/10}} = 4.24\,10^{-2}.$$

These values give an indication of the minimum displacement which can be measured, provided the intensity or photon number is high enough so that the sensitivity of the measurement is small enough.

2.4.4.4 Measurement procedure

For a run of measurement, we first measured the crosstalk reference of the alignment. As we explained, the quality of this alignment is assessed by how close to the reference crosstalk values we are. Table 2.4 presents the crosstalk values of the reference alignment and of every measurement run. We can first note that the values for the HG_{02} mode). We think that this is due to a slight ellipticity of the beam: in one of the directions, the waist size does not correspond to the theoretical waist size the MPLC system is designed for. Thus, even in the correct position, the beam does not perfectly overlap with the HG_{00} mode and some energy is projected on the HG_{0j} modes. However, the quality of the alignments is satisfactory.

Table 2.4: Crosstalk values (in dB).

	HG_{10}	HG_{01}	HG_{20}	HG_{02}	HG_{11}	HG_{21}	HG_{12}	HG_{22}
reference	-26.29	-27.45	-26.86	-26.40	-45	-31.06	-38.39	-24.33
long scan	-25.69	-25.83	-25.50	-22.53	-37.63	-30.81	-37.92	-27.78
short scan 1+2	-26.06	-26.50	-25.51	-22.66	-37.40	-31.64	-38.59	-27.74
short scan 1	-26.65	-29.25	-25.27	-22.75	-37.22	-31.62	-37.54	-27.72
short scan 2	-26.34	-24.90	-25.59	-22.52	-35.77	-31.15	-37.18	-27.52

After the alignment step, a measurement run is performed by introducing displacement in steps using the translation stage. For each step, we record the intensity of all the output modes. In order to have another measure of the introduced displacement, a quadrant detector was installed. We also record the signal of the quadrant detector for each step. The data for both detectors was three-fold: one set of intensities for just one arm of the interferometer, another set for the second arm, and a third for the superposition of both arms. In the remaining of this section we will refer to those three sets as 1, 2 and 1 + 2.

2.4.5 Experimental results and analysis

2.4.5.1 Measurements

We present here the results for two different scans: a short scan ($d \in [0, 0.4w_0]$) and a long scan ($d \in [0, 3w_0]$). The quality of the alignment of both scans is assessed by comparing their crosstalk values to the reference ones (see table 2.4).

In the case of the short scan, we show both values for the "two incoherent sources" situation, that is the measurements performed with both arms of the interferometer (both arms of the interferometer on (1+2)) — but we also show the measurement for each arm of the interferometer separately (values were recorded with one arm on (1) or the other one on (2)). We plot the values of the normalized intensities measured in all the modes in figures 2.16, 2.17 and 2.18.

2.4.5.2 Analysis

First order analysis for displacement measurement At the first order, in the reference basis of the Hermite-Gaussian mode basis, we give the following expression of a displaced gaussian beam:

$$HG_{00}(d\cos\beta, d\sin\beta) = HG_{00}(0,0) + \frac{d\cos\beta}{w_0} HG_{10}(0,0) + \frac{d\sin\beta}{w_0} HG_{01}(0,0) \quad (2.86)$$

with $d \ll w_0$ and β giving the direction of the displacement. The measurements we perform are intensity measurements. For small displacements, the intensities which are detected in the HG_{10} and HG_{01} modes, I_{10} and I_{01}, are proportional to $d^2\cos^2\beta$ and $d^2\sin^2\beta$. Thus, we can extract the measured values of d and β:

$$d_m = w_0\sqrt{I_{01} + I_{10}} \quad (2.87)$$

$$\beta_m = \arctan\sqrt{\frac{I_{01}}{I_{10}}}. \quad (2.88)$$

This analysis is valid for $d \ll w_0$ but does not use all the information that the MPLC system provides. It is nonetheless useful to gain a quick estimation of the displacement (again in the case of $d \ll w_0$).

Numerical fit We can also numerically compute the projections on each mode as a function of d and β: $I_{nm}^{theo}(d,\beta) = \int HG_{00}(d\cos\beta, d\sin\beta)\, HG_{nm}(0,0)\, dxdy$. The translation stage has micrometer-precision screw: we thus have access to a theoretical displacement d_{theo}. We can compare this value to the measured displacement.

To do so, we introduce the parameter b such that $d_m = b \times d_{theo}$. This proportionnality constant will be used to adjust the fits between the functions I_{nm}^{theo} and the measured intensities. The 4-quadrant detector also gives an information on the actual displacement that we can compare to d_m. We also introduce mode-dependent gain parameters g_{nm} which will be used to take into account the mode-dependent losses of the system. The intensity measured in mode HG_{nm} will be fitted with two different numerical models: $I_{nm}^{theo}(d, \beta)$ and $g_{nm} \times I_{nm}^{theo}(d, \beta)$. We used this numerical model to fit the measurements we present in figures 2.16, 2.17, 2.18 and 2.19. The fitting parameters we used for the short scan are presented in table 2.5, while the gains we used in the fitting operation are given in table 2.6.

2.4.5.3 Results and limitations

Figure 2.15 gives the comparison between the displacement measured by the quadrant detector and the displacement measured by the MPLC system. In the x direction, the measurements of both techniques are in good agreement for both arms. There seems to be a slight displacement difference between the two arms. In the y direction, the displacement measured by the quadrant detector is equal for both arms. For small displacements, the displacement measured by the MPLC for arm 2 seems to show a small error — at large displacement, the displacements measured for both arms becomes equal. In this direction we observe a small discrepancy between the amount of displacement measured by one technique with respect to the other. Finally, we plot the total measured displacement for both arms and both techniques and compare it with the theoretical displacement (given by the translation stage). First, the displacement measured by the quadrant detector for both arms is very close. It nonetheless shows a discrepancy in slope with the reference value. The displacement measured by the MPLC system for arm (1) — which had the best alignment crosstalk values — agrees remarkably well with the reference value for very small displacements. We observe that for displacement greater than $0.3\, w_0$, the estimation of the displacement is not as accurate anymore. This is consistent with the fact that outside of the range $d/w_0 \ll 1$, higher orders of the Taylor expansion need to be taken into account. The agreement is not as good with arm (2) of the interferometer. As we indicated, the alignment of this arm with the MPLC was not as good (see table 2.4 and in particular the value for HG_{01}) which gives an explanation to this poorer performance. Finally, we note that overall, the measurements made with both techniques are in good agreement, which validates the MPLC as a distance measurement device.

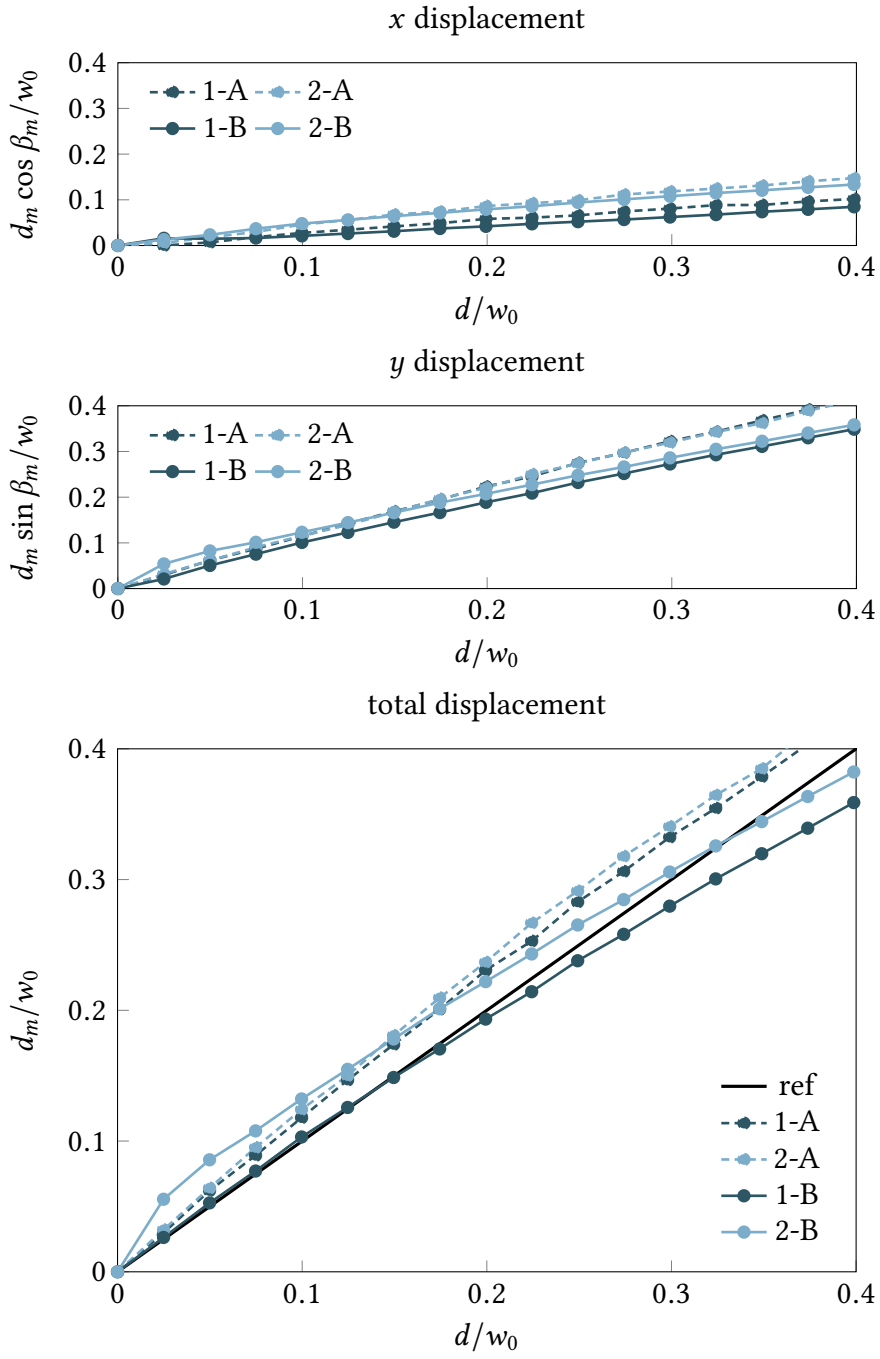

Figure 2.15: *Short scan* — Displacement measurement using (A) a 4-quadrant detector and (B) a spatial mode demultiplexer for both arms of the Mach-Zehnder interferometer (1, 2).

In figure 2.16, we plot the normalized recorded intensities in each of the output modes with both arms of the interferometer on. Figures 2.17 and 2.18 plot the same curves for arms (1) and (2) of the interferometer. The first observation we make is that as we stated, the measurement of the displacement of one source is equivalent to the measurement of the distance between two incoherent sources — the curves for these two measurements are very similar. One should note that the scales are not the same for every mode. On these curves, we also plot two theoretical fitting curves. The first one, dashed, shows the theoretical projections in each mode using the parameters we derived in table 2.5. The second one, dash-dotted, uses these same parameters but adds an additional degree of liberty in the form of a free gain parameter g_{nm}. These fitting parameters are displayed in table 2.6. For the modes on which we measure significant amounts of energy, the theoretical curves fit nicely with the data, and the gain parameters are not too far from unity. As the intensity measured in a mode decreases, the quality of the fit drops.

Table 2.5: Displacement parameters — short scan.

	β	b	relative intensity
1+2	1.095	0.974	1
1	1.007	0.925	0.47
2	1.138	1.0150	0.53

Table 2.6: Gain fit parameters — short scan.

	g_{00}	g_{10}	g_{01}	g_{20}	g_{02}	g_{11}	g_{21}	g_{12}	g_{22}
1+2	0.966	1.004	1.155	1.029	1.428	0.310	0.904	1.703	0.021
1	0.973	0.869	0.989	1.127	1.443	0.245	0.660	1.829	0.018
2	0.959	1.011	0.805	0.994	1.505	0.363	1.067	1.549	0.021

In figure 2.19, we again plot the normalized recorded intensities in each of the output modes with both arms of the interferometer on for the "long" scan. The theoretical curves are computed in the same way as those of the short scan. The parameters used are shown in table 2.7. We also include figure 2.20 in which we plot all the normalized intensity measurements on a single plot. We can clearly see how the different modes come into play as the value of the displacement increases.

The fit parameters for the long scan are $\beta = 1.055$ and $b = 0.925$.

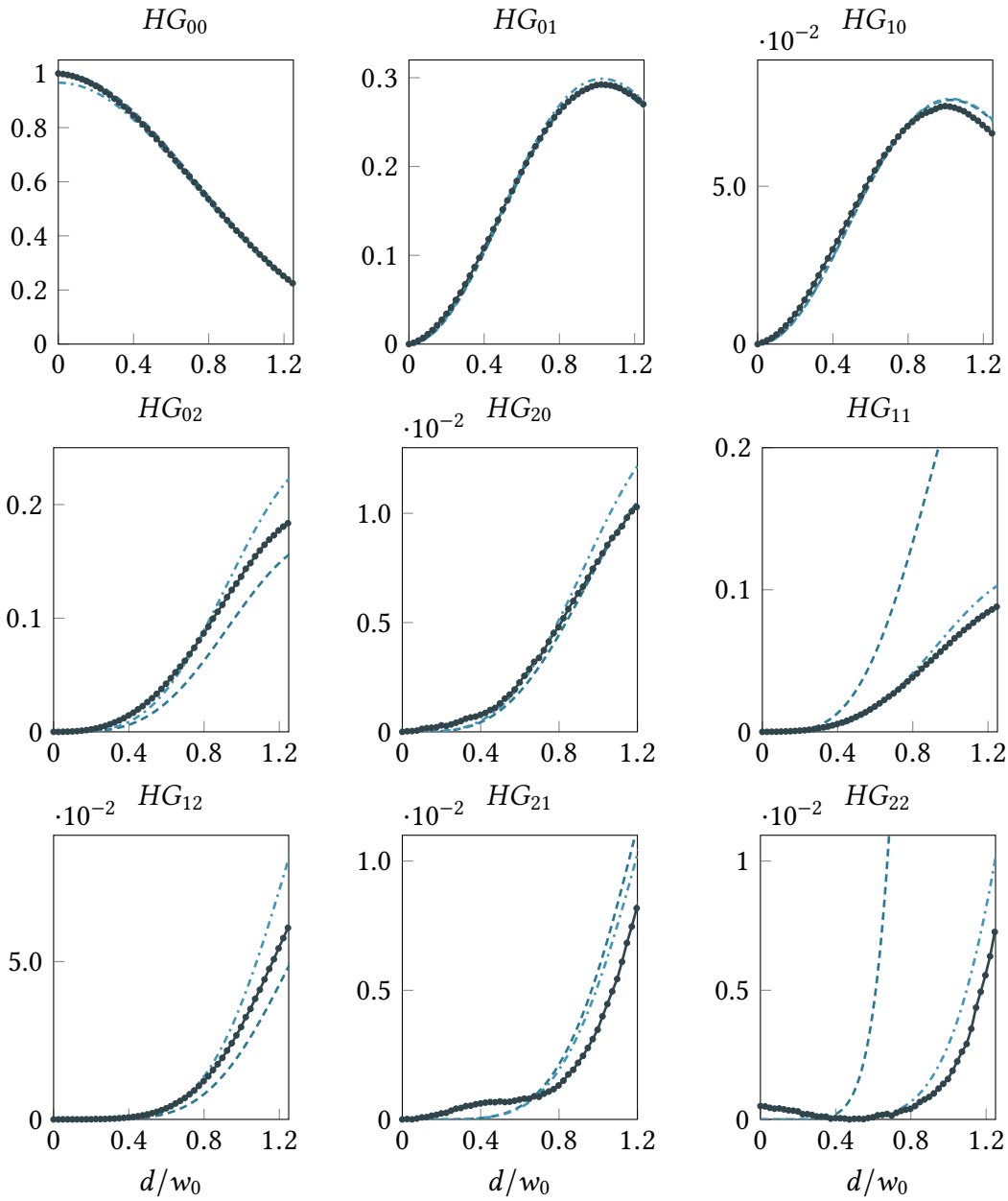

Figure 2.16: *Short scan* — Normalized intensity measured on each of the MPLC output (solid line) with both arms of the interferometer. In dashed line is the theoretical model for the extracted parameter of displacement and angle. In dashed dotted, we fit the data with an additional gain which depends on the channel.

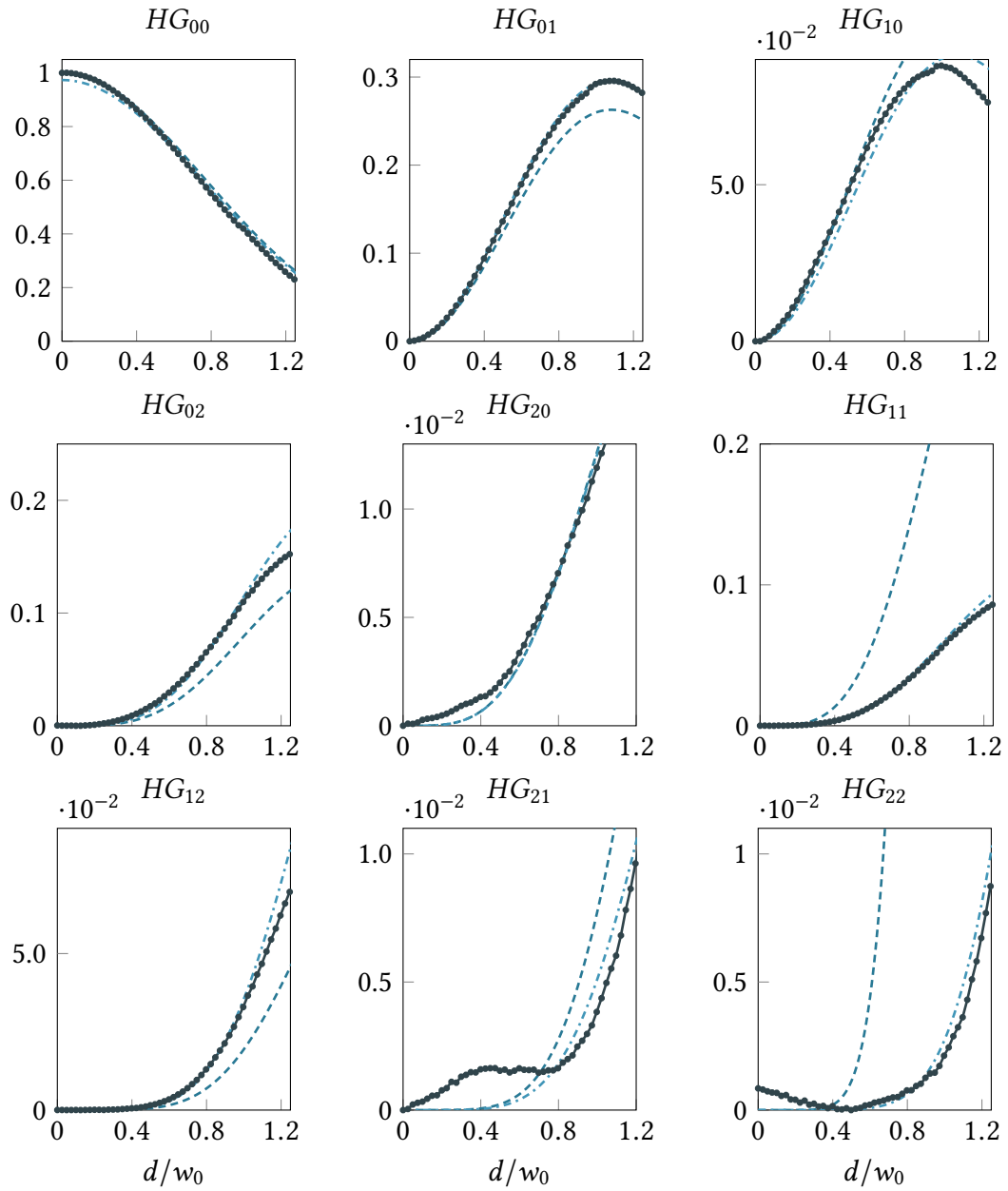

Figure 2.17: *Short scan* — Normalized intensity measured on each of the MPLC output (solid line) with arm 1 of the interferometer. In dashed line is the theoretical model for the extracted parameter of displacement and angle. In dashed dotted, we fit the data with an additional gain which depends on the channel.

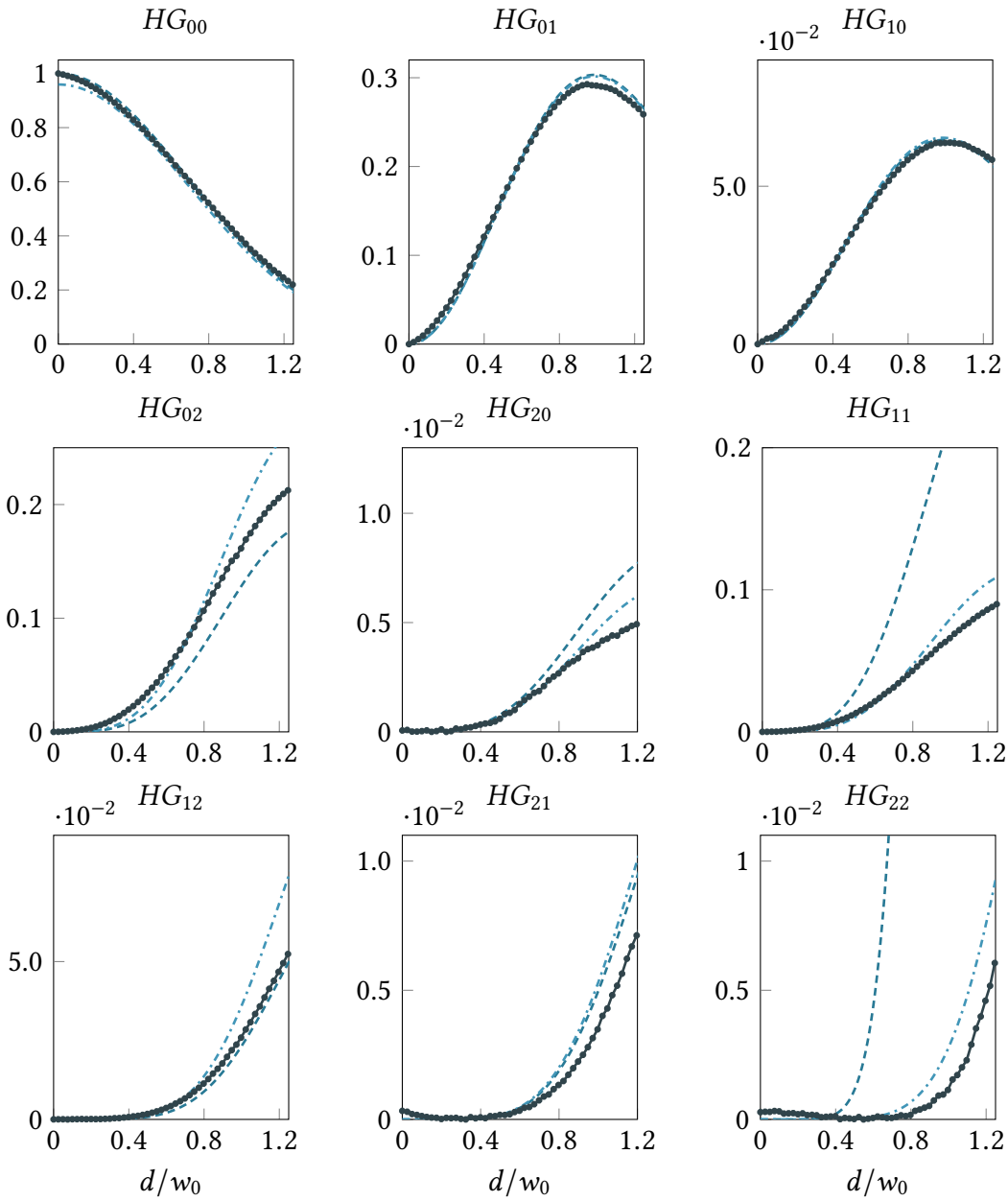

Figure 2.18: *Short scan* — Normalized intensity measured on each of the MPLC output (solid line) with arm 2 of the interferometer. In dashed line is the theoretical model for the extracted parameter of displacement and angle. In dashed dotted, we fit the data with an additional gain which depends on the channel.

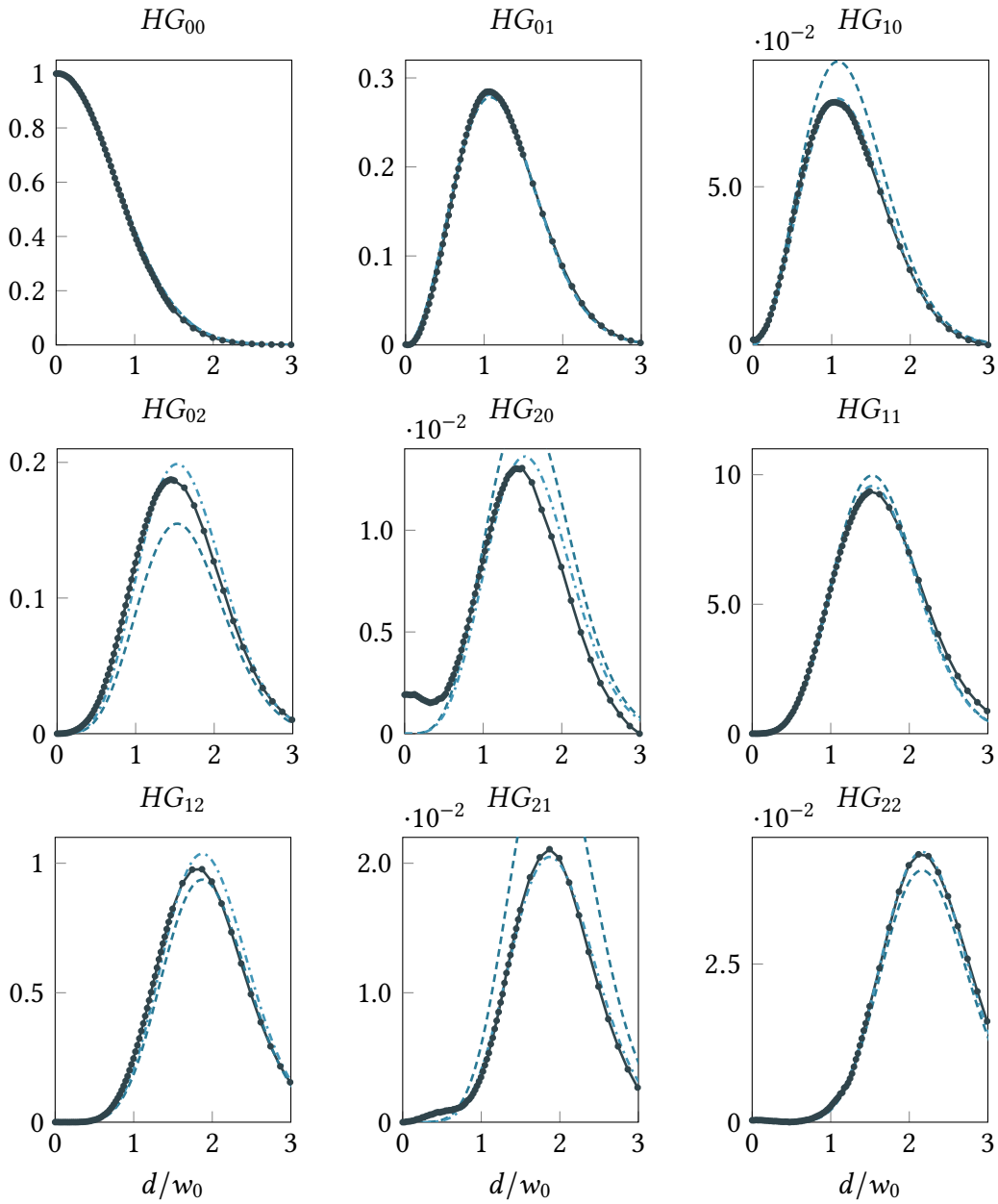

Figure 2.19: *Long scan* - Normalized intensity measured on each of the MPLC output (solid line) with both arms of the interferometer. In dashed line is the theoretical model for the extracted parameter of displacement and angle. In dashed dotted, we fit the data with an additional gain which depends on the channel.

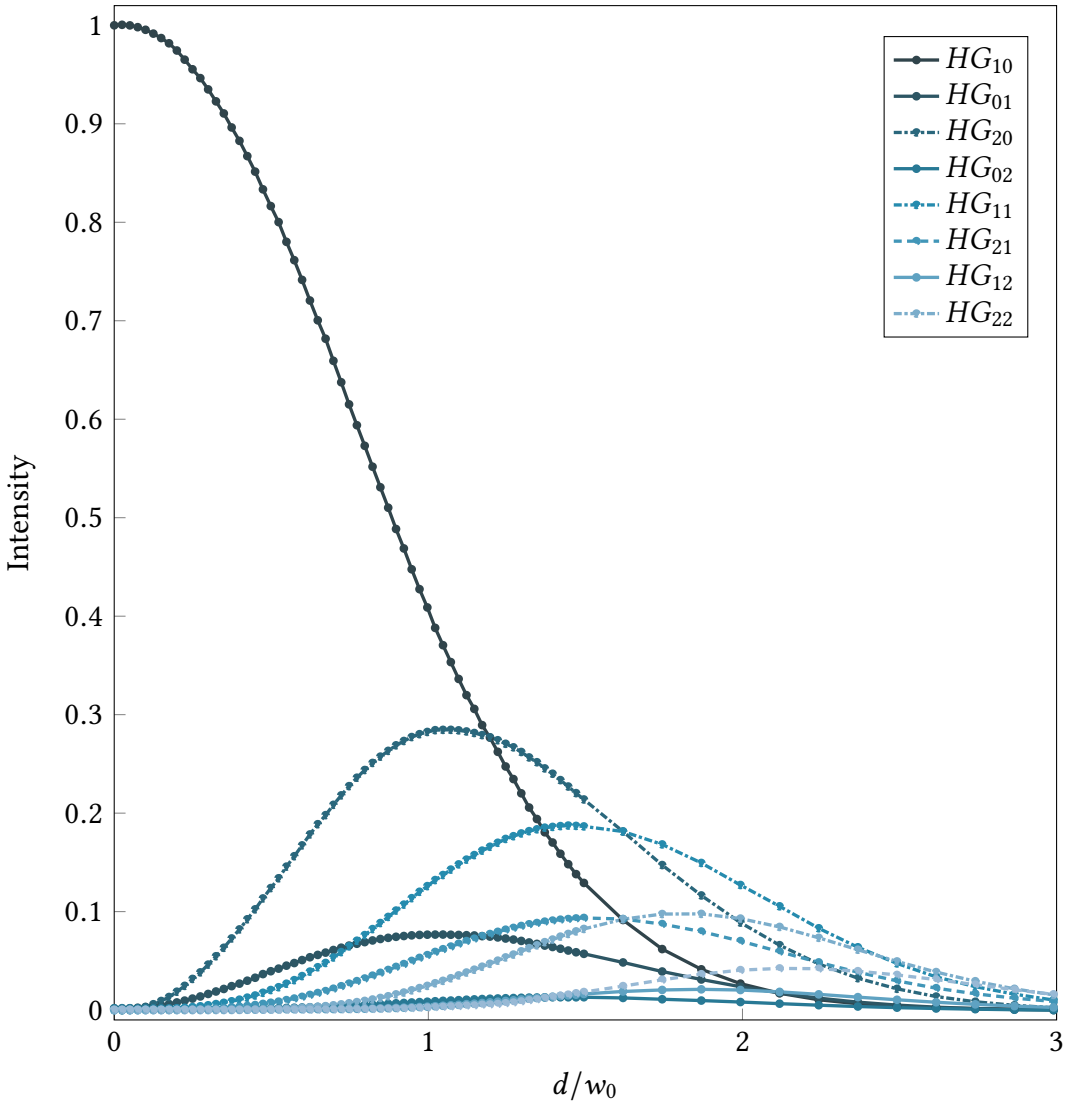

Figure 2.20: *Long scan* - Experimental projections of displaced beam.

Table 2.7: Gain fit parameters - long scan.

g_{00}	g_{10}	g_{01}	g_{20}	g_{02}	g_{11}	g_{21}	g_{12}	g_{22}
1	1.012	0.869	1.284	0.848	0.959	1.107	0.679	1.073

2.4.5.4 New measurements and outlooks

The results we present are encouraging but call for some supplementary measurements which we did not have time to make before the end of this Ph.D. The minimum signal we can measure with our apparatus should be quantified and compared to the Cramér-Rao bound. One should also characterize the noise of the source we used.

These experimental results were taken at intensities of the order of the μW. An interesting regime — which we did not have time to explore — in which to test this setup is at low photon numbers. Indeed, as the Cramér-Rao bound states, at low photon numbers, the sensitivity of distance measurement decreases. It is in this regime that we hope to measure a significant performance increase for the spatial demultiplexing system compared to other measurement schemes. The physical implementation of our setup makes it readily compatible for probing low intensity regimes: as the output of the MPLC are fibered, they can be directly connected to avalanche photodiodes - which can count photons individually - whilst avoiding the loss of photons, or the contamination by stray photons.

The setup we present is also compatible with the measurement of correlations between the different modes.

2.5 A small detour through compressed sensing

We make in this final section a small digression toward the field of compressed sensing. We wish to introduce the main tools and most important results of this new and exciting field. Compressed sensing has introduced a paradigm shift in the way one can think about information retrieval (see for instance [Bobin 08] for the application of this theory to astronomy). Since the limits of classical imaging were derived from information theory, compressed sensing theory brings new elements to the imaging field and calls into question its limitations . The link which can be made between the MPLC system and compressed sensing will be clarified in the conclusion of this chapter.

2.5.1 Quick overview

In a general context, data acquisition goes as follows: massive amounts of information are collected (as an example, good quality smartphones record images of more than 2 millions pixels while the number of pixels on good quality cameras easily exceeds 15 millions) to record as much information as possible. More than often,

much of this information is superfluous. In any case, the amount of recorded data is too formidable to be stored or processed as such, and these data must first go through an essential compression stage. The many compression algorithms which have emerged take advantage of the redundancy of the recorded signal to diminish the amount of capacity storage needed. The redundancy of the signal expresses the fact that in some well chosen basis the signal is *sparse*. This means that the signal, represented in this basis, only has a few significant coefficients. Sparse approximation algorithms make use of this and only record the location and value of the highest coefficients. Many well known formats (JPEG, MPEG, MP3) use this process, often taking into account the limited human perception abilities to fix the minimum thresholds. However useful and successful, this process is thus inherently wasteful since information is recorded before being discarded.

Compressed sensing takes the opposite view by aiming at acquiring only the useful information. In other words, the recorded data should *already* be compressed. The central idea of compressed sensing is the following: while the Nyquist-Shannon sampling theorem states that a minimum given number of samples n should be recorded in order to perfectly capture a given bandlimited signal, if we know that in a given basis this signal is sparse (meaning represented by $k \ll n$ non-zero coefficients), we can greatly reduce the number of measurements that need to be stored. We want to directly "sense" the signal in a compressed form. A direct consequence of this approach is to greatly diminish the number of sensors needed for the same acquisition quality.

The field of compressed sensing originates in the works of Candès, Romberg, Tao and Donoho [Candès 06a, Candès 06b, Candès 06c, Candès 06d, Donoho 06]. Very good introductions to the subject can be found in [Davenport 12, Candès 07b], from which the following section is largely inspired.

2.5.1.1 Useful concepts

We will detail in this section some of the mathematical concepts which are at the core of compressed sensing.

Sparsity A signal $\boldsymbol{x}$ expressed in basis Υ is sparse if its representation in a mode basis Ψ possesses at most k non zero coefficients — the signal is said to be k-sparse. The signal in itself may not be sparse in its representation basis Υ as such but it admits a sparse representation in some basis Ψ. If M is the basis change matrix between basis Ψ and Υ then $\boldsymbol{x} = M\boldsymbol{c}$ where the vector $\boldsymbol{c}$ only has k non-zero coefficients. Most natural images have large portions which are uniform with a small

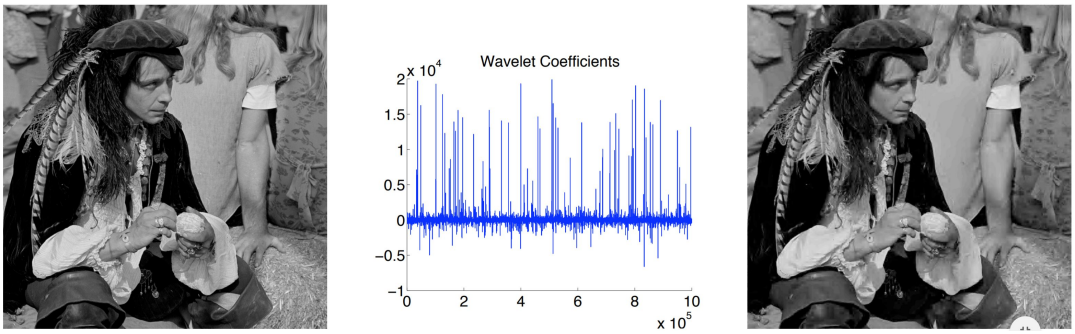

Figure 2.21: The image on the left displays a 1 megapixel image and its wavelet transform coefficients (displayed in a randomized order). One can see that a few wavelet coefficients make the most important contributions. The rightmost image shows the image reconstructed by keeping only the 25 000 largest coefficients and putting all the other coefficients to 0 (that is, by discarding 97.5% of the original information. One can see that the changes are not easily noticeable. This procedure underlies most modern compression algorithms such as JPEG-2000. (Figure taken from [Candès 07b].)

number of edges between them. These types of images are *almost* sparse. A tool which is particularly useful for natural images compression, and more broadly for sparse representations, is wavelet analysis (see figure 2.21 for an illustration): in wavelet basis, natural images are often sparse. An excellent reference on the subject is [Mallat 08].

Incoherence The second concept at the core of compressed sensing is incoherence and relates to the sensing modality. It broadens the concept of time-frequency duality: a signal which is sparse in frequency (for instance a signal which has a Dirac function representation) is spread out in the time domain and vice versa. We define Ψ as the sparse representation basis and Φ as the measurement basis. Incoherence expresses and quantifies the fact that the sensing basis Φ must have a dense representation on the basis Ψ on which the object is sparse. The measure of the coherence between mode basis Ψ and Φ of $\mathbb{R}^n$ is defined as [Candès 07b]

$$\mu(\Psi, \Phi) = \sqrt{n} \max_{1 \leq, k, l \leq n} |\langle \boldsymbol{\phi}_k, \boldsymbol{\psi}_l \rangle| . \tag{2.89}$$

If bases Ψ and Φ have correlated vectors, $\mu(\Psi, \Phi)$ is large. On the contrary, if Ψ and Φ are uncorrelated, it is small. We note that $\mu(\Psi, \Phi) \in \left[1, \sqrt{n}\right]$. Also, random

unitary matrices are largely incoherent to most bases: if Φ is a random orthonormal basis, then $\mu(\Psi, \Phi) \simeq \sqrt{2 \log n}$ with high probability.

2.5.1.2 Main results

We will now present some of the main results of compressed sensing. The question behind compressed sensing is (for a noiseless signal): what is the minimum number of measurements one needs to make to be sure to be able to reconstruct the signal perfectly? For example, in the context of time-varying signals, the well-known Shannon sampling rate gives the minimum sampling rate for a bandlimited signal. But this bound is in fact derived only for one specific type of measurements: in this case measurements which are made at fixed frequency — or equivalently measurements for which samples are taken regularly in time. This type of measurement makes the reconstruction of the measured signal straightforward. On the contrary, compressed sensing factors in all the possible types of measurements. This implies that the reconstruction of the signal is not unequivocal anymore. We will develop in this section three main compressed sensing results. In these results, the choice of the type of measurement is represented by the choice of the measurement matrix. A measurement operation can be modeled by the multiplication of a signal vector $\boldsymbol{x}$ with a measurement matrix. Let $A \in \mathbb{R}^{m,n}$ (with $m < n$) be the measurement or sensing matrix. Its columns are vectors of the measurement basis Φ and $\boldsymbol{y}$ is the measurement vector:

$$\boldsymbol{y} = A\boldsymbol{x}. \tag{2.90}$$

The reconstruction of the signal $\boldsymbol{x}$ can be expressed as an optimization problem: knowing the measurement matrix A and the measurement result $\boldsymbol{y}$, we want to access the signal vector $\boldsymbol{x}$.

We will present three measurement situations which correspond to three different optimization problems.

Sparse noiseless signal recovery We recall that we consider a signal which can be expressed as $\boldsymbol{x} = M\boldsymbol{c}$ with M the mode basis change matrix between basis Υ of $\boldsymbol{x}$ and basis Ψ of $\boldsymbol{c}$, with $\boldsymbol{c}$ a k-sparse vector: only the first k coefficients of $\boldsymbol{c}$ are non-zero. For any vector $\tilde{\boldsymbol{x}} \in \mathbb{R}^n$, we can also define $\tilde{\boldsymbol{c}} \in \mathbb{R}^n$ such that $\tilde{\boldsymbol{x}} = M\tilde{\boldsymbol{c}}$. We define the following optimization problem:

$$\min_{\tilde{\boldsymbol{x}} \in \mathbb{R}^n} \|\tilde{\boldsymbol{x}}\|_{l_1} \qquad \text{subject to} \qquad \boldsymbol{y} = A\tilde{\boldsymbol{x}} \tag{2.91}$$

with $\|\tilde{\boldsymbol{x}}\|_{l_1} = \sum_i |\tilde{x}_i|$. Among all the vector $\tilde{\boldsymbol{x}}$ which are consistant with the measurement vector $\boldsymbol{y}$, we wish to find the vector which has the minimum l_1 norm. This problem is a convex optimization problem. These types of optimization problems have a large body of literature detailing many types of efficient algorithms which allow their resolution.

Theorem 1 The first theorem we wish to present is the following: we suppose once again that $\boldsymbol{x} \in \mathbb{R}^n$ and that its representation $\boldsymbol{c}$ in the basis Ψ is k-sparse. We select a matrix A whose columns are sampled uniformly at random in the measurement basis Φ. Then, if

$$m \geq C\mu^2(\Phi, \Psi)\, k \log n \tag{2.92}$$

where C is a positive constant, the solution to the optimization problem 2.91 is exact with overwhelming probability (see [Candès 07a, Candès 07b]). If a signal is k-sparse, the degree of incoherence between the measurement basis Φ and representation basis Φ sets an inferior bound to the number of measurements m which need to be taken in order to exactly recover the signal. For instance, if the basis Ψ and Φ are perfectly incoherent we have $\mu(\Psi, \Phi) \simeq 1$, which means one needs to acquire as few as $k \log n$ samples (instead of n) to reconstruct the signal. This result can be condensed as such: by sampling a signal in a nonadaptive way (for example using a random sensing matrix) from an incoherent domain, and using a linear optimization algorithm, one can recover a signal perfectly.

Figure 2.22: On the left hand side is the image reconstructed with the 25 000 highest coefficient on the sparse basis, while on the right hand side is the image reconstructed with 96 000 incoherent measurements. The two images are identical, meaning the recovery with a l_1 minimization is perfect. The ratio between the sparsity of the image and the number of incoherent samples is roughly 1/4. (Taken from [Candès 07b].)

Robust compressive sampling The first result we presented is quite strong but does not deal with noise. Yet, as we have seen throughout this chapter, noise plays a major role in any type of measurement scheme. In this problem, noise comes into two forms. First, signals are often not exactly k-sparse: $n-k$ coefficients are close to zero but not exactly zero. We define the signal vector $\boldsymbol{x}^{(k)} = \Psi \boldsymbol{c}^{(k)}$ where $\boldsymbol{c}^{(k)}$ is the vector constituted of the k largest coefficients of $\boldsymbol{c}$. Secondly, during the acquisition process, noise will corrupt the data. A realistic measurement process model is in fact

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{z}, \tag{2.93}$$

the vector $\boldsymbol{z}$ being the noise error term. To continue our introduction, we need to introduce the concept of the isometry constant of a matrix.

Definition For s in $\mathbb{N}$, the isometry constant δ_s of a matrix A is the smallest number such that

$$(1 - \delta_s) \|\boldsymbol{x}\|_{l_2}^2 \leq \|A\boldsymbol{x}\|_{l_2}^2 \leq (1 + \delta_s) \|\boldsymbol{x}\|_{l_2}^2 \tag{2.94}$$

holds for all s-sparse vectors $\boldsymbol{x}$ ($\|\boldsymbol{x}\|_{l_2} = \sqrt{\sum_i |x_i|^2}$). A is said to obey the *restricted isometry property* of order s if δ_s exists and $\delta_s < 1$.

In order to take into account the noise, we examine two new optimization problems. We recall that now place ourselves in the case where the sparse representation $\boldsymbol{c}$ of the signal $\boldsymbol{x}$ is no longer k-sparse but *almost* k-sparse. The first optimization problem we look at is the same as the preceding one:

$$\min_{\tilde{\boldsymbol{x}} \in \mathbb{R}^n} \|\tilde{\boldsymbol{x}}\|_{l_1} \qquad \text{subject to} \qquad \boldsymbol{y} = A\tilde{\boldsymbol{x}}. \tag{2.95}$$

In the case of a signal which is not perfectly sparse, we need to impose a condition on the sensing matrix A.

Theorem 2 If the sensing matrix A admits an isometry constant which verifies $\delta_{2k} < \sqrt{2} - 1$, the solution $\boldsymbol{x}^*$ to problem 2.95 obeys

$$\|\boldsymbol{x}^* - \boldsymbol{x}\|_{l_2} \leq C_0 \|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|_{l_1} / \sqrt{k} \qquad \text{and} \qquad \|\boldsymbol{x}^* - \boldsymbol{x}\|_{l_1} \leq C_0 \|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|_{l_1} \tag{2.96}$$

with C_0 a constant. The distance between the solution $\boldsymbol{x}^*$ and the signal $\boldsymbol{x}$ is can be measured using the l_1 or l_2 norm. In both cases, it is bounded by the distance between the signal and its k-sparse representation measured with the l_1 norm. In

the case of the l_2 norm, this bound is made even tighter by a $1/\sqrt{k}$ factor. A first remark is that if $\boldsymbol{x}$ is indeed k-sparse, then the reconstruction is of the signal is again exact. If not, the reconstruction we perform is as good as if one knew ahead of time which where the largest k coefficients describing $\boldsymbol{x}$ and measured exactly those. The major difference with theorem 1 is that this result is deterministic. Figure 2.22 illustrates this result.

We examine a final optimization problem for noisy data using relaxed constraints for reconstruction defined as

$$\min_{\tilde{\boldsymbol{x}}\in\mathbb{R}^n} \|\tilde{\boldsymbol{x}}\|_{l_1} \qquad \text{subject to} \qquad \|A\tilde{\boldsymbol{x}} - \boldsymbol{y}\|_{l_2} \leq \epsilon \tag{2.97}$$

where ϵ is the upper bound to the level of noise of the data. In this case, we take into account the fact that additive noise is present and relax the constraints on $\tilde{\boldsymbol{x}}$. We again need to impose a condition on the sensing matrix A.

Theorem 3 Assuming that $\delta_{2k} \leq \sqrt{2} - 1$, the solution $\boldsymbol{x}^*$ of 2.97 obeys

$$\|\boldsymbol{x}^* - \boldsymbol{x}\|_{l_2} \leq C_0 \|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|_{l_1}/\sqrt{k} + C_1\epsilon \tag{2.98}$$

where C_0 and C_1 are constants (for example, one can show that if $\delta_{2k} = 1/4, C_0 \leq 5.5$ and $C_1 \leq 6$). This theorem stats that the quality of the reconstruction is bounded by the sum of two terms: the first one is linked to the fact that the signal is not exactly sparse, while the second one is simply proportional to the noise level.

Provided we know can perform a measurement whose matrix representation A obeys the restricted isometry condition, compressed sensing states that we can reconstruct a signal as well as what we could have done knowing in advance the k-largest coefficients of the signal in the basis in which is it (almost) sparse. The reconstruction of the signal is done using optimization algorithms.

Sensing matrices Efficient optimization algorithms are easily available. In the context of compressed sensing, the complexity of the problem of signal retrieval now lies in the design of suitable sensing matrices. Many different strategies can be employed but we will only present one of them for illustration (see [Candès 07b] for more details). If one constructs a sensing matrix using one of the three routines

- sampling n column vectors at random on the unit sphere of $\mathbb{R}^n$.
- sampling independent identically distributed entries from the normal distribution with zero mean and $1/m$ variance.

then with overwhelming probability, the resulting matrix A obeys the restricted isometry property provided that

$$m \geq C\, k \log n/k. \tag{2.99}$$

In short, compressed sensing with random matrices allows to obtain signals which are as good as the best k-term approximation by using only $O(k \log n/k)$ samples. Moreover, the acquisition setup is fixed — it needs not be modified depending on the structure of the signal. As soon as this theory was developed, it found numerous applications. For instance, reducing the number of samples per image allows to speed up the measurement process in magnetic resonance imaging [Lustig 07] allowing for significant improvements for patients. Compressed sensing, by reducing the required number of sensors, also allows for data compression [Duarte 05]. Finally, the principles of compressed sensing can be turned around to design efficient channel coding algorithms [Candès 05].

2.5.2 Compressed sensing for optical resolution

As we described in the previous part, many images satisfy the sparsity constraint, meaning that there exists a basis in which their sparsity k is such that $k \ll n$ where n is the number of pixels used to represent them. It is only natural that compressed sensing should be applied to the field of optical imaging [Gazit 09, Shechtman 10, Willett 11]. The result we introduced in 2.2 can be rewritten in the formalism of compressed sensing: in the prolate spheroidal function basis, the sparsity of the image is given by $k = S \times \pi/2$ where S is the Shannon number of the imaging operator. In [Wang 12], discrete sparse objects are imaged on discrete prolate spheroidal sequences and reconstructed exactly using compressed sensing.

Compressed sensing opens promising perspectives for the improvement of resolution in imaging: for a given number of sensors, the information gathered through compressed sensing is considerably higher than using traditional pixel imaging. In this sense, compressed sensing does provide some kind of super-resolution. The challenge lies on the construction of the sensing basis/measurement apparatus. Indeed, one needs to collect the signal on an appropriate mode basis. This is where the use of a demultiplexer such as the MPLC, which allows to choose any desired measurement basis, should prove to be a most interesting tool.

2.6 Conclusion

We presented in this chapter the demonstration of spatial demultiplexing as a new method for the measurement of the displacement of one source or of the separation between two incoherent sources. The sensitivity of this method saturates the theoretical Cramér-Rao bound, which means it is an efficient estimator for this parameter. The MPLC giving access to several spatial modes, more than one parameter can be estimated. We made the experimental demonstration of a distance measurement in two dimensions. The fact that we can access the higher order modes also means we can maintain optimal sensitivity even outside the regime of very small displacement. The demultiplexing MPLC system was tailored to give the best sensitivity for displacement and separation measurements. We introduced the concept of noise or measurement modes - these modes are given by the derivative of the spatial profile of the field with respect to the parameter we want to estimate. The MPLC system can be built to demultiplex virtually any mode basis. This means that any parameter, in the limit of small variation of this parameter, can be estimated with optimal precision, as long as the spatial profile of the probe beam does depend on this parameter. Probing the efficiency of the MPLC in estimating different types of parameters, with different kinds of illumination could lead to many new promising experiments.

We have also introduced the theory of compressed sensing, which states that signals with a limited amount of information can be estimated with far less coefficients than what is stated by the sampling theorem. In imaging, we have introduced two types of sensing mode basis which can verify the hypothesis required in compressed sensing: random mode bases, or the prolate spheroidal function mode basis. Imaging through random medium has already been demonstrated and has shown very promising results. While the MPLC is in theory capable of demultiplexing a spatial mode basis composed modes with random spatial profiles, it should stand out in the demultiplexing of the prolate spheroidal spatial modes. An interesting prospect is to assess, in an imaging context, the performances of a MPLC system with a reasonable number of modes ($\simeq 50$[17]) for natural images. Such systems may prove useful for multiparameter estimation, or for instance pattern recognition. As of today, the performances of the millions of pixels of available cameras will likely not be exceeded by a MPLC imaging system in a traditional imaging context. However, the use of a MPLC system might prove advantageous in at least two different situations. Camera pixels are limited in the speed at which they operate: the enor-

[17] The current MPLC systems can shape up to 45 modes [Bade 18].

mous amount of pixels complexifies information retrieval. Reducing the number of sensors should allow to operate at higher speeds. Alternatively, in a context where reducing the number of sensors might give an economic advantage (for instance at specific wavelengths where traditional sensors are very expensive — for instance in the infrared), this imaging technique could prove useful.

The two approaches we presented are opposite strategies to information retrieval. In the case of parameter estimation, we tailor the mode basis to the information we want to measure. In this case, a few modes carry all the relevant information. On the contrary, in compressed sensing, we aim at designing a mode basis which is incoherent with the mode basis in which the signal we want to measure is expressed. Both techniques open the way for new exciting applications of the MPLC system.

Chapter 3

Manipulation of the spatial degrees of freedom

"And if the dam breaks open many years too soon [...]
I'll see you on the dark side of the moon."

– Pink Floyd

Contents

This chapter produces a detailed outline of the results presented in [Boucher 17] (see also B). In the previous chapter we have presented how a mode basis change could allow the precise measurement of a spatial property of the electromagnetic field: as the field profile is modified, energy is transferred into higher-order modes in which it can be detected and measured. The central idea of this chapter is now to couple this energy into another mode basis than the detection mode basis. The output mode basis should now correspond to another spatial degree of freedom. This way, the spatial degrees of freedom to which correspond the input and output mode bases are different, and the modification of the former has a direct effect on the later.

We will introduce in a first part the second order Taylor series describing the displacement, tilt, defocus and waist size change of a Gaussian beam, as well as give a detailed description of how one can establish a link between two distinct spatial degrees of freedom. In a second part, we will present how we designed a MPLC system allowing the coupling of the displacement and tilt of a Gaussian beam to its defocus. Finally, we will present the results of the experiment we constructed to validate this analysis.

3.1 Mode basis change as a parameter transfer tool

The idea that the perturbative treatment of a modified beam involves given spatial mode bases was introduced in [Anderson 84] and [Hsu 04, Delaubert 06b, Lassen 07, Labroille 14] and developed in chapter 2. We recall in the following section the expressions of the Taylor series of a Gaussian beam for four different types of spatial degrees of freedom. The field we will use throughout this chapter is a Gaussian beam:

$$E^{(+)}(\boldsymbol{r}) = HG_{00}(\boldsymbol{r}) = LG_0^0(\boldsymbol{r}). \tag{3.1}$$

In the remaining of the chapter we will refer to this field as "the beam" to simplify discussions and use the notation $E^{(+)}(\boldsymbol{r}) = u(\boldsymbol{r})$. The parameters which we will consider in this section are first the transverse displacement of the beam and the tilt of the beam compared to the optical axis. In a second part we will examine the waist size modification as well as the defocusing of the beam.

3.1.1 Displacement and tilt of a beam

In the previous chapter we have studied in detail the transverse displacement of the beam. We recall here the expressions for this transverse displacement d and also give the equations for the tilt p of the beam. We define $\tilde{u}\,(d, p; x, y, z)$ which we will sometimes simplify as $\tilde{u}\,(d, p)$:

$$\tilde{u}\,(d, p; x, y, z) = u\,(x - d, y, z)\, e^{ipx} \tag{3.2}$$

with $p = 2\pi\theta/\lambda$ and θ is the tilt angle in radians.

3.1.1.1 First order analysis

At the first order in d and p, the Taylor development for the displaced beam gives

$$\tilde{u}(d, p) \simeq \tilde{u}\,(0, 0) + d\,\frac{\partial \tilde{u}}{\partial d}\,(0, 0) + p\,\frac{\partial \tilde{u}}{\partial p}\,(0, 0) \tag{3.3}$$

One can show that

$$\frac{\partial \tilde{u}}{\partial d}\,(d, p; x, y, z) = \frac{1}{w_0} HG_{10}\,(x, y, z) \tag{3.4}$$

$$\frac{\partial \tilde{u}}{\partial p}\,(d, p; x, y, z) = \frac{i w_0}{2} HG_{10}\,(x, y, z) \tag{3.5}$$

Thus we can rewrite equation 3.3 as:

$$\tilde{u}(d, p; x, y, z) \simeq HG_{00}\,(x, y, z) + \left[\frac{d}{w_0} + \frac{ipw_0}{2}\right] HG_{10}\,(x, y, z) \tag{3.6}$$

3.1.1.2 Second order analysis

At the second order in d and p, we have

$$\begin{aligned}\tilde{u}\,(d, p) =& \tilde{u}\,(0, 0) + d\,\frac{\partial \tilde{u}}{\partial d}\,(0, 0) + p\,\frac{\partial \tilde{u}}{\partial p}\,(0, 0) \\ &+ \frac{d^2}{2}\,\frac{\partial^2 \tilde{u}}{\partial d^2}\,(0, 0) + \frac{p^2}{2}\,\frac{\partial^2 \tilde{u}}{\partial p^2}\,(0, 0) + d \times p\,\frac{\partial^2 \tilde{u}}{\partial d \partial p}\,(0, 0)\end{aligned} \tag{3.7}$$

We can derive:

$$\frac{\partial^2 \tilde{u}}{\partial d^2}(d,p;x,y,z) = \frac{\sqrt{2}}{w_0^2} HG_{20}(x,y,z) - \frac{1}{w_0^2} HG_{00}(x,y,z) \tag{3.8}$$

$$\frac{\partial^2 \tilde{u}}{\partial p^2}(d,p;x,y,z) = -\frac{w_0^2}{2\sqrt{2}} HG_{20}(x,y,z) - \frac{w_0^2}{4} HG_{00}(x,y,z) \tag{3.9}$$

$$\frac{\partial^2 \tilde{u}}{\partial d \partial p}(d,p;x,y,z) = \frac{-i}{\sqrt{2}} HG_{20}(x,y,z) + \frac{i}{2} HG_{00}(x,y,z)\,. \tag{3.10}$$

From these expressions we have

$$\begin{aligned}\tilde{u}(d,p;x,y,z) \simeq & \left[1 + \frac{idp}{2} - \frac{d}{2w_0^2} - \frac{p^2 w_0^2}{8}\right] HG_{00}(x,y,z) \\ & + \left[\frac{d}{w_0} + \frac{ipw_0}{2}\right] HG_{10}(x,y,z) \\ & + \left[\frac{d^2}{\sqrt{2}w_0} - \frac{p^2 w_0^2}{4\sqrt{2}} - \frac{idp}{\sqrt{2}}\right] HG_{20}(x,y,z)\,.\end{aligned} \tag{3.11}$$

3.1.2 Defocusing

To defocus a beam is to change the position of the plane in which the waist size is at its minimum. In the definition we introduced in chapter 1, this plane corresponds to $z = z_0$ (see equation 1.32). In much of the derivations of the previous chapters we have set $z_0 = 0$, and we will keep this convention in this chapter. The defocus will be designated by δz_0. A first order Taylor expansion gives

$$\tilde{u}(\delta z_0;x,y,z) = u(x,y,z-\delta z_0) \simeq u(x,y,z) - \delta z_0 \frac{\partial u}{\partial z}(x,y,z) \tag{3.12}$$

We can express the first order derivative along z with modes of the Laguerre-Gauss mode basis:

$$\frac{\partial u}{\partial z}(x,y,z) = \frac{i}{2z_R}\left[\left(1-(kw_0)^2\right) LG_0^0(x,y,z) + LG_1^0(x,y,z)\right]. \tag{3.13}$$

It may be noted that in this case, the first order derivative is not orthogonal to u. Finally, we can express the first order Taylor expansion as:

$$u(x,y,z+z_0) \simeq \left[1 - \frac{i\delta z_0}{2z_R}\left(1-k^2 w_0^2\right)\right] LG_0^0(x,y,z) - \frac{i\delta z_0}{2z_R} LG_1^0(x,y,z)\,. \tag{3.14}$$

We can also show that

$$\frac{\partial^2 u}{\partial z^2}(x,y,z) = \frac{-\lambda^2}{2\pi^2 w_0^4}\left[LG_2^0(x,y,z) + \left(3 + \frac{4\pi^2 w_0^2}{\lambda^2}\right) LG_1^0(x,y,z) + \left(\frac{1}{2} - \frac{12\pi^2 w_0^2}{\lambda^2} + \frac{8\pi^4 w_0^4}{\lambda^4}\right) LG_0^0(x,y,z)\right]. \tag{3.15}$$

3.1.3 Waist size change

Another spatial degree of freedom which one might want to modify is the waist size of a beam. The waist size of the beam can be defined as

$$w(z)^2 = 2 \times \frac{\iint \left(x^2+y^2\right) \tilde{u}^*(x,y,z)\,\tilde{u}(x,y,z)\,dxdy}{\iint \tilde{u}^*(x,y,z)\,\tilde{u}(x,y,z)\,dxdy}. \tag{3.16}$$

For a gaussian beam, the expression of $w(z)$ is known and completely specified by $w(z=0) = w_0$. We consider now the beam $\tilde{u}(\delta w_0; x, y, z)$ with a waist given by $w_0 + \delta w_0$:

$$\tilde{u}(\delta w_0; x, y, z) \simeq u(x,y,z) + \delta w_0 \frac{\partial u}{\partial w_0}(x,y,z) \tag{3.17}$$

$$\frac{\partial u}{\partial w_0}(x,y,z) = -\frac{1}{w_0} LG_1^0(x,y,z) \tag{3.18}$$

$$\tilde{u}(\delta w_0; x, y, z) \simeq LG_0^0(x,y,z) - \frac{\delta w_0}{w_0} LG_1^0(x,y,z). \tag{3.19}$$

We can also show that

$$\frac{\partial^2 u}{\partial w_0^2}(x,y,z) = \frac{2}{w_0^2}\left[LG_2^0(x,y,z) + \frac{3}{2} LG_1^0(x,y,z) - LG_0^0(x,y,z)\right]. \tag{3.20}$$

To complete this description, as we have seen that the derivatives along z and w_0 involve the same Laguerre-Gauss mode basis, we can express the following quantity:

$$\frac{\partial^2 u}{\partial z \partial w_0}(x,y,z) = \frac{-i\lambda}{\pi w_0^3}\left[LG_2^0(x,y,z) + \left(\frac{5}{2} - \frac{2\pi^2 w_0^2}{\lambda^2}\right) LG_1^0(x,y,z)\right] \tag{3.21}$$

Finally, we can write

$$\begin{aligned}\tilde{u}\left(\delta z_0, \delta w_0; x, y, z\right) = \Bigg[1 - \frac{i\delta z_0}{2z_R}\left(1 - k^2 w_0^2\right) - \frac{\delta w_0^2}{w_0^2} \\ - \frac{\delta z_0^2 \lambda^2}{4\pi^2 w_0^4}\left(\frac{1}{2} - \frac{12\pi^2 w_0^2}{\lambda^2} + \frac{8\pi^4 w_0^4}{\lambda^4}\right)\Bigg] LG_0^0\left(x, y, z\right) \\ + \Bigg[-\frac{i\delta z_0}{2z_R} - \frac{\lambda^2 \delta z_0^2}{4\pi^2 w_0^4}\left(3 + \frac{4\pi^2 w_0^2}{\lambda^2}\right) - \frac{\delta w_0}{w_0} + \frac{3\delta w_0^2}{2w_0^2} \\ - \frac{i\lambda \delta w_0 \delta z_0}{\pi w_0^3}\left(\frac{5}{2} - \frac{2\pi^2 w_0^2}{\lambda^2}\right)\Bigg] LG_1^0\left(x, y, z\right) \\ + \left[-\frac{\lambda^2 \delta z_0^2}{4\pi^2 w_0^4} + \frac{\delta w_0^2}{w_0^2} - \frac{i\delta w_0 \delta z_0 \lambda}{\pi w_0^3}\right] LG_2^0\left(x, y, z\right). \end{aligned} \tag{3.22}$$

3.1.4 Coupling of spatial degrees of freedom

From the expressions we have introduced in this section, we see that the Taylor series of the beam with respect to one spatial parameter brings up a given spatial mode basis. It is of course possible to express a tilted beam in any mode basis. However, we see that in the Hermite-Gauss mode basis, the expression at the first or second order of a tilted mode is simple and more importantly, only involves a small number of spatial modes. We developed this idea in the previous chapter and showed that in order to measure a small displacement, the Hermite-Gauss mode basis was optimal: one only needs to measure the intensity of the HG_{10} mode. In the present chapter, we have also shown that the defocus of a beam as well as the waist size change of a beam can be preferentially associated with the Laguerre-Gauss mode basis.

The central idea of this chapter is the following. We consider the situation where two spatial degrees of freedom are linked to two distinct mode bases (which for this demonstration both admit the Gaussian beam as the first mode of the basis). In this paragraph, we call the input and output degrees of freedom p_1 and p_2 and the mode bases $\{u_i\}$ and $\{v_i\}$ respectively. We also consider an ideal mode basis change system between those two mode bases. In this situation, a Gaussian beam launched towards the system defines the *input* beam E_{in}, while the Gaussian beam exiting the system defines the *output* beam E_{out} (see figure 3.1). The decomposition of the

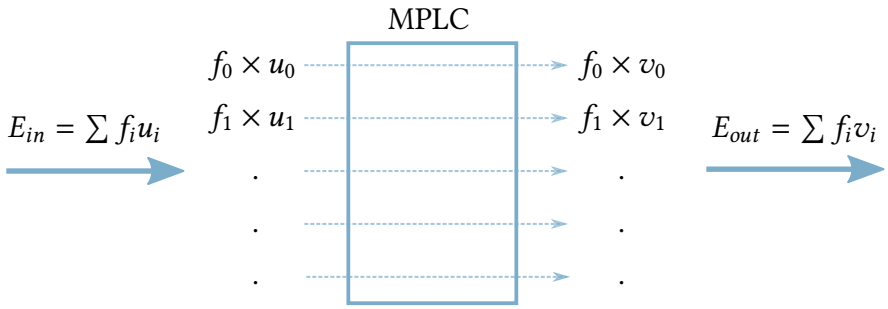

Figure 3.1: Relation between the input and output beam modal decompositions.

input beam on the input mode basis $\{u_i\}$ gives the vector $(f_0, f_1, f_2, \cdots)$

$$E_{in} = \sum_i f_i u_i \text{ with } \sum_i |f_i|^2 = 1. \tag{3.23}$$

The fact that the transform is perfect and unitarity means that the decomposition of the output beam on the output mode basis $\{v_i\}$ is given by the same vector $(f_0, f_1, f_2, \cdots)$:

$$E_{out} = \sum_i f_i v_i. \tag{3.24}$$

In the case we consider where u_0 and v_0 both correspond to a Gaussian beam, if $E_{in} = u_0$ then we have $(f_0, f_1, f_2, \cdots) = (1, 0, 0, \cdots)$ and also $E_{out} = v_0$. If the input beam is now spatially modified along the spatial parameter p_1, the coefficients of its decomposition change from $(1, 0, 0, \cdots)$ to a different vector $\left(f_1^{p_1}, f_2^{p_1}, f_3^{p_1}, \cdots\right)$ which can for instance be derived using a Taylor expansion. In this case,

$$E_{in} = \sum_i f_i^{p_1} u_i \text{ with } \sum_i \left|f_i^{p_1}\right|^2 = 1. \tag{3.25}$$

The unitarity of the transform gives:

$$E_{out} = \sum_i f_i^{p_1} v_i. \tag{3.26}$$

The output beam is thus no longer the Gaussian beam v_0. In this configuration, modifying the parameter p_1 for the input beam directly modifies the coefficients of the output beam in the output mode basis. At the first order in p_1 and p_2, the modification of the parameter p_1 for the input beam implies a modification of the

parameter p_2 for the output beam. This means that one can gain control of one spatial degree of freedom by modifying another spatial property of the field.

A first nuance to add to this analysis is the following. As we have derived in the expressions of this section, it is often the case that two different spatial degrees of freedom are coupled to the same mode basis. In this case, if for the input beam one modifies the parameter p_1, in the general case, the output beam sees the parameter p_2 *and* p_2' modified.

Moreover, the design of the mode basis change system is in fact not completely straightforward. Indeed, in the case where a MPLC system is used, one needs to define two orthogonal mode basis. This was the case in the analysis we conducted with the Taylor expansion of the first order only. With these expressions, there is — in the case of displacement, tilt and waist size change — orthogonality between the spatial profile of the mode and its first derivative. Moreover, in the case of displacement and tilt, which are linked to the same mode basis, the projection coefficients are decoupled in the sense that one is real (d/w_0) and the other is imaginary ($ipw_0/2$) as is expressed by equation 3.6. However we included the second order expansions in the previous subsections to stress the following point: by the second order, this orthogonality property no longer applies, as well as this decoupling of degrees of freedom associated to the same mode basis. For the defocus of a beam, this property is not even true at the first order as equation 3.13 illustrates.

3.2 Design of a parameter transfer MPLC

In order to experimentally test the coupling of spatial degrees of freedom through mode basis change we decided to use, as input degrees of freedom, the displacement and tilt of the beam, and as the output degree of freedom, the defocus of the beam.

3.2.1 Existing defocusing techniques

Changing the focal position of a beam is not in itself a difficult task. The simplest method consists in the mechanical scan of an optical element — for example a lens, or an axicon [Druart 08]. However, like all mechanical movements, the frequency at which this position can be modified is limited by inertia. Indeed, the fast, precise and continuous scanning of the focal position of a beam has a wide variety of applications ranging from autofocusing systems, adaptive optics and imaging to laser processing and optical inspection as well as different types of microscopy (multipho-

ton, confocal, light sheet microscopy, endoscopy, etc.) [Reddy 07, Sancataldo 17]. In all these fields, a precise, high-frequency and continuous technique of scanning a beam's focal point could prove quite beneficial. Several techniques exist which meet some of these requirements although not often all at the same time. The MEMs technology introduced in chapter 1 allows to perform the defocusing of a beam [Yang 06], but the actuating speed, while greater than that of bulky systems, is still limited by inertia. Electro-optic [Shibaguchi 92] and acousto-optic [Kaplan 01] systems offer good alternatives with speeds reaching several hundred of kHz. However, their use is not universal. Electro-optic lenses are polarization dependent and require the use of polarized light. In the case of acousto-optic systems, two synchronized crystals are required in order to mimic a cylindrical lens, which means that in order to reproduce the action of a spherical lens, four acousto-optic systems need to be synchronized. This stacking of components introduces increased levels of losses as well as complexity in the operation of the system. Liquid crystal lenses constitute another possible technology [Sato 85, Lin 10] — although they too require the use of polarized light. Moreover, in order for those lenses to have long focal lengths, the thickness of the liquid crystal constituting the lens must be increased, which in turn reduces the response time of the system. This phenomenon introduces a trade-off between long scanning lengths and high scanning frequencies. In the pulsed regime, a sophisticated technique was developed [Zhu 05, Durst 06, Du 09, Oron 05]. It relies on the manipulation of the spectral bandwidth of the light pulses. Another technique allows to construct a lens with a tunable focal length: these lenses are constructed using two non-miscible liquids, such as oil and water, and electrowetting techniques. The electric charges at an interface allow the modification of the contact angle of a drop on a metal surface [Berge 00]. Finally, TAG or acoustically driven liquid lenses can be used to defocus a beam [Duocastella 13, Mermillod-Blondin 08, Duocastella 14]. The focal length of such lenses is continuously tuned, which produces periodic defocusing.

3.2.2 Principle of the experiment

Our approach differs from the works described above in the sense that it does not try to mimic the working principle of a lens. Instead, we propose to directly couple transverse degrees of freedom to the defocusing of the beam. Several efficients techniques of beam steering are readily available [Römer 14]. In the design we propose, the quality of control of the input spatial parameters is directly transferred on the output parameter since the mode basis change is a passive optical operation. This means that the system described in 3.2 could control precisely, at high frequency

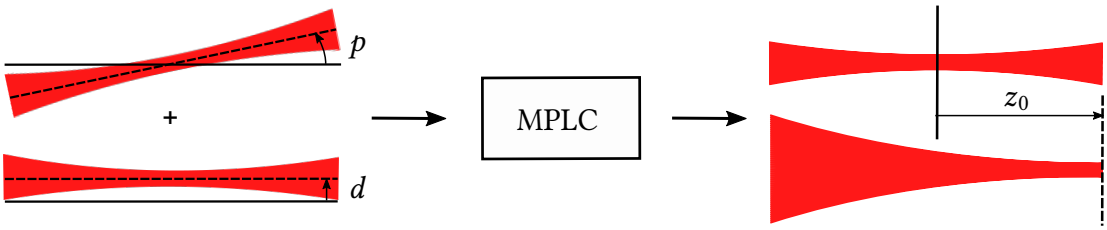

Figure 3.2: The MPLC system couples the tilt and displacement of the input beam to the defocusing of the output beam.

and continuously the position of the focus point of a beam, without any theoretical restriction on the polarization or power of the input beam. Indeed physical implementations of MPLC systems have proven to be compatible with high powers (up to several kW [Garcia 17]) and shape both polarizations identically.

3.2.3 Mode bases definition

We introduced in 3.1.4 the fact that the "derivative" mode bases, which are natural description mode bases when writing the Taylor series of a spatially modified beam, are not in fact suitable as such for the definition of a MPLC system. Indeed, order two derivatives are not orthogonal to the zero and first order derivatives. Suitable orthogonal mode bases can be defined analytically using orthonormalization algorithms. However, the number of modes necessary to correctly describe large parameter modifications scales up quickly.

In order to be able to realize large parameter modifications while using only a small number of modes, we chose to use a numerical method for the definition of the two mode bases necessary to the definition of the MPLC. This choice was motivated by the fact that, in order to produce large displacements, high order derivatives must be taken into account, involving a large number of theoretical modes.

The complexity in terms of phase plates for a MPLC system is linked to the number of modes N_m which defines the transformation (see chapter 1). In this case, we chose $N_m = 5$ so as to keep the number of phase plates small ($N_p = 10$). We found that with $N_m = 5$, we could theoretically obtain an overlap with a perfectly defocused beam of more than 88% on a $4\,z_R$ range. The algorithm we devised for the definition of the mode bases is depicted in figure 3.3 in the case of defocus.

The first step is the definition of the $\{v_i\}$ modes in figure 3.3: we chose beams which are regularly defocused on the range we aim at (here $4\,z_R$) and define the $\{v_i\}$ modes on a common plane $\mathcal{P}$. These modes are not orthogonal. To define the

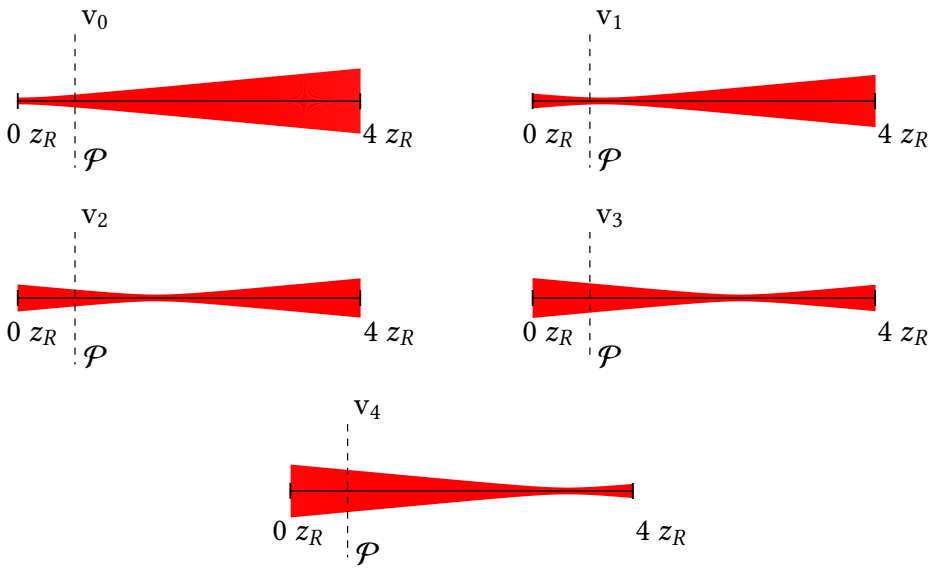

Figure 3.3: Definition of the target modes $\{\mathrm{v}_i\}$ — before orthonormalization.

output mode basis $\{v_i\}$, we apply the Gram-Schmidt algorithm to the $\{\mathrm{v}_i\}$ mode family. We note that in this way, in contrary to what was the case in the previous section, the first mode of the basis v_0 is not a Gaussian beam.

The input mode basis $\{u_i\}$ is defined in a similar fashion, replacing the role of defocus by that of transverse displacement. The corresponding $\{u_i\}$ and $\{v_i\}$ are represented in figure 3.4. As inferred by the analytical approach presented in 3.1, the mode bases bear a close resemblance to the Hermite-Gauss and Laguerre-Gauss mode bases. However, they are not equal.

3.2.4 Definition of the parameter trajectory

Once the input and output mode bases are defined, we can calculate the projection coefficients on each mode for all the positions within the $4z_R$ range we chose. For every position $z_0^{(k)} \in [0z_R, 4z_R]$ we want $E_{out}^{(k)} = u\left(x, y, z - z_0^{(k)}\right)$. By projecting $u\left(x, y, z - z_0^{(k)}\right)$ on the output mode basis $\{v_i\}$, we define the coefficients $\left(f_i^k, f_2^k, f_3^k, \cdots\right)$. These coefficients define in the continuous limit the "projection

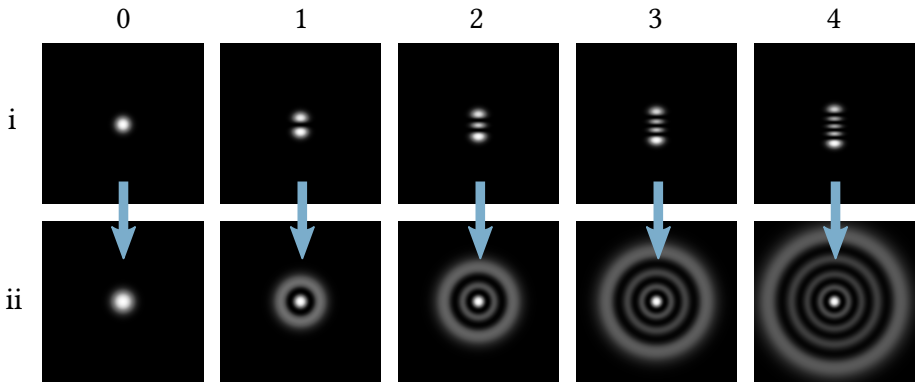

Figure 3.4: Intensity profiles of (i) the input and (ii) output mode bases. The MPLC system shapes mode (i, n) into mode (ii, n).

functions" $\{f_i(z_0)\}$ and we have:

$$E_{out}^{(k)} = \sum_{i=0}^{4} f_i\left(z_0^{(k)}\right) v_i. \tag{3.27}$$

If we consider a perfect unitary transform, we thus need to have the following field at the input of the system:

$$E_{in}^{(k)} = \sum_{i=0}^{4} f_i\left(z_0^{(k)}\right) u_i. \tag{3.28}$$

We now need to determine which couple of displacement $d_i(z)$ and tilt $p_i(z)$ of the input beam allow to produce the correct field $E_{in}^{(k)}$ so that

$$E_{in}^{(k)} = \tilde{u}\left(d_i\left(z_0^{(k)}\right), p_i\left(z_0^{(k)}\right); x, y, z\right) \tag{3.29}$$

- if such a mode exists. To do so, we define the overlap function:

$$G^{(k)}(d,p) = \left|\langle \tilde{u}(d,p;x,y,z) \,|\, \sum_{i=0}^{4} f_i\left(z_0^{(k)}\right) u_i\rangle\right|^2 \tag{3.30}$$

The values of d and p which maximize the function $G^{(k)}$ are defined as $d^{(k)}$ and $p^{(k)}$. We used a least-square numerical method to perform this optimization. In figure 3.5 (left part) we plotted the trajectory which the input beam needs to describe in the

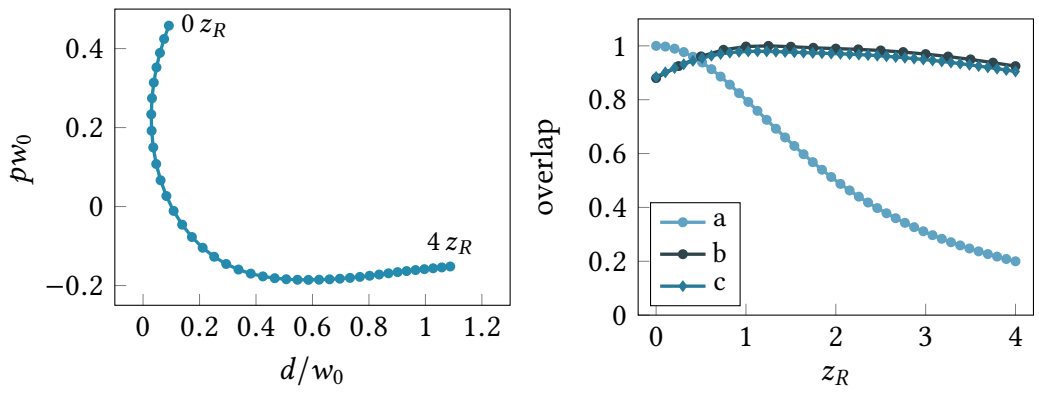

Figure 3.5: Left: Calculated trajectory in normalized units. d is the transverse displacement and p is the tilt imposed on the beam. w_0 is the waist of the beam. This parameter trajectory allows to scan the position of the beam focus from 0 to $4\,z_R$ with z_R the Rayleigh length of the beam. Right: Overlap between the desired output (perfectly defocused gaussian beam) and (a) a "reference" beam (Gaussian beam focused at $z = 0\,z_R$), (b) the beam produced with an ideal unitary operation and (c) the beam produced with the actual MPLC system.

(d, p) plane for the output beam to see its focal point describe the $[0\,z_R, 4\,z_R]$ range. We stress the fact that this trajectory is continuous: any position between 0 and $4\,z_R$ can be attained following this trajectory. The amplitude of the displacement and tilt are directly proportional to the waist size of the beam. For instance, in order to move the focus point of the output beam to $4\,z_R$, one needs to shift the input beam by $d/w_0 = 1.09$ and tilt it by $pw_0/2 = -0.15$. Figure 3.5 (right) displays the overlap between a perfectly defocused gaussian beam and the beam produced by a perfect unitary operation when the trajectory of figure 3.5 (left) is followed. As a mean of comparison, we also plot the overlap between a perfectly defocused Gaussian beam and a fixed Gaussian beam.

3.3 Experimental demonstration of beam defocusing using beam steering

3.3.1 MPLC system

We built the MPLC system corresponding to the mode basis change from $\{u_i\}$ to $\{v_i\}$ to experimentally verify the theory presented in this chapter. In table 3.1, the

convergence of the MPLC algorithm is given for all the modes of the unitary transform. The mean overlap value, 0.991, indicates a very good quality transform. In

Table 3.1: Convergence of the algorithm.

mode index	0	1	2	3	4
overlap	0.989	0.990	0.992	0.993	0.993

figure 3.5 (right), we also plot the overlap between a perfectly defocus beam and the beam produced with the optimized modes defined by the MPLC system. We can observe a very small drop of efficiency of maximum 2% compared to the performances of an ideal unitary system.

3.3.2 Experimental setup

Figure 3.6 presents the experimental setup we built to test the validity of our analysis and numerical results. To have a precise control on the displacement and tilt of the input beam, we used a liquid-crystal SLM. Acousto-optic and galvanometers can serve the same purpose and have superior scanning speeds. In this experiment, the maximum operating frequency was set by the SLM to 203 Hz. The speed of the total system is set by the scanning element since the MPLC is a passive component.

We used a continuous wave laser at $\lambda = 1064$ nm coupled to a single mode fiber. At the output of the fiber, the beam is collimated and we select one polarization of the beam using a polarizing beam-splitter in order for it to be compatible with the SLM. The beam produced by this system is very close to a Gaussian beam. We then use a lens to mode-match the beam so as to produce a Gaussian beam with a waist size $w_0 = 126\,\mu$m. This waist size corresponds to that which we used to define the MPLC system. The SLM surface is divided into two independant parts which play the role of two orientable mirrors, with respective angles α_1 and α_2. The beam thus reflects twice on the SLM. The values of distances L_1, L_2 and L_3 allow to compute the displacement d and tilt p imposed on the beam at its waist. A 4-f telescope was built to image the waist position of the beam. At the output of the telescope, the beam is launched into the MPLC system. The output beam is imaged on a scanning slit beam profiler using a 2-f—2-f setup. The waist size of the output beam is $w_0 = 113\,\mu$m (which corresponds to $z_R = 37.7$ mm).

The z axis is defined by the direction of light propagation. The x and y are respectively the horizontal and vertical directions in the plane normal to the z axis. In our implementation of this experimental setup, the footprint of the optical elements

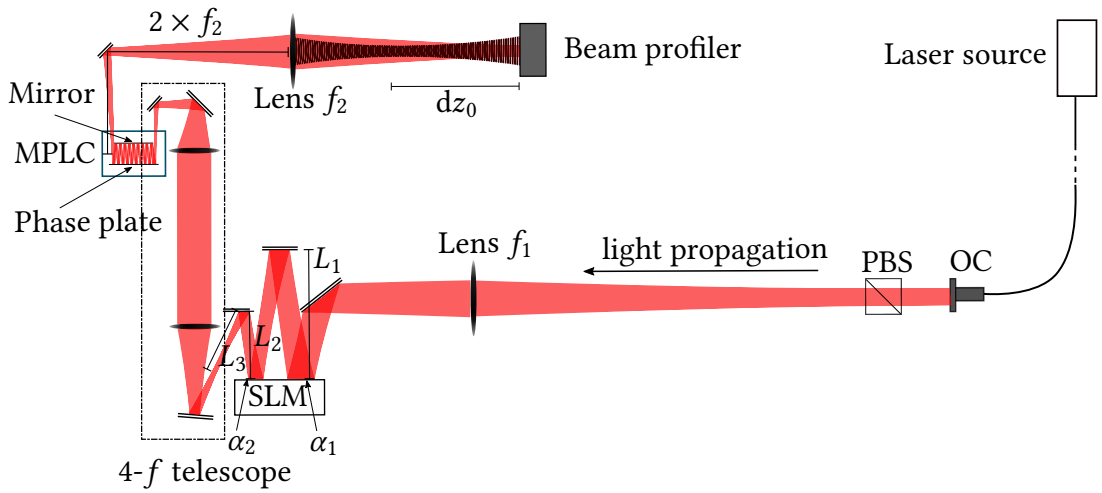

Figure 3.6: Experimental setup. The output of a single-mode fiber is collimated using an output coupler (OC) and mode-matched using a lens of focal length f_1. The polarization compatible with the SLM is selected using a polarizing beam splitter (PBS). Two reflections on the surface of the SLM allow to control the displacement and tilt of the beam. The "tilt" angles of the virtual mirrors which constitute the two halves of the SLM are respectively α_1 and α_2. The focus position of the input beam is imaged using a 4-f telescope and injected into the MPLC system. The output beam is imaged on a beam profiler using a 2-f imaging system.

led to a small clip of one side of the beam - it is a particular problem of the experimental setup and not a fundamental limit. For this reason, the quality of the beam was degraded on the horizontal, or x, direction. The results obtained along the x direction were consequently degraded compared to those of the y direction, and we present here only the results on the y direction. Mirror holders with a smaller footprint might have allowed to avoid this problem.

The experimental procedure we carried out was the following. For a given defocus z_0, we positioned lens f_2 at a distance $z_0 + 2f_2$ (f_2 being the focal length of lens f_2) and the beam profiler at a distance $2f_2$ of the lens. We adjusted the parameters d and p through the angles α_1 and α_2 so as to minimize the waist size measured by the beam profiler. The correct parameters α_1 and α_2 were then fixed and the position of the beam profiler was scanned along the propagation axis to record the evolution of the beam waist size. The results of those measurements are presented on figure 3.7. A Gaussian fit to the recorded data is performed in order to calculate the waist size (dotted lines on figure 3.7). For every position, the experimental data is plotted as well as the corresponding numerical fit (superimposed solid lines) with the waist size formula:

$$w(z) = w_0 \sqrt{1 + M^2 \left(\frac{(z - z_0)}{z_R} \right)^2} \tag{3.31}$$

The theoretical waist size which we aimed at producing is plotted on figure 3.7 as well. We observe a clear defocus on a $4\,z_R$ range, confirming the validity of our approach. The measured waist size varies between 113 μm and 154 μm, while the M^2 factor of the fit to the measurements varies between 1.17 and 1.36[1]. The experimental values, while in good agreement with the theory, do not exactly match with our model. We believe this to be mainly due to two effects. First, as mentioned before, the beam was clipped in the x direction: this introduces a mismatch with respect to a perfect Gaussian beam. Furthermore, we suspect a waist size mismatch between the input beam experimentally produced and the beam used to produce the MPLC device. Those two elements both reduce the overlap between the input beam and the input mode basis used to define the MPLC device.

In this experimental setup, the transmission efficiency is limited by two elements: the efficiency of the steering system and that of the MPLC system. The transmission losses for the MPLC take their origin in the fact that we use finite-size phase plates, discrepancies between the theoretically calculated phase plates and the printed phase plates, as well as from the losses due to the quality of the coating used on the phase plate and mirror. On a typical telecommunication system manu-

[1] A perfect gaussian beam has $M^2 = 1$.

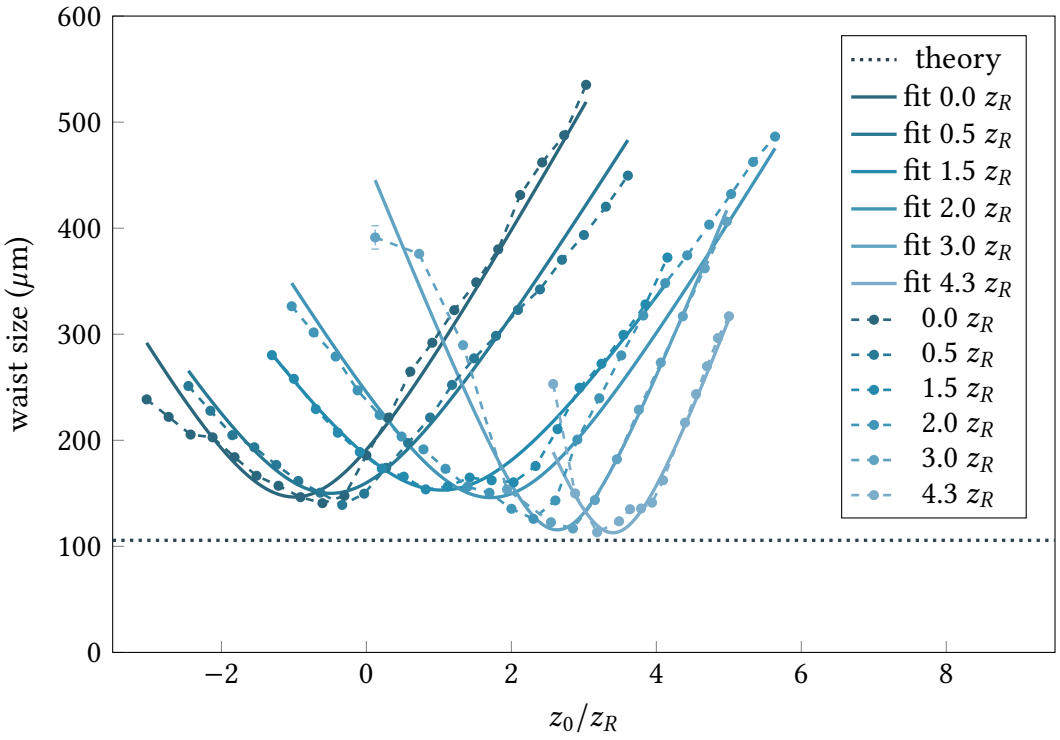

Figure 3.7: Experimental demonstration of defocusing over a $4\,z_R$ range using the displacement and tilt of a beam combined with a MPLC system. The waist size evolution on the y direction is plotted for six different defocusing positions. The dotted line represents the theoretical waist size.

factured by CAILabs, with standard quality evaporated gold reflective coatings and 14 reflections ($N_p = 14$), the overall transmission efficiency of the system is of the order of 70%.

3.4 Conclusion

We demonstrated in this chapter the fact that we can couple different spatial parameters of a laser beam. The theoretical analysis of the Taylor series of modified Gaussian beams seeded the numerical method we developed to address large displacements. To test the validity of this approach, we defined two orthogonal mode bases, one associated with a transverse beam displacement or tilt, the other with a defocus. Using a MPLC system to perform a unitary transform between those two mode bases, we were able to numerically and experimentally couple the angle and displacement of an input beam to the defocus of an output beam over a $4\,z_R$ range. Over the whole scanning range, the overlap between the numerically simulated mode and a perfect defocused Gaussian mode is superior to 88%. An experiment was built to validate this approach. We observed good performances inside the design range which validate our approach. However, the experimental setup was not ideally designed and introduced some errors which reduced the quality of the output beam compared to the numerical experiment. Outside the design range, the quality of the beam deteriorates. We demonstrated that with this system, we can control the defocus of a beam with the control quality one can have on the displacement and tilt of a beam. Using the most precise techniques available for the displacement and tilt of a beam, the scan of the focus point of a Gaussian beam in a nearly continuous fashion in a given range is made available: in order to control the position of the beam with a $z_R/10$ precision (in this case 3.7 mm) along a $4\,z_R$ defocusing trajectory, a precision of $d/w0 = 0/14\%$ and $pw_0/2 = 0.032\%$ (which in this experiment corresponds to $d = 0.15\mu$m and $\theta = 1\mu$rad) is necessary. This scanning can be done at high-frequency since no mechanical element is involved in this system. The theoretical and numerical approaches we put forward in this work are not limited to displacement/tilt and defocus and can be adapted to other spatial degrees of freedom.

Chapter 4

Numerical study of the MPLC system properties

"When in doubt, use brute force."

– Ken Thompson, Bell Labs

Contents

We have introduced in section 1.3 the principle of the MPLC technology and how such a system was designed and built. One noteworthy fact is that a typical MPLC system has a *lot* of degrees of freedom. A typical system which shapes 6 modes has 14 phase plates of 320×500 pixels each. This means that 6 modes are shaped using with more than 2 millions degrees of freedom. Furthermore, the patterns which are written on each phase plate are not — except in the case of very simple transforms where an analytical solution is already known — simple shapes which can be described as functions of a few coefficients. The large number of degrees of freedom is the main motivation for a complex description of a MPLC system. A given MPLC performs a mode basis change. The size of the mode bases is finite and orders of magnitude smaller than the number of degrees of freedom of the system (6 modes with 2 millions pixels). For the modes which are shaped, the transform implemented by the MPLC is unitary. One of the questions we would like to address is: what about the *other* modes? We wish to study how the shaping of a MPLC system affects the modes for which the transform is not defined. Another question which we are interested in is to investigate how a MPLC system reacts to misalignments (change in the length or angle between the phase plate and the mirror for instance): how does it affect the modes which are shaped as well as those for which the transform is not designed?

We say that the MPLC system is — in a certain measure — complex. While complexity is a common concept, giving a definition of it is no easy task. A complex system is more than often characterized by the fact that it possesses a large number of degrees of freedom. One important aspect of complexity is the fact that it may depend on the nature of the description of the system we choose, as well as the scale of the description. A thorough study of complexity is outside the scope of this thesis.

Complex and disordered systems have been thoroughly studied in a number of fields, among which is notably quantum transport. The transport properties of such systems have been derived using a number techniques, such as the diagrammatic approach for instance. We will in this chapter make use of the tools introduced by

random matrix theory.

Our goal in this chapter is to develop a modelisation of MPLC systems compatible with the description of complex and disordered systems. We aim at describing the transport properties of the MPLC systems. We also would like to understand how the different constraints we put on the construction of such systems influence their transport properties. We will in a first part introduce the field of random matrix theory, before we describe the formalism and some results of the study of scattering medium. Finally, we will present how we built a numerical model for the MPLC and our efforts to model its transport properties.

4.1 Random matrix theory

Random matrix theory is a tool which was developed for the purpose of describing the statistical properties of complex many-body systems. It derives the statistical properties of large matrices whose entries are randomly distributed. The symmetries of the matrices or the distributions of the matrix elements are used to derive the singular values of the matrices, the correlations between the singular values and other correlation functions of the matrix elements. This theory does not take into account the microscopic or fine structure of the system it studies nor does it aim at describing these structures. Rather, this theory leans on the large scale symmetries of the system. The success of this modelization lies in the fact that the microscopic degrees of freedom are so plentiful that stochastic treatments provides a near exact description of the system.

4.1.1 Historical overview of random matrix theory

Random matrices have first been introduced in multivariate statistical analysis by Wishart [Wishart 28]. In an attempt to describe the experimentally measured spectra of energy levels (and in particular energy level spacings) of complex nuclei, Wigner, Dyson, Mehta and Gaudin replaced the hamiltonian of the nucleus by random Hermitian matrices (this work is described in [Mehta 91]). They showed that, asymptotically, the spectrum of the correlations of the energy levels only depends on the symmetries of the system — demonstrating through this means universality in spectral correlations.

Another field where random matrix was extremely successful is that of quantum chaos. The Bohigas, Giannoni and Schmit conjecture [Bohigas 84] although not yet formally demonstrated has proven very successful. It states that the statistical

spectral properties of chaotic systems with time-reversal symmetry are predicted by the random matrix theory of orthogonal Gaussian matrices.

At the end of last century, a renewed interest in the field was brought by the development of a random matrix theory of quantum transport, which is concerned with electronic mesoscopic systems. The mesoscopic scale designates the lengths that range between the microscopic and the macroscopic scales. In the mesoscopic study of electron transport, quantum coherence effects due to the wavelike nature of electrons need to be taken into account, while the large number of atoms also requires a statistical treatment [Beenakker 97, Akkermans 07]. The development of the field of quantum transport was of course boosted by the miniaturization of electronic components.

Because only very general concepts are used to develop this formalism, it is easily transferable to other types of wave physics — classical or quantum. Random matrix theory was used to study acoustic [Weaver 89, Aubry 09] and microwave systems [Stöckmann 90, Doron 90], where its predictions proved very successful. More recently, it has been applied to light, whether to describe random lasers (see for example [Goetschy 11]) or light scattering in disordered medium [Vellekoop 08b, Rotter 17].

Finally, random matrix theory is a mathematical field of study as such [Akemann 11]. Its tools and results find applications in a wide variety of disciplines — such a neural networks, numerical linear algebra, information theory or signal processing. As an example, [Tulino 04] gives a good review of the use of random matrix theory to study the fundamental limits of communication channels.

4.1.2 Elements of algebra and multivariate statistical analysis

We give in this section the definitions of some classical mathematical tools which we will make use of in this chapter.

4.1.2.1 Singular value decomposition

The matrices used in random matrix theory are not necessarily square (although it often is the case) and can be real or complex. We define H as an $M \times N$ matrix: $H \in \mathbb{K}^{M\times N}$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. The *singular value decomposition* of H is given by

$$H = UDV^{\dagger} \tag{4.1}$$

with D a $M \times N$ diagonal matrix with real non-negative diagonal entries, U a $M \times M$ unitary[1] matrix and V a $N \times N$ unitary matrix as well. The diagonal entries of D define the singular values of D.
The first $P = \min\{M, N\}$ columns of U and V are respectively the left- and right-singular vectors of the corresponding singular values. With the previous definition, we can write the following relations:

$$H^\dagger H = V \left(D^\dagger D \right) V^\dagger \tag{4.2}$$

$$HH^\dagger = U \left(DD^\dagger \right) U^\dagger. \tag{4.3}$$

Hence, the non-zero singular values of H are the square root of the non-zero eigenvalues of $H^\dagger H$ and $HH^\dagger$.

4.1.2.2 Joint probability distribution

As we stated, the entries of a random matrix are random numbers. These random variables follow probability distributions. The conditional probability is defined as $P(X = x \mid Y = y)$. The joint probability distribution of two discrete random variables X and Y is defined as

$$\begin{aligned} P(X = x \text{ and } Y = y) &= P(Y = y \mid X = x)\, P(X = x) \\ &= P(X = x \mid Y = y)\, P(Y = y) \end{aligned} \tag{4.4}$$

and can easily be generalized to n random variables

$$\begin{aligned} P(X_1 = x_1,\ \ldots, X_N = x_N) = {} & P(X_1 = x_1) \\ & \times P(X_2 = x_2 \mid X_1 = x_1) \\ & \ldots \\ & \times P(X_N = x_N \mid X_1 = x_1,\ \ldots, X_{N-1} = x_{N-1}). \end{aligned} \tag{4.5}$$

If all the X_i random variables are independent, we can rewrite the joint probability distribution as:

$$P(X_1 = x_1,\ \ldots, X_n = x_n) = P(X_1 = x_1)\ P(X_2 = x_2)\ \ldots\ P(X_n = x_n) \tag{4.6}$$

[1] A matrix U is unitary if $UU^\dagger = U^\dagger U = I$.

4.1.3 Some random matrix theory results

We introduce in this section some tools and results which we will make use of in the next sections. However, there is no need to thoroughly understand them in order to follow the rest of the chapter. A hurried reader may very well skip the following section.

4.1.3.1 Probability density

A key symmetry property random matrices must often verify is hermiticity — it is for instance the case when random matrices are used to describe complex hamiltonians. It is however not always the case. In the current subsection we will be interested in such Hermitian matrices, that is matrices for which $H = H^\dagger$. Under this condition, H is necessarily a square matrix with $H_{ii} \in \mathbb{R}$ and $H_{ij} = H_{ji}^{*}$[2]. We define the probability density of sampling the matrix H by using the same notation as for the probability of a single random variable: $P(H)$. To simplify notations we take here the example of a 2×2 hermitian matrix. The three matrix elements are defined as independent which means they are uncorrelated. This translates to

$$P(H) = P_{11}(H_{11})\, P_{22}(H_{22})\, P_{12}(H_{12})\,. \tag{4.7}$$

Moreover, the probability density P for the three independent matrix elements H_{11}, H_{22} and H_{12} should be normalized:

$$\int_{-\infty}^{\infty} P(H)\, dH_{11} dH_{22} dH_{12} = 1. \tag{4.8}$$

We will now introduce two different ensembles of random hermitian matrices which obey supplementary condition: the probability density of H must be invariant upon any canonical transformation:

$$P(H) = P\left(UHU^{-1}\right) \tag{4.9}$$

The ensemble to which U belongs defines the ensemble to which H belongs.

If H is real and the matrix U is orthogonal, that is if

$$U \in \mathbb{R}^{N\times N} \text{ and } U^t U = UU^t = I, \tag{4.10}$$

H is said to belong to the Gaussian Orthogonal Ensemble (GOE).

[2]This notation designates complex conjugation: for $(a, b) \in \mathbb{R}^2$, $(a + ib)^* = a - ib$

If H is complex and U is unitary, that is if

$$U \in \mathbb{C}^{N\times N} \text{ and } U^\dagger U = UU^\dagger = I, \tag{4.11}$$

H belongs to the Gaussian Unitary Ensemble (GUE).

In both cases, one can show (see [Haake 06]) that the probability density of the matrix elements of a random hermitian matrix H admits a gaussian distribution:

$$P(H) = Ce^{-A \operatorname{Tr} H^2} \tag{4.12}$$

where A and C are constants. This results holds true for a $N\times N$ matrix with $N \in \mathbb{N}$.

The eigenvalues of a matrix, not its elements, are often the object in which one is interested when studying a matrix which is why we will now derive the probability density of the eigenvalues of a matrix H. For any dimension N, a Hermitian matrix $H \in \mathbb{C}^{N\times N}$ can be diagonalised by a unitary matrix U: $H = UDU^\dagger$ with D a diagonal matrix. Equation 4.12 thus implies that the eigenvalues of a random hermitian matrix H have a reduced distribution: indeed, the description of D uses two coefficients while H uses four real numbers. For the sake of simplicity, we go back to the $H \in \mathbb{C}^{2\times 2}$ case to present the following demonstration. If we set the determinant of U to unity (which up to a phase factor is always possible), we can write the unitary transform U using two parameters Θ and ϕ:

$$U = \begin{pmatrix} \cos\Theta & -e^{-i\phi}\sin\Theta \\ e^{i\phi}\sin\Theta & \cos\Theta \end{pmatrix} \tag{4.13}$$

and the matrix D takes the following expression

$$D = \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix}. \tag{4.14}$$

E_+ and E_- are the eigenvalues of H. The independent elements of H can thus be expressed as:

$$H_{11} = E_+ \cos\Theta^2 + E_- \sin\Theta^2 \tag{4.15}$$

$$H_{22} = E_+ \sin\Theta^2 + E_- \cos\Theta^2 \tag{4.16}$$

$$H_{12} = (E_+ - E_-)\, e^{i\phi} \cos\Theta \sin\Theta \tag{4.17}$$

To make a change of variable in equation 4.8 from the matrix elements to the matrix eigenvalues, we introduce the jacobian $J = (E_+ - E_-)^2 \cos\Theta \sin\Theta$. By integrating

the angles Θ and ϕ, one can show (see [Haake 06]) that the probability distribution of the eigenvalues of a 2×2 random hermitian matrix is given by

$$P(E_+, E_-) \propto C\,|E_+ - E_-|^2\, e^{-A\left(E_+^2 + E_-^2\right)}. \tag{4.18}$$

If now the matrix H belongs to the GOE, the same derivation can be performed to show that

$$P(E_+, E_-) \propto C\,|E_+ - E_-|\, e^{-A\left(E_+^2 + E_-^2\right)}. \tag{4.19}$$

These expressions can be generalized to higher dimensions of the matrix H. The joint distribution of the eigenvalues E_μ of a hermitian matrix takes the following form [Haake 06]:

$$P\left(\{E_\mu\}\right) = C \prod_{\mu<\nu}^{1\ldots N} \left|E_\mu - E_\nu\right|^\beta \exp\left(-A \sum_{\mu=1}^{N} E_\mu^2\right) \tag{4.20}$$

β is equal to 1 or 2 if the matrix belongs respectively to the GOE or the GUE[3]. The first factor in equation 4.20 indicates clearly that there can be no degeneracies of the eigenvalues: the probability of having two equal eigenvalues is necessarily equal to zero. This phenomenon is known as energy level repulsion[4].

4.1.3.2 Density of states

The normalized *density of states* of the system — or density of levels — described by the matrix H is defined as

$$\rho(E) = \frac{1}{N} \sum_{i=1}^{N} \delta(E - E_i)\,. \tag{4.21}$$

This function describes the behaviour of the typical eigenvalue. In solid-state physics, it quantifies the number of electronic states which are found at a given energy. It is for instance used to define Fermi's golden rule. This expression of the density of states is valid for one particular matrix. We would like to know what is the typical density of states of a matrix when this matrix is sampled in a given ensemble (for instance the GUE). We thus define the ensemble averaged density of state ρ as

$$\rho(E) = \int P(E, E_2, \ldots, E_N)\, dE_2 \ldots dE_N \tag{4.22}$$

[3] Although its detailed description is outside the scope of this thesis, another ensemble with $\beta = 4$ can be defined. It is called the Gaussian Symplectic Ensemble.

[4] Indeed, if H is the hamiltonian of a system, the eigenvalues represent the energy levels.

which holds for all the samples of a given ensemble. We use the same letter in equations 4.21 and 4.22, since as N goes to infinity, one matrix's density of states converges quickly to that of the ensemble it belongs to.

Semi-circle law For this integral to be well-defined when $N \longrightarrow \infty$, a normalisation can be chosen such that the second-order moments of the matrix elements $\langle H_{ij} H_{ji} \rangle$[5] (for H in the GUE) take the value $1/(4N)$. If so, the density of levels converges almost surely to the famous Wigner's semicircle law [Haake 06]:

$$\rho(E) = \begin{cases} (2/\pi)\sqrt{1-E^2} & \text{if } |E| \leq 1 \\ 0 & \text{if } |E| > 1 \end{cases} \tag{4.23}$$

Marčenko-Pastur law Another famous distribution can be derived using random matrix theory for a different class of random matrices. We will now consider not Hermitian matrices but rather $N \times M$ matrices H with independent identically distributed entries whose mean value is equal to zero, variance to $1/N$ and fourth-order momentum of order $O\left(1/N^2\right)$. The ratio $M/N \longrightarrow \beta$ as $N, M \longrightarrow \infty$. The level density of the matrix $H^\dagger H$ converges almost surely to the Marčenko-Pastur distribution [Marčenko 67]

$$\rho(E) = \left(1 - \frac{1}{\beta}\right)^+ \delta(E) + \frac{\sqrt{(E-a)^+ (b-E)^+}}{2\pi\beta E} \tag{4.24}$$

with $a = \left(1 - \sqrt{\beta}\right)^2$ and $b = \left(1 + \sqrt{\beta}\right)^2$ (see [Tulino 04])[6].

4.1.4 Free probability theory

A significant development in random matrix theory has been the application of the non-commutative *free-probability* theory to random matrices. Free-probability was introduced by Voiculescu [Voiculescu 83]. This theory was developed to tackle the following problem: if we know the spectral properties of two matrices H_1 and H_2, can we have information on the spectral properties of the sum $H_1 + H_2$ or $H_1 H_2$?

[5] The average $\langle \ldots \rangle$ is taken as an ensemble average.

[6] We define x^+ as

$$x^+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{4.25}$$

In the usual context, unless H_1 and H_2 commute, we cannot say anything about the properties of $H_1 + H_2$, $H_1 H_2$ or any combination of H_1 and H_2. Free probability studies noncommutative random variables. The "freeness" character is the equivalent of the classical notion of independence of random variables.

4.1.4.1 Resolvent

Several types of transforms can be defined to describe the spectrum properties of random matrices (see appendix C). If we consider a Hermitian $N \times N$ matrix H with eigenvalues E_i, the resolvent of the matrix is defined as (see [Goetschy 11])

$$g(z) = \frac{1}{N} \left\langle \mathrm{Tr} \frac{1}{z - H} \right\rangle = \frac{1}{N} \left\langle \sum_{i=1}^{N} \frac{1}{z - E_i} \right\rangle \tag{4.26}$$

where the average $\langle \ldots \rangle$ is taken over a number of realizations of H. The resolvent can be linked to the level density of the matrix by

$$g\left(E + i\epsilon\right) = \mathbf{P} \int_{-\infty}^{\infty} dE' \frac{\rho\left(E'\right)}{E - E'} - i\pi\rho\left(E\right) \tag{4.27}$$

with P the principal value of the integral[7]. The expression of $\rho\left(E\right)$ can be reciprocally derived from g:

$$\rho\left(E\right) = -\frac{1}{\pi} \lim_{\epsilon \to 0^+} \mathrm{Im}\left[g\left(E + i\epsilon\right)\right]. \tag{4.29}$$

The resolvent is a very useful tool in other areas, for instance in physics for the interactions of atoms with the electromagnetic field [Cohen-Tannoudji 68].

4.1.4.2 Free probability

The following paragraph is greatly inspired from [Goetschy 11]. We briefly recall some properties of independent random variables before defining their free counterparts. The probability density of a random variable x is given by p_x. Its cumulant

[7] The principal value is defined as

$$\mathbf{P} \int_a^b f\left(x\right) dx = \lim_{\epsilon \to 0} \left(\int_a^{c-\epsilon} f\left(x\right) dx + \int_{c+\epsilon}^b f\left(x\right) dx \right). \tag{4.28}$$

generating function is given by $r_x(z) = \ln(\langle e^{zx} \rangle)$. If two random variables x_1 and x_2 are independent, they verify the following properties:

$$\langle x_1 x_2 \rangle = \langle x_1 \rangle \langle x_2 \rangle \tag{4.30}$$

$$p_{x_1+x_2} = p_{x_1} p_{x_2} \tag{4.31}$$

$$r_{x_1+x_2} = r_{x_1} + r_{x_2}. \tag{4.32}$$

We define the expectation value $\langle \dots \rangle_\Lambda$ for a $N \times N$ random matrix X as:

$$\langle X \rangle_\Lambda = \frac{1}{N} \langle \mathrm{Tr}\, X \rangle \tag{4.33}$$

Two Hermitian matrices X_1 and X_2 are asymptotically free if for all $l \in \mathbb{N}$ and all polynomials P_i and Q_i $(1 \leq i \leq l)$ we have [Tulino 04]:

$$\forall i \ \langle P_i(X_1) \rangle_\Lambda = \langle Q_i(X_2) \rangle_\Lambda = 0 \implies \langle P_1(X_1) Q_1(X_2) \dots P_l(X_1) Q_l(X_1) \rangle_\Lambda = 0. \tag{4.34}$$

This equation can be interpreted as follows: two random matrices X_1 and X_2 are free if their eigenvectors are almost surely orthogonal. If X_1 and X_2 are free, then we can write

$$\langle X_1 X_2 \rangle_\Lambda = \langle X_1 \rangle_\Lambda \langle X_2 \rangle_\Lambda. \tag{4.35}$$

The transforms defined in appendix C allow to derive interesting results for free matrices [Goetschy 11]. The R-transform of the sum of two free matrices writes

$$\mathrm{R}_{X_1+X_2}(z) = \mathrm{R}_{X_1}(z) + \mathrm{R}_{X_2}(z). \tag{4.36}$$

Likewise, the S-transform of the product of two free matrices writes

$$\mathrm{S}_{X_1 X_2}(z) = \mathrm{S}_{X_1}(z)\, \mathrm{S}_{X_2}(z). \tag{4.37}$$

The parallel with classical probability can be extended by noting two interesting properties. The semicircle law can be seen as the free analog of the Gaussian distribution (in the case of the central limit theorem): the spectrum of the normalized sum of free random matrices converges to the semi-circle law [Tulino 04]. Likewise, the spectrum of the normalized sum of free random matrices of unit rank converges to the Marčenko-Pastur law which can thus be seen as a free analog of the Poisson distribution

4.2 Scattering in complex media

In this section we aim at describing systems in which a wave propagates through a disordered or complex medium. Even though the equations which describe the propagation of a wavefront in such a medium could in theory be written down and solved, it is in practice often not the case. Just like specifying all the positions and momenta of molecules in a gas is both unrealistic and unnecessary, giving the positions of all scatterers in an opaque material, is practically infeasible and is in fact not at all necessary to describe the interference phenomena which take place when a wave propagates through such systems. *Mesoscopic scattering theory* tackles such problems. It was originally developed to describe the quantum transport properties of electrons, and latter adapted to describe the propagation of light in complex media [Beenakker 97, Akkermans 07]. The parallel between the two fields can be made by comparing the Schrödinger equation for the evolution of an electron in a potential and the scalar Helmholtz equation for an electric field in a medium described by the dielectric function $\epsilon(\boldsymbol{r})$ (a complete derivation can be found in [Rotter 17]). This correspondence between the equations describing electronic and light scattering is essential to transfer results and concepts from one field to another — the validity of this parallel being set by the respective validity limits of both the Helmholtz and Scrödinger equations.

4.2.1 Scattering matrix formalism

We do not aim at describing the propagation medium through its microscopic components. Instead, we consider it as a bulk connected to a number of ports — in this thesis we only consider the case of two connecting ports. Light or electronic waves launched through these ports can be either reflected or transmitted. We will thus describe the scattering medium by its reflection and transmission coefficients. The scattering matrix encompasses such a modelisation. A good introduction to this formalism can be found in [Brouwer 97, Rotter 17] — the next section is largely based on these references. The system we consider is described in figure 4.1. We suppose that the left input port can support N modes — or channels — and that the right input port can support M modes. In the electronic case, this quantization of the number of modes comes from the finite transverse dimensions of the wire. In optics, these ports can for instance be few or multi-mode fibers. We define the $(N+M)\times(N+M)$ scattering matrix S which relates incoming and outgoing modes

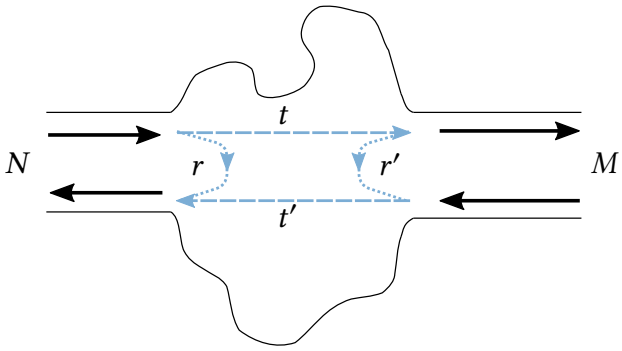

Figure 4.1: Scattering matrix components. t and t' are transmission matrices and r and r' reflection matrices.

as

$$S = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \tag{4.38}$$

$t \in \mathbb{C}^{M\times N}$ and $r \in \mathbb{C}^{N\times N}$ respectively characterize the amplitudes of the modes incoming from the left in figure 4.1 which are transmitted and reflected. $t' \in \mathbb{C}^{N\times M}$ and $r' \in \mathbb{C}^{M\times M}$ characterize the amplitudes of the transmitted and reflected modes incoming from the right in figure 4.1. The matrix S gives the relation between the modes which come in and out of the system. We use the subscripts l and r to designate the left and right parts in figure 4.1 and the superscript + and − to designate the rightward and leftward propagating modes. We can then write:

$$\boldsymbol{e}_{in} = \begin{pmatrix} \boldsymbol{e}_l^+ \\ \boldsymbol{e}_r^- \end{pmatrix}, \ \boldsymbol{e}_{out} = \begin{pmatrix} \boldsymbol{e}_l^- \\ \boldsymbol{e}_r^+ \end{pmatrix} \text{ and } \boldsymbol{e}_{out} = S\boldsymbol{e}_{in}. \tag{4.39}$$

The off-diagonal blocks t and t' respectively contain the transmission amplitudes from left to right and right to left. We make the assumption that there are no losses nor absorption. Flux conservation requires the matrix be unitary:

$$S^\dagger S = SS^\dagger = 1. \tag{4.40}$$

For the mode of index n incoming on the left, its transmission and reflection coefficients are $T_n = \sum_{m=1}^{N} |t_{mn}|^2 = \left(t^\dagger t\right)_{nn}$ and $R_n = \sum_{m=1}^{N} |r_{mn}|^2 = \left(r^\dagger r\right)_{nn}$.

The total transmission T and reflection R for modes incoming from the left in

figure 4.1 can be expressed as

$$T = \mathrm{Tr}\left(t^\dagger t\right) = \sum_k \tau_k \tag{4.41}$$

$$R = \mathrm{Tr}\left(r^\dagger r\right) = \sum_k v_k \tag{4.42}$$

where τ_k and v_k are the eigenvalues of $t^\dagger t$ and $r^\dagger r$[8]. In general, T_n is not equal to any of the τ_k, meaning the transmission coefficient of mode n is different from that of the transmission eigenmode k. In the case where $N = M$, it can also be shown [Rotter 17] that the matrices $t^\dagger t$ and $r^\dagger r$ can be simultaneously diagonalized and that in this case their eigenvalues obey the following relation $\tau_k = 1 - v_k$.

However, t is not necessarily square. Using singular value decomposition introduced in 4.1.2.1, we can write $t = UDV^\dagger$. The nonzero singular values σ_k of t are the square roots of the eigenvalues of $t^\dagger t$: $\sigma_k = \sqrt{\tau_k}$.

Using the scattering matrix formalism, concatenated systems can be described simply: the transmission matrix t allows to describe the transmission properties of two scattering media connected by a common port. In this configuration, the total transfer matrix is expressed as a product [Brouwer 97]:

$$t = t_2\, t_1 \tag{4.44}$$

with t_1 the left-most scattering system transfer matrix, t_2 the right-most one and t that of the total system. The transmission properties of large scattering systems can in this way be derived from those of smaller parts of the system.

4.2.2 Mesoscopic scattering theory

We present in the following section some of the most remarquable results derived in mesoscopic scattering theory. As mentioned before, the goal of our approach is not to describe every coefficient of the scattering matrix S. Rather, we wish to deduce the transport properties from the ensembles to which the matrix S belongs. We

[8] A physical quantity which is linked to T is the conductance. It is a central quantity in mesoscopic physics. It is defined using the transmission matrix:

$$G = G_0 \mathrm{Tr}\, t^\dagger t = G_0 T. \tag{4.43}$$

with $G_0 = e^2/h$. In electricity, this quantity measures the capacity of a material to transmit electric current — it is the inverse of the resistivity. The definition we give here allows to study this object both in mesoscopic physics and optics (see 4.5).

have for instance indicated in the previous section that S must be unitary when no absorption or loss are present. We will present two specific cases which have been studied in the context of mesoscopic scattering theory and for which this question has been solved — the quantum dot and the disordered wire. The different statistics we derive for these two models stem from the origin of disorder: in one the chaos[9] comes from boundary scattering whereas in the other the disorder is due to impurity scattering.

Quantum dot A quantum dot usually designates a semiconductor structure in which the electronic wave function is confined in all three dimensions. We are interested here in quantum dots which consist of structures etched on semiconductors: they form a cavity for a two dimensional electron gas[10]. A quantum dot is linked to electron reservoirs. In the case where two of them are present, a quantum dot corresponds to the geometry studied in 4.2.1. The shape of the confining potential may lead to ballistic chaotic trajectories. Billiards are good examples of such potentials: in a stadium-shaped billiard (see figure 4.2 for illustration), a classical particle has chaotic trajectories. In the case where the coupling is ideal (there is

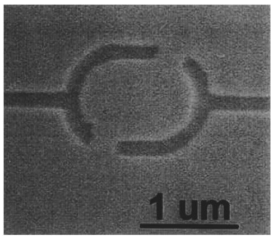

Figure 4.2: Stadium-shaped quantum dot. Image taken from [Chang 94].

no barrier between the quantum dot and the electron reservoirs), the probability density of the scattering matrix S of the quantum dot at the Fermi energy[11] is taken to be constant (derivation of this result can be found in [Beenakker 11]). This property implies that the scattering matrix belongs to the Dyson circular ensemble of

[9] Chaos is associated with systems where trajectories are sensitive to initial conditions: a slight change in initial conditions leads to exponentially diverging trajectories.

[10] These structures are particularly interesting because the electron wavelength at the Fermi energy is two orders of magnitude greater than that of an electron in a metal. It is thus easier to study the statistics of the level spacing in these systems compared to metallic structures since the energy resolution required to probe the energy levels is not as sharp.

[11] The Fermi energy refers to the top most filled energy level at 0 K.

random matrices[12]. If we now suppose that the number of transmission and reflection channels are equal and that N is much greater than 1 (i.e. we suppose that S is a very large matrix), the normalized density of transmission eigenvalues can be derived [Beenakker 11]:

$$\rho(\tau) = \frac{N}{\pi} \frac{1}{\sqrt{\tau}\sqrt{1-\tau}}. \tag{4.45}$$

This distribution has a *bimodal* distribution and its mean transmission value is $\bar{\tau} = 1/2$ (see left part of figure 4.4).

Quantum wire The diffusive regime in scattering systems is defined by the fact that multiple scattering is the dominant process — as opposed to the ballistic regime. In the case of a metal wire, the diffusive regime can be observed in the case where the length L of the wire is large compared to the mean free path l while also small compared to the localization length ξ[13]: $l \ll L \ll \xi$. The mean free path l gives a measure of the length after which the propagation direction of an electron is randomized — it characterizes the scattering properties of the medium. In a disordered

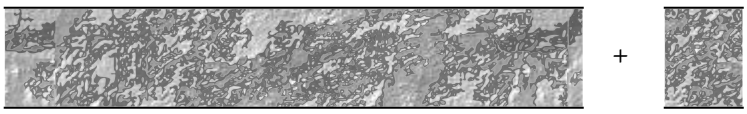

Figure 4.3: Weakly scattering segments are recursively added to the wire.

wire, the degree of disorder experienced by an incoming wave depends strongly on the "time" it spends in the medium: the thickness or length of the wire plays an important role. To describe a disordered wire, Dorokhov, Mello, Pereyra and Kumar (DMPK) proposed a model [Dorokhov 82, Mello 88] which breaks down the wire into thin scattering slices and stacks them together in a recursive way (see figure 4.3). The length of each scattering element is taken to be much larger than the wavelength while shorter than the transport mean free path. The so-called DMPK

[12]This ensemble is composed of unitary matrices which columns are orthogonal and have a norm equal to 1 [Dyson 62, Beenakker 11].

[13]The localization length is given by $\xi = (\beta N + 2 - \beta)\, l$. β is the symmetry index and N the number of modes [Beenakker 11].

equation describes the evolution of the transmission eigenvalues τ_n as a function of the total length of the wire L (see [Beenakker 97]). The random matrix we consider in this case is not the scattering matrix S but the transmission matrix t. For a metallic wire of cross-section A and Fermi wavelength λ_F[14], the number of propagating modes can be approximated as $N \simeq A/\lambda_F^2$. t is thus a $N \times N$ matrix (typically, $N \gg 1$). It can be shown that in the diffusive regime, the DMPK equation can be solved. One can deduce from it the eigenvalue density of the system (see [Beenakker 97] for a complete derivation):

$$\rho(\tau) = \frac{Nl}{2L} \frac{1}{\tau\sqrt{1-\tau}}, \quad \text{for } \tau_{min} \leq \tau < 1 \tag{4.46}$$

with $\tau_{min} \simeq 4e^{-L/2l}$ (see right part of figure 4.4). This distribution is again bimodal with a mean transmission value $\bar{\tau} = l/L$. The inverse values of the localization lengths ξ_n must be evenly distributed [Beenakker 97]. The localization lengths and the transmission intensity eigenvalues are related through the following equation: $\tau_n = \cosh^{-2}(L/\xi_n)$. This bimodal distribution was experimentally measured using elastic waves in a disordered waveguide [Gérardin 14].

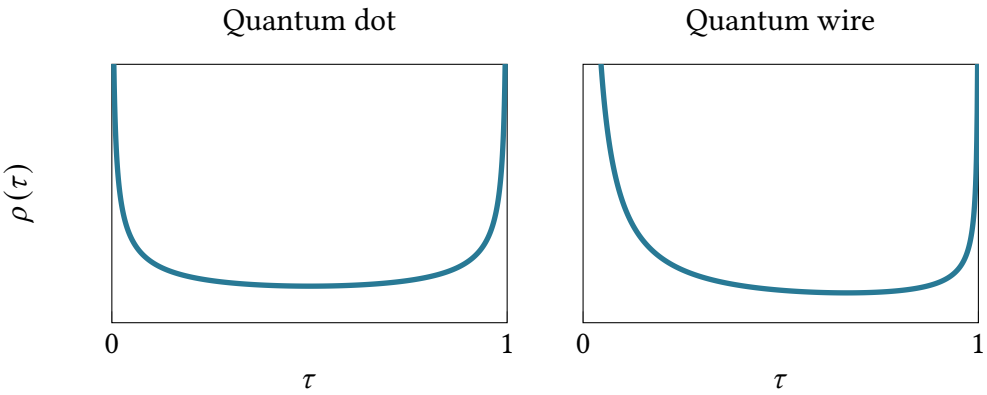

Figure 4.4: Bimodal distributions

Open and closed channels In both distributions defined in 4.45 and 4.46 and represented in figure 4.4, we can observe that the transmission coefficients are peaked around 0 and 1. This means that nearly all the transmission channels are "closed"

[14]The Fermi wavelength is the de Broglie wavelength of the electrons present near the Fermi energy level.

($\tau_n \simeq 0$) or "open" ($\tau_n \simeq 1$). The number of such "open" channels is discussed in [Imry 86]. These distributions, like all objects of random matrix theory, are to be interpreted as a statistical result: they do not specify the shape of the eigenvalue distribution of one system but rather the shape of the eigenvalue distribution of an ensemble of realizations of similar systems.

Conductance quantization The following result, although not connected to the work conducted in this thesis, illustrates nicely the correspondence which can be made between electronics and optics. Mesoscopic scattering theory makes the demonstration of conductance quantization. When the size of the contact to a quantum dot is continuously tuned, the conductance G increases in steps of size G_0. As the size of the contact increases, the number of guided modes will increase by steps of one: the size of the contact when a step is reached corresponds to the threshold for a new mode to be guided. Hence, according to equation 4.43, the total conductance will increase by a factor e^2/h. Since this effect relies entirely on the wave-like nature of electrons, it can also be implemented with electromagnetic fields, as shown in figure 4.5.

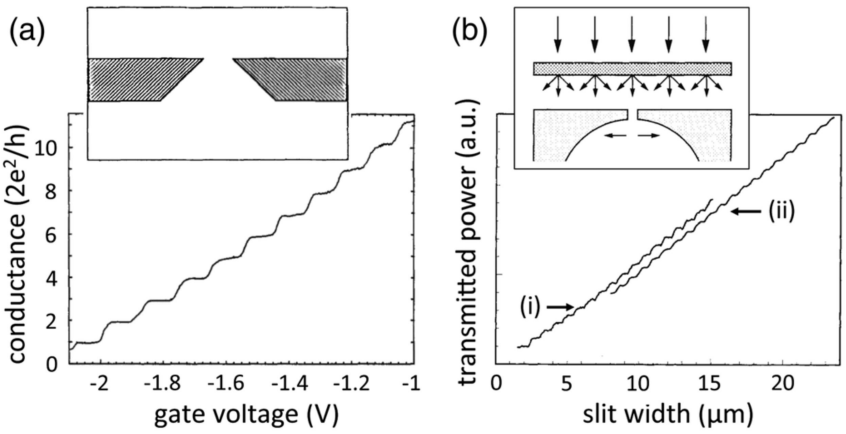

Figure 4.5: Conductance quantization for electronic transport through a quantum point contact ((a), taken from [van Wees 88]) and light intensity transmission through a slit of variable size ((b), taken from [Montie 91]). The complete figure was taken from [Rotter 17].

4.2.3 Mesoscopic effects in optical systems

As we have introduced in the previous section, many phenomena first discovered in the field of mesoscopic transport find their counterpart in optics. A good review of the correspondence between the two fields can be found in [Rotter 17]. However, the field of optics offers one advantage to push this study further. Indeed, in the case of electronic transport, the coherence of the electronic wave packet is at the heart of the phenomena we describe. The systems under study should thus be protected from decoherence, which is a technical challenge. It is also harder to individually control modes in electronic system compared to optical systems. On the contrary, in optics, one can manipulate electromagnetic modes individually and thus access their reflection or transmission coefficients separately. The drawback of optical systems is that the number of modes is very large which means that the scattering matrix of a system cannot usually be completely measured. We note that a large amount of work was conducted in the fields of microwaves [Aubry 10] and acoustic waves, which offer the advantage of a reduced number of modes compared to optical systems.

Speckle A well known optical phenomenon is speckle. Speckle is formed when a coherent field is shone through, or reflected from, a thin diffusing object (for example a sheet of paper). The resulting intensity pattern has a random aspect and follows — in the case of perfect coherence of the source — the following probability distribution [Goodman 76]:

$$p(I) = (1/\langle I \rangle)\, e^{-I/\langle I \rangle}. \tag{4.47}$$

This behaviour can be derived from the fact that the intensity on each point is the sum of a large number of random but identically distributed random variables. The field $E = X + iY$ on point x is the sum of a large number of transmitted fields which have followed all the possible trajectories S to reach the observation point: $E = \sum_S E_S$. The fields E_s are independent and identically distributed. This means that $X = \sum_S X_s$ and $Y = \sum_S Y_s$ will, according to the central limit theorem, follow the distributions:

$$p(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-X^2/2\sigma^2} \tag{4.48}$$

$$p(Y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-Y^2/2\sigma^2} \tag{4.49}$$

with $\sigma^2 = \langle X^2 \rangle = \langle Y^2 \rangle = \langle I \rangle / 2$. We can derive the joint distribution

$$p(X, Y) = \frac{1}{\pi \langle I \rangle} e^{-I/\langle I \rangle}. \tag{4.50}$$

Finally, the distribution of intensities obeys

$$p(I) = \iint p(X, Y)\, \delta\left(X^2 + Y^2 - I\right) dX dY = \frac{1}{\langle I \rangle} e^{-I/\langle I \rangle}. \tag{4.51}$$

The speckle phenomenon can also be studied in the context of random matrix theory and equation 4.47 derived using a different demonstration (see [Kogan 95] for the complete derivation).

Optical experiments using wavefront shaping Equation 4.46 was originally derived for electronic transport in a waveguide. It was shown that this expression is valid in a broader context [Nazarov 94], in particular in a slab geometry. This is of prime interest since a large number of free-space optics experiment, in particular those in which light is sent through disordered media, can be described as slab geometries. Optics experiment are of particular interest to test the results we introduced thanks to the development of beam shaping techniques such as those introduced in chapter 1. Indeed, tools such as a SLM allow to send through the disordered medium any one mode at a time. Theoretically, this allows the measurement of the t matrix (or of the r matrix). Singular value decomposition of this measured matrix then allows to find the eigenmodes of propagation — that is the "open" channels for a given disordered medium. The measurement of the transmission matrix of a sample was demonstrated in [Popoff 10, Popoff 11]. The effects of an incomplete control over all the channels is studied in [Goetschy 13]. However, the singular value distribution of the transmission matrix measured in [Popoff 10] does not correspond to a bimodal distribution. This is due to the fact that the matrix measured in the experimental setup is not the complete transmission matrix but rather a filtered version of the matrix. One of the most noteworthy experiments conducted with optical systems is presented in [Vellekoop 08a]. The transmission matrix coefficients are measured and the variance of the eigenvalues distribution is probed. The transmission matrix is first used to adjust the intensity and phase of each of the available input modes so as to maximize the intensity at a given point in a plane lying behind the scattering medium (see figure 4.6 left). The experiment is conducted a large number of times (by changing the disordered medium) to compare the predictions of random matrix theory with the experimental results. The

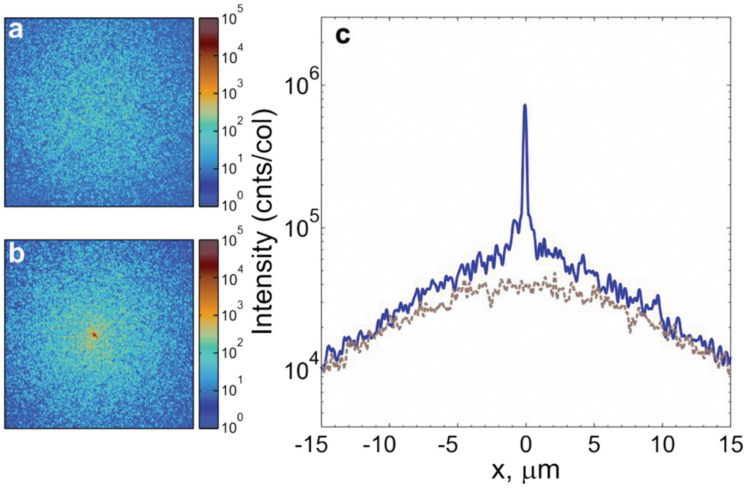

Figure 4.6: Intensity distribution of light in (a) a nonoptimized wavefront and (b) the optimized wavefront. In (c), the intensity is summed in the y direction to average over speckle. The dashed brown curve corresponds to the nonoptimized wavefront while the blue solid curve displays the optimized wavefront. The complete figure was taken from [Vellekoop 08a].

bimodal distribution of disordered media predicts that in the case where the incident wavefront is perfectly shaped to match with transmission eigenmodes, the ratio of the incident and transmitted intensities is equal to $2/3$. This ratio was experimentally measured to 0.68 ± 0.07 in excellent agreement with theory (see figure 4.6). This experiment is therefore a first experimental confirmation of the bimodal distribution of transmission coefficients. In a silicon waveguide, the energy density distribution of modes corresponding to open and closed channels was measured in [Sarma 16].

4.3 Modelization of the MPLC

We have introduced a number of tools, among which random matrix theory, to study the transmission properties of complex systems and want to use them to study the transmission properties of MPLC systems. In the introduction of this chapter, we explained that the MPLC system possesses a number of degrees of freedom much greater than the number of constraints of the system. The transmission properties of the MPLC system are only known for a very small number of modes: the number of modes the MPLC system is designed to shape. In order to study the transmission

properties of the MPLC systems for other modes, we want to gain access to the transmission matrix of the MPLC. The experimental measurement of such a matrix is quite a difficult task which is why we turned to a numerical approach.

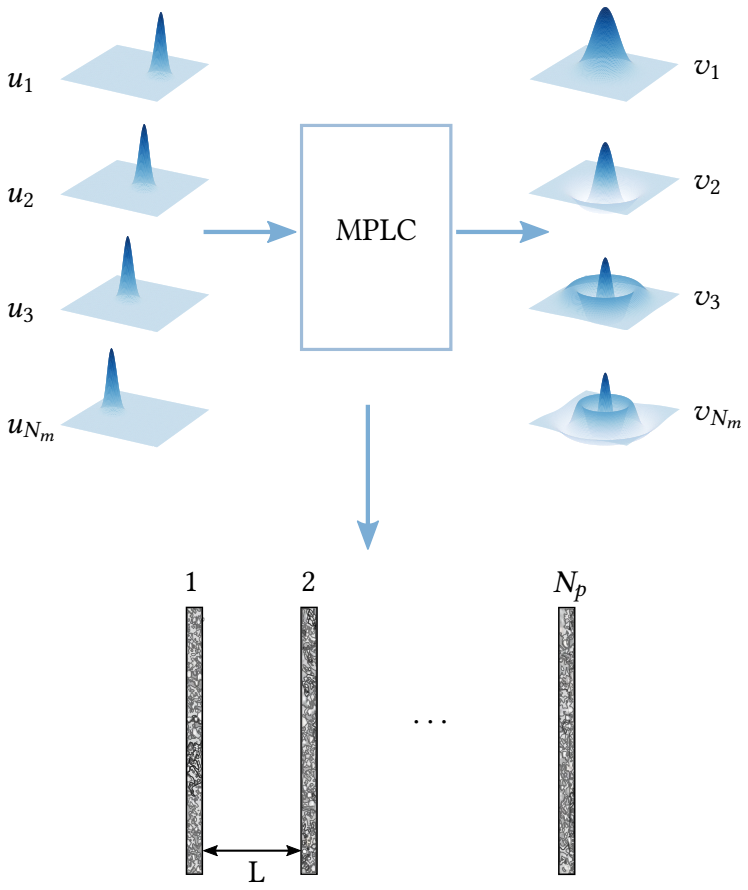

Figure 4.7: Steps of the construction of a MPLC system.

4.3.1 Construction of the MPLC transmission matrix

A MPLC system is characterized by a collection of inputs (see figure 4.7). First we define the number of modes to be shaped N_m and their spatial profiles: they are the input $\left(\{u_i\}_{1\leq i\leq N_m}\right)$ and output $\left(\{v_i\}_{1\leq i\leq N_m}\right)$ modes. The number of phase plates $\left(N_p\right)$ and their size ($m_x \times m_y$ pixels), as well as the length of the separation (L) between each phase plate are chosen accordingly. As a first step, we chose to work

with existing systems, meaning we used the data of MPLC systems which had already been simulated and constructed physically. In this way, we can be positive that the numerical propagation model we use is adequate since the agreement between the simulation and the physical system has already been tested. We use this model to build the transmission matrices of several different MPLC systems.

We present here the method we used to construct the transmission matrix of a MPLC system. As we presented in 4.2.1, a transmission matrix t relates N input modes with M output modes. We will in this part use the integers P and Q in place of N and M to avoid any confusion with the number of modes N_m or of phase plates N_p. We consequently first need to define a mode basis of size P to characterize the input of the MPLC system and similarly a mode basis of size Q for the output of the system. The mode basis we chose at the output of the system is the pixel mode basis $\left(\{\psi_i\}_{1\leq i\leq Q}\right)$ with $Q = m_x \times m_y$, with $m_x \times m_y$ the number of pixels of the actual phase plates. This mode basis is natural since it is the one which is used in the algorithm which calculates the phase plates of a MPLC system. It could also be used as the mode basis for the input of the system. However, in the system we use, Q is quite a large number: $150\,000 \leq Q \leq 400\,000$. The transmission matrix of the system is a matrix of size $Q \times P$. If one chooses the pixel mode basis for both the input and the output, we have $P = Q$ and the size the matrix t reaches makes it too big to be stored or processed. We therefore chose as the mode basis for the input of the system the $(\{\phi_i\}_{1\leq i\leq P})$ Hermite-Gauss mode basis. As was defined in chapter 1, the Hermite-Gauss mode basis is constructed as a product of n_x and n_y modes in the x and y directions, meaning $P = n_x \times n_y$. In practice, we choose values of P and Q such that $P \ll Q$. Typically, $P \simeq 650$ ($n_x = 15$ and $n_y = 43$). The choice of the Hermite-Gauss mode basis as the basis for the input of the system is motivated by the fact that in most of the systems we will consider, the input modes which defined the MPLC system, the modes $\left(\{u_i\}_{1\leq i\leq N_m}\right)$ can be well described by Hermite-Gauss modes. More than often, the input of MPLC systems are a collection of spatially separated Gaussian beams. We have shown in the previous chapters that Gaussian beams which see their spatial parameters vary from the design ones (displacement, tilt, waist size, defocus) can very well be described as a combination of a small number of Hermite-Gauss modes. We have mentioned in the beginning of the chapter that one of the motivation for the study we conduct here is to investigate how misalignment of the physical MPLC systems translates in terms of its transmission properties. Errors in the positioning of the input beams as well as in the distance or angle between the phase plate and the mirror can precisely be described as slightly modified Gaussian beams. We would like to stress that we do not have as much a priori information on how to describe the output modes $\left(\{\tilde{v}_i\}_{1\leq i\leq N_m}\right)$ of a misaligned MPLC system.

Nevertheless, the experience acquired by technicians at CAILabs tells us that such $\tilde{v}_i$ modes intensity profiles often resemble speckle patterns. This is the reason why we choose to keep the pixel mode basis as the mode basis at the output of the MPLC system, in order to keep the possibility of having a fine description of the transmitted output modes. Finally, we had that we have tested different types of mode basis and sizes of mode basis without noticing notable differences linked to the choice of the mode bases.

Thus, we constructed the transmission matrix t of a MPLC system by propagating modes of basis $(\{\phi_i\}_{1\leq i\leq P})$ through the complete system and projecting the result on basis $\left(\{\psi_i\}_{1\leq i\leq Q}\right)$. The matrix t schematically gives the link $\{\phi_i\} \rightarrow \{\psi_i\}$. We have seen in the previous section that a number of models derived for transmissions matrices give in fact information on the spectrum of $t^\dagger t$. In our case, $t^\dagger t$ is a square matrix of size $P \times P$. If we were to complete the mode bases $\{u_i\}$ and $\{v_i\}$ for them to respectively sizes P and Q we could also express the matrix t using these mode bases. In this case, by construction, the phase masks are optimized such that a sub-block of size $N_m \times N_m$ of $t^\dagger t$ is the identity matrix:

$$t^\dagger t = \begin{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} & & \\ & ? & \\ & & ? \end{bmatrix}. \tag{4.52}$$

Finally, we mention that another approach can be pursued to define the transmission matrix t of a MPLC system — although we did not use it. The constituting elements of a MPLC system can all be represented with their own transmission matrices. The free-space propagation on distance L between phase plates defines a transmission matrix t_0. Each phase plate sets a specific transmission matrix t_i with $i \in [1, N_p]$ (the matrices t_i were defined in 1). The total transmission matrix is

$$t = t_{N_p} t_0 t_{N_p-1} t_0 \cdots t_0 t_1. \tag{4.53}$$

4.3.2 Framework

The first step of our approach is to derive the spectral properties of the transmission matrices of existing MPLC systems. To study the spectral properties of a transmission matrix t, we perform the singular value decomposition of the matrix (SVD as

defined in 4.1.2.1):

$$t = UDV^{\dagger} \tag{4.54}$$

with U of size $Q \times Q$, V of size $P \times P$ and D a $Q \times P$ diagonal matrix. Hence, we have

$$t^{\dagger}t = V\left(D^{\dagger}D\right)V^{\dagger}. \tag{4.55}$$

We would like to understand if some of the models we introduced in the previous section appear to fit the distributions we observe for existing MPLC systems and if different regimes appear depending on the number of modes N_m or phase plates N_p. The development of such a theoretical model of the MPLC transmission matrices would allow us to identify the classes of transforms we can implement with MPLC systems.

In a second step, we wish to understand the influence of the following parameters on the transmission spectra of $t^{\dagger}t$ matrices:

- N_m the number of modes for which the system is designed
- N_p the number of phase plates
- L the distance between phase plates

Through this study we wish to investigate whether the parameters we use in the way we currently design MPLC systems are optimal or not. This optimality criterion can itself vary depending on the types of transforms which are implemented.

In order to meet realistic industrial or even physical manufacturing requirements, some limitations — such as relative smoothness of the phase plates — are integrated in the phase plate calculation algorithm. Another interesting area of investigation which we were not able to pursue in this thesis is the study of these fabrication constraints on the transmission properties of MPLC systems. However, the tools we developed can be used to study this question.

4.3.3 Properties of the transmission matrix

We plot in figure 4.8 the matrix $t^{\dagger}t$ of a MPLC system with $N_m = 10$ (and $N_p = 14$). The first observation we can make is that the matrix is almost diagonal. When looking at a close up of this matrix, we can see a block structure appearing. The input modes basis $\{\phi_i\}_{1 \leq i \leq P}$ is a Hermite-Gauss basis with number of modes n_x and n_y in the x and y directions respectively. The size of the blocks in $t^{\dagger}t$ is equal to $n_x \times n_x$ and reflects the way the input mode basis was constructed. This structure

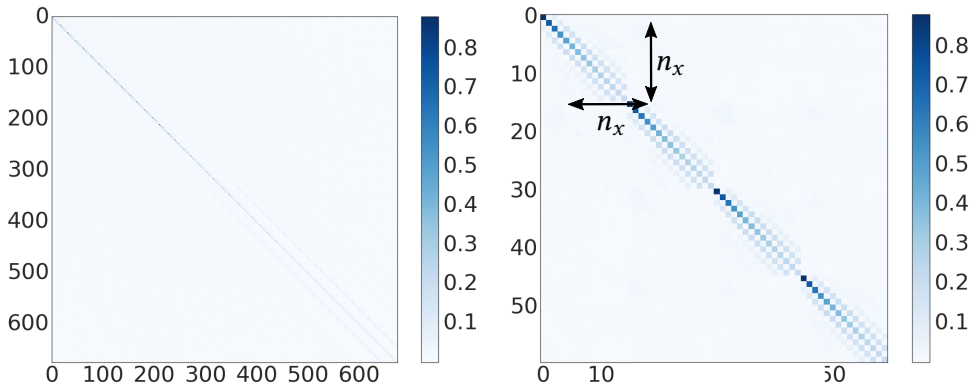

Figure 4.8: $t^\dagger t$ for a transform with $N_m = 10$ (and $N_p = 14$). On the right we plot the top-left corner of the $t^\dagger t$ matrix.

is found whatever the number of modes shaped N_m or the number of phase plates N_p.

The second interesting structure we can observe comes from the singular values of t. In figure 4.9 we plot the singular values for three different values of N_m ($N_m = 6, 10, 15$, the corresponding N_p values are 14, 14 and 20). We can observe the first N_m singular values stand out from the rest of the singular values: a gap appears. This feature is not surprising since the system is optimized to shape and transmit preferentially N_m modes. But this phenomenon is not systematic. On figure 4.10, we show a system with $N_m = 9$ (and $N_p = 20$) where no clear difference can be seen between the first N_m singular values and the rest of them. There is however one property which differentiates the systems of figure 4.9 and that of figure 4.10. In figure 4.9, the output modes are fiber modes, whereas in figure 4.10, the output modes are Laguerre-Gauss modes. In the case where the first N_m singular values stand out, this behaviour leads us to make the hypothesis that theses modes correspond to the "open" channels of this system — which raises the following question: we have seen in the previous section that open channels can be found in disordered and chaotic systems. Can we consider a MPLC system to be disordered or chaotic?

Finally, we observe on figure 4.9 that the overall repartition of the singular values also seems to change when the number of modes which is shaped increases.

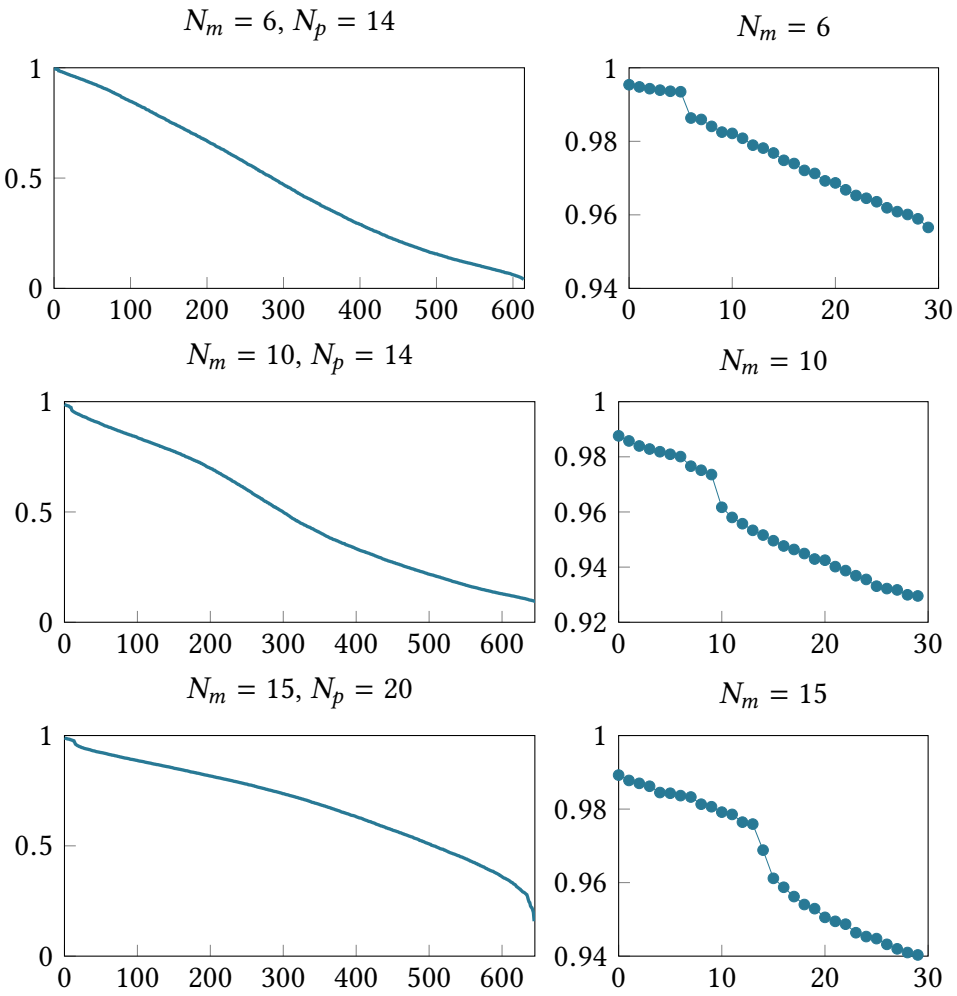

Figure 4.9: Singular values of different MPLC systems with $N_m = 6, 10, 15$ and $N_p = 14, 14, 20$. On the left we plot all the singular values while on the right we only plot the first 30 singular values.

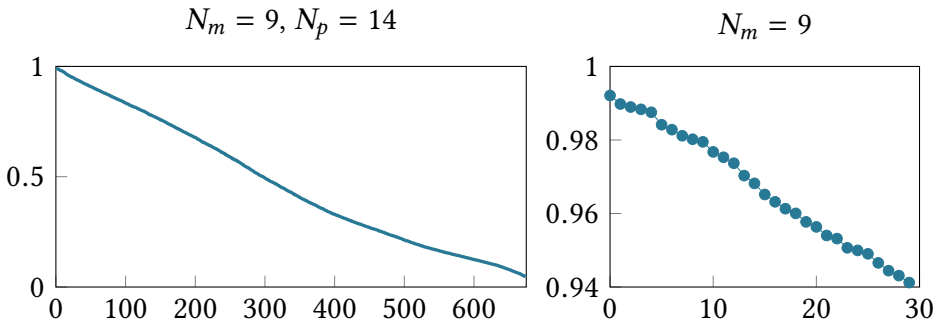

Figure 4.10: Singular values of a MPLC systems with $N_m = 9$ (and $N_p = 14$). On the left we plot all the singular values while on the right we only plot the first 30 singular values.

4.3.4 Comparison with theoretical models

We derive models for the transmission matrices of MPLC systems and compare their predictions to the actual distribution of the singular values of t.

4.3.4.1 Random Gaussian matrix

The first type of random matrix with which we can compare the matrix t is a random matrix whose entries have Gaussian distributions. In making this hypothesis, we do not impose any constraint of energy conservation on the system — since we know that in the case we consider (scattering without absorption), energy should be conserved. This is a coarse modeling of the system. The Marčenko-Pastur distribution describes the eigenvalue distributions of such matrices. Using a method described in [De Cock 99] to fit the Marčenko-Pastur law with a distribution, we compare the eigenvalue distribution for MPLC systems with Marčenko-Pastur distributions. We can see on figure 4.11 that the eigenvalue distribution of a MPLC system is not well described by the Marčenko-Pastur distribution. As a comparison, we plot the distribution of $h^\dagger h$ for h a simulated Gaussian random matrix and see that the fit is in this case very good.

4.3.4.2 Filtered matrix model for the MPLC system

We now introduce a filtered random matrix approach to derive the distribution of the eigenvalues of the matrix we measure. We recall that a scattering system can be modeled using a scattering matrix S. For a MPLC system, we give again the

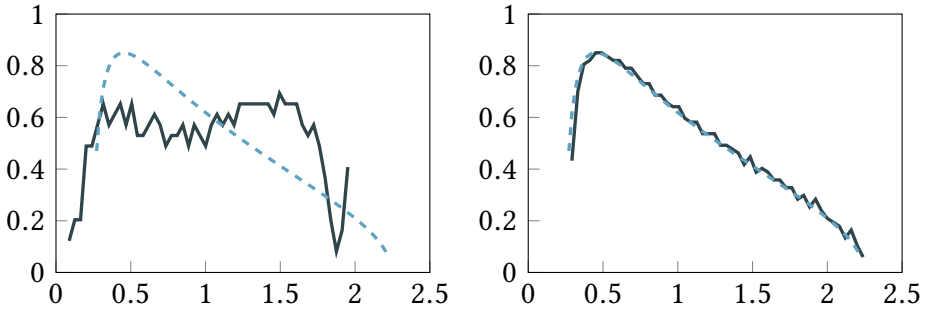

Figure 4.11: Histograms of the eigenvalues of a matrix $t^\dagger t$ (for $N_m = 15$ and $N_p = 20$) and of a random Gaussian matrix of the same size. The dotted line corresponds to the corresponding Marčenko-Pastur distribution.

expression of S:

$$S = \begin{pmatrix} r_0 & t_0' \\ t_0 & r_0' \end{pmatrix} \tag{4.56}$$

We start from the fact that the complete scattering matrix S is necessary unitary since a perfect MPLC system does not have losses. Energy conservation imposes: $S^\dagger S = 1$. Because S is unitary, the resolvent of $S^\dagger S$ is

$$g_0(z) = \frac{1}{N+M}\mathrm{Tr}\left\langle \frac{1}{z-1} \right\rangle = \frac{1}{z-1}. \tag{4.57}$$

The unitary character of S is a constraint that we impose to respect energy conservation. Since we have access to the resolvent of the matrix, we can derive the singular value distribution of matrix S. We define the projection matrices $\tilde{P}_1$ of size $(M+N)\times N$ and $\tilde{P}_2$ of size $M\times(N+M)$ such that

$$t_0 = \tilde{P}_2 S \tilde{P}_1. \tag{4.58}$$

We also introduce the parameters $m_1 = N/(N+M)$ and $m_2 = M/(N+M)$. However, we in fact do not have access to the whole transmission matrix but we measure only a limited number of the system's modes. The matrix we measure is t of size $N_2 \times N_1$[15]. We thus introduce the matrices P_1 of size $(M+N)\times N_1$ and P_2 of size $N_2\times(M+N)$. The matrix t is equal to

$$t = P_2 S P_1. \tag{4.59}$$

[15]In this derivation, $P = N_1$ and $Q = N_2$.

Finally, we make the assumption that the matrix S is random and that the filtering matrices P_1 and P_2 are free. In this case, a model exists to derive the singular value distribution of t. The general approach is detailed in [Goetschy 13] and the derivation presented here was made by the author of this work. We define $N(z)$ and $D(z)$ which are functions of $g(z)$:

$$N(z) = m_1 z\, g(z) + 1 - m_1 \tag{4.60}$$

$$D(z) = m_1 g(z) \left[z m_1 g(z) + m_2 - m_1 \right]. \tag{4.61}$$

The resolvent $g(z)$ of the matrix t can be obtained using the following implicit equation which relates D and N to the resolvent g_0 of matrix S

$$N(z)\, g_0\left(\frac{N^2(z)}{D(z)}\right) = D(z) \tag{4.62}$$

The solution to this equation system is

$$g^{\pm}(z) = \frac{m_1 - m_2 + (1-2m_1)\, z \pm \left[(m_1-m_2)^2 + z\,(z - 2m_1 - 2m_2 + 4m_1 m_2)\right]^{1/2}}{2m_1 z\,(1-z)}. \tag{4.63}$$

In the case where $m_1 + m_2 = 1$, this equation can be rewritten as

$$g^{\pm}(z) = \frac{1-\alpha}{2z} \pm \frac{1+\alpha}{2z}\frac{\sqrt{z-\tau_{min}}}{\sqrt{z-1}}. \tag{4.64}$$

[16] However, we are here interested in the case where the number of channels we control when measuring S is smaller than the total number of channels. We make

[16] **Filtered chaotic cavity** The distribution of the transmission coefficients of the scattering matrix of a chaotic cavity in the case where the number of transmission and reflection channels is not equal can be derived [Brouwer 96]. We define $\alpha = M/N$ as the ratio of the number of transmitted (M) and reflected (N) channels. The distribution writes:

$$\rho(\tau) = \frac{1+\alpha}{2\pi}\frac{\sqrt{(\tau-\tau_{min})^+}}{\tau\sqrt{1-\tau}} \quad \text{with} \quad \tau_{min} = \left[\frac{\alpha-1}{\alpha+1}\right]^2. \tag{4.65}$$

This distribution has a mean value $\overline{\tau} = \frac{\alpha}{1+\alpha}$ which is greater than $1/2$. In the limiting case where $N = M$, we find again the bimodal distribution of equation 4.45. The resolvent of this distribution can be shown to be

$$g(z) = \frac{1-\alpha}{2z} + \frac{1+\alpha}{2z}\frac{\sqrt{z-\tau_{min}}}{\sqrt{z-1}}. \tag{4.66}$$

We note that corresponds exactly to equation 4.64: in the case where the filtering matrices P_1 and P_2 are such that $m_1 + m_2 = 1$, the filtered matrix model presented in this section describes a chaotic cavity.

the assumption that $N = M$ to simplify the problem. Moreover, the case $N = M$ corresponds to the bimodal distribution with which we wish to compare our system. Since $t \neq t_0$, we introduce the quantities $\tilde{m}_1 = N_1/N$ and $\tilde{m}_2 = N_2/N$. The spectrum of $t^\dagger t$ is directly described by equation 4.63 with the parameters

$$m_1 = \frac{N_1}{2N} = \frac{\tilde{m}_1}{2} \tag{4.67}$$

$$m_2 = \frac{N_2}{2N} = \frac{\tilde{m}_2}{2} \tag{4.68}$$

The solution of this derivation gives the expression of the distribution of the eigenvalues of $t^\dagger t$:

$$\rho(\tau) = \frac{1}{\pi} \frac{\sqrt{(\tau^+ - \tau)(\tau - \tau^-)}}{\tilde{m}_1 \tau (1 - \tau)} + \max\left(1 - \frac{\tilde{m}_2}{\tilde{m}_1}, 0\right) \delta(\tau) \tag{4.69}$$

with

$$\tau^\pm = \frac{\tilde{m}_1 + \tilde{m}_2 - \tilde{m}_1 \tilde{m}_2 \pm \sqrt{\tilde{m}_1 \tilde{m}_2 (2 - \tilde{m}_1)(2 - \tilde{m}_2)}}{2}. \tag{4.70}$$

The model for the spectrum of $t^\dagger t$ given by equation 4.69 is the model we will use in the following part.

4.3.4.3 Comparison of the filtered matrix model with data

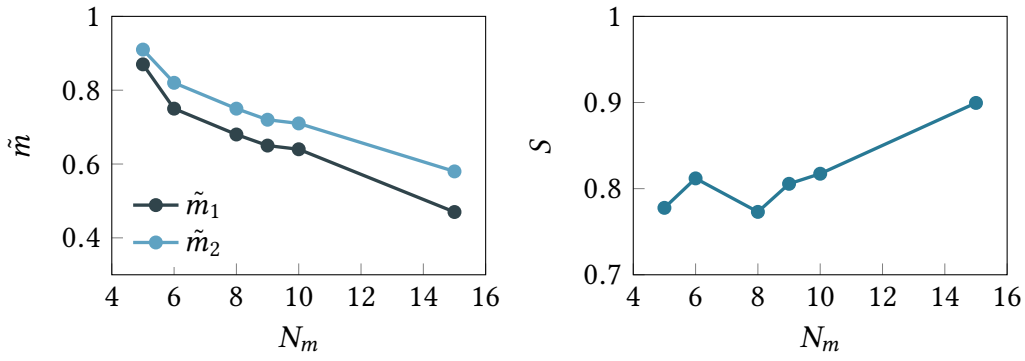

Figure 4.12: Fit coefficients between model 4.69 and eigenvalues distributions of actual MPLC systems. The corresponding plots are shown in 4.13. For all the systems, $N_p = 14$ except for $N_m = 15$ where $N_p = 20$.

We fitted the distribution model 4.69 with the actual eigenvalue distributions of $t^\dagger t$ we calculate. To perform the fitting of the $\tilde{m}_1$ and $\tilde{m}_2$ parameters, we used the similarity function defined in the supplementary material of [Froufe-Pérez 17]. We introduce a similarity function S which is defined as the area under the curve of the minimum between the data and the model curves. We normalize S by the maximum area of the two curves such that $S \in [0, 1]$. The results are plotted in figure 4.13. From equation 4.69 we can derive the expression of the singular value distribution of t. We also plot on figure 4.13 the singular value distributions and their fits with the parameters $\tilde{m}_1$ and $\tilde{m}_2$ of the eigenvalue fit. The MPLC systems have $N_m = 5, 6, 8, 9, 10$ and 15. For all the systems, $N_p = 14$ except for $N_m = 15$ where $N_p = 20$. The fit parameters $\tilde{m}_1$ and $\tilde{m}_2$ are presented in figure 4.12.

We first observe that as N_m grows, at constant N_p, the number of modes we "control" seems to decrease. This can be due to the fact as the system needs to shape more and more modes, the area of the pattern written on the phase plates grows, introducing more and more diffraction. Diffraction means that a portion of the light exits the spatial extent of the phase plates on which we measure the transmitted field. The second observation is that the fit for the singular values of t does not give very good results at small N_m while the quality of the fit improves as N_m grows. For $N_m = 15$, our model describes the distribution quite well.

Finally, we note that the singular value plots are easier to read than the eigenvalue ones: the peak around 0 of the eigenvalue distribution flattens the dynamic which make comparisons difficult. For this reason, we will in the next cases plot only the singular value plots.

4.3.4.4 Role of N_m

We investigate the role of the number of shaped modes N_m on the singular values distributions. We made two similar numerical experiments. We took two MPLC systems with $N_m = 10$ ($N_p = 14$) and $N_m = 15$ ($N_p = 20$). Each of these systems had associated parameters values (L, N_p, size of the phase plates $m_x \times m_y$). We kept those parameters constant and constructed a series of MPLC systems with different N_m. We then calculated their singular value distributions and compared these with our model (see figures 4.16 and 4.17). The corresponding fitting parameters are presented in figures 4.14 and 4.15. We observe that our derived model seems to fit best systems where the number of shaped modes N_m is large. This may be due to the fact that as the number of modes we shape increases, the disorder the system needs to introduce to shape the modes becomes more pronounced: hence, our model of random projections describes the system better in these cases. This analysis is

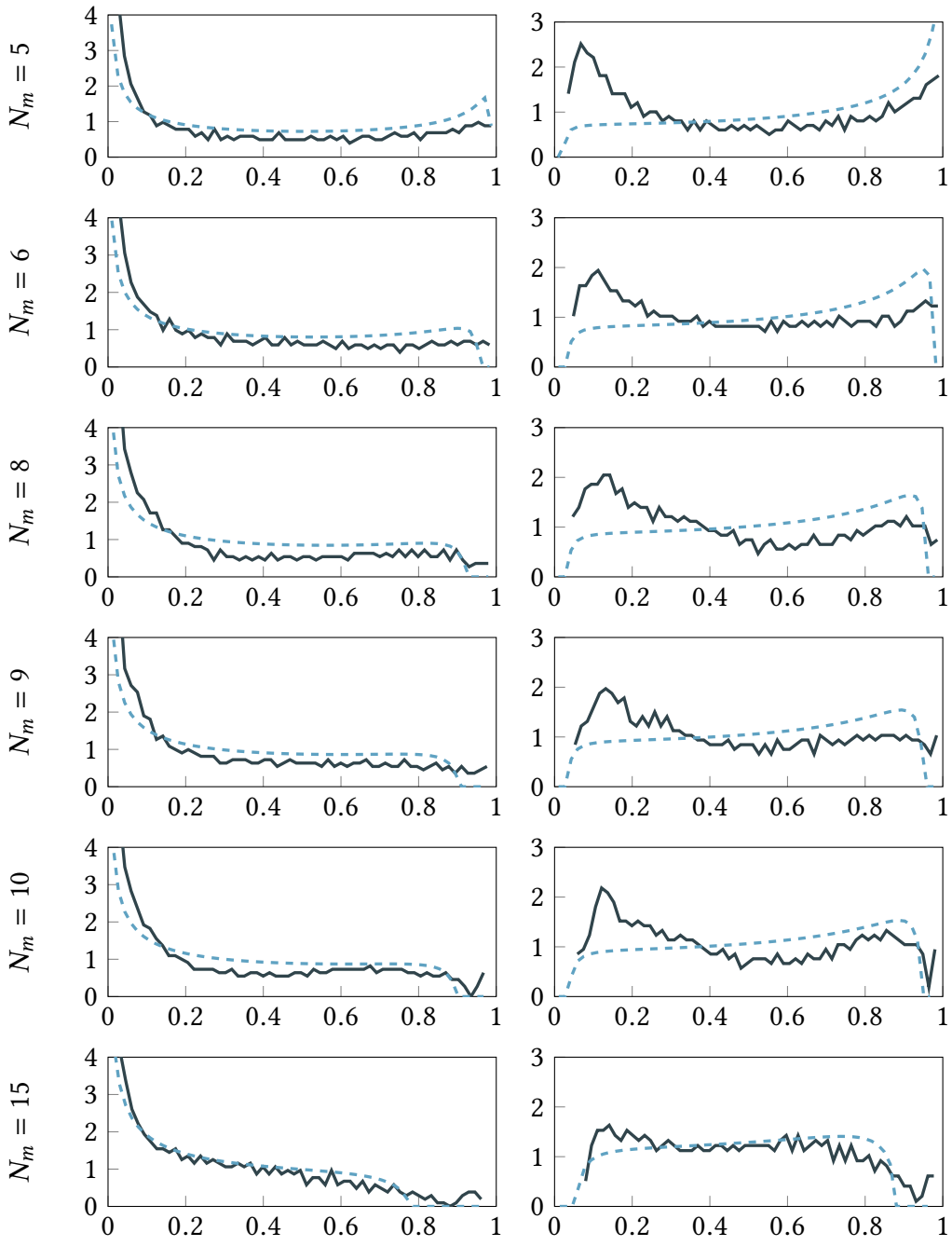

Figure 4.13: Full line: distribution of the eigenvalues of $t^\dagger t$ (left) and singular values of t (right) for MPLC systems with different values of N_m. Dashed line: fit of the filtered matrix model. The coefficients of the fit are plotted in 4.12. For all the systems, $N_p = 14$ except for $N_m = 15$ where $N_p = 20$.

supported by the fact that the fits are better in the case of the system with $N_m = 15$ than in that with $N_m = 10$.

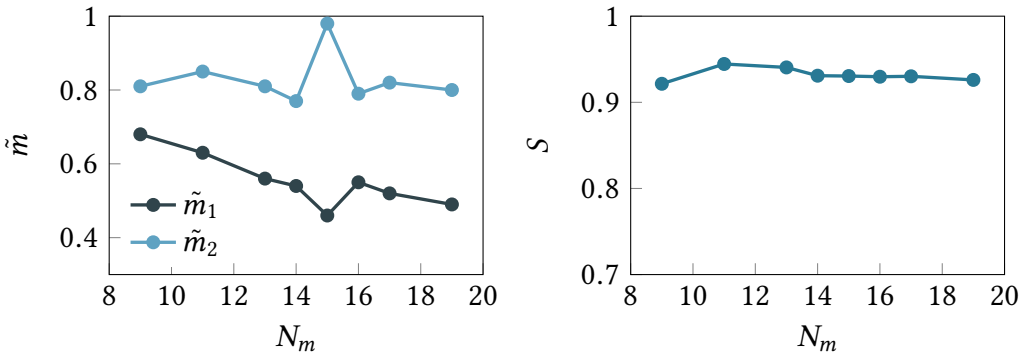

Figure 4.14: Reference model $N_m = 15$ ($N_p = 20$). Fit coefficients between model 4.69 and eigenvalues distributions of actual MPLC systems. The corresponding plots are shown in 4.16.

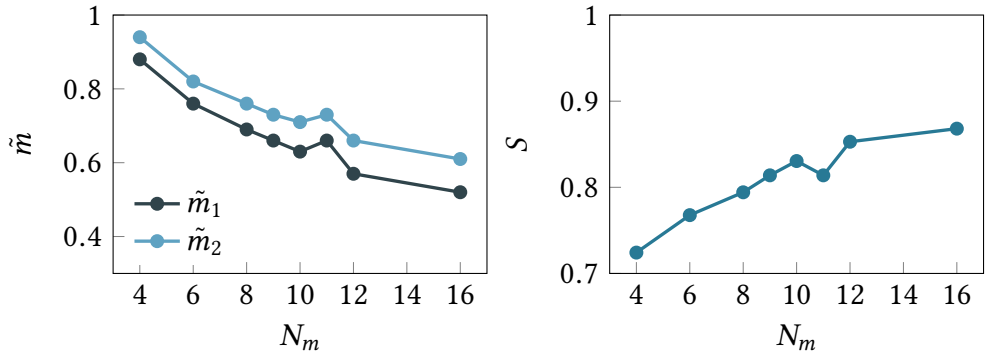

Figure 4.15: Reference model $N_m = 10$ ($N_p = 14$). Fit coefficients between model 4.69 and eigenvalues distributions of actual MPLC systems. The corresponding plots are shown in 4.17.

4.3.4.5 Role of N_p

We also investigate the role of the number of shaped modes N_m on the singular values distributions. In this case we took a MPLC system with $N_m = 6$. We constructed a series of MPLC systems with a constant number of shaped modes $N_m = 6$

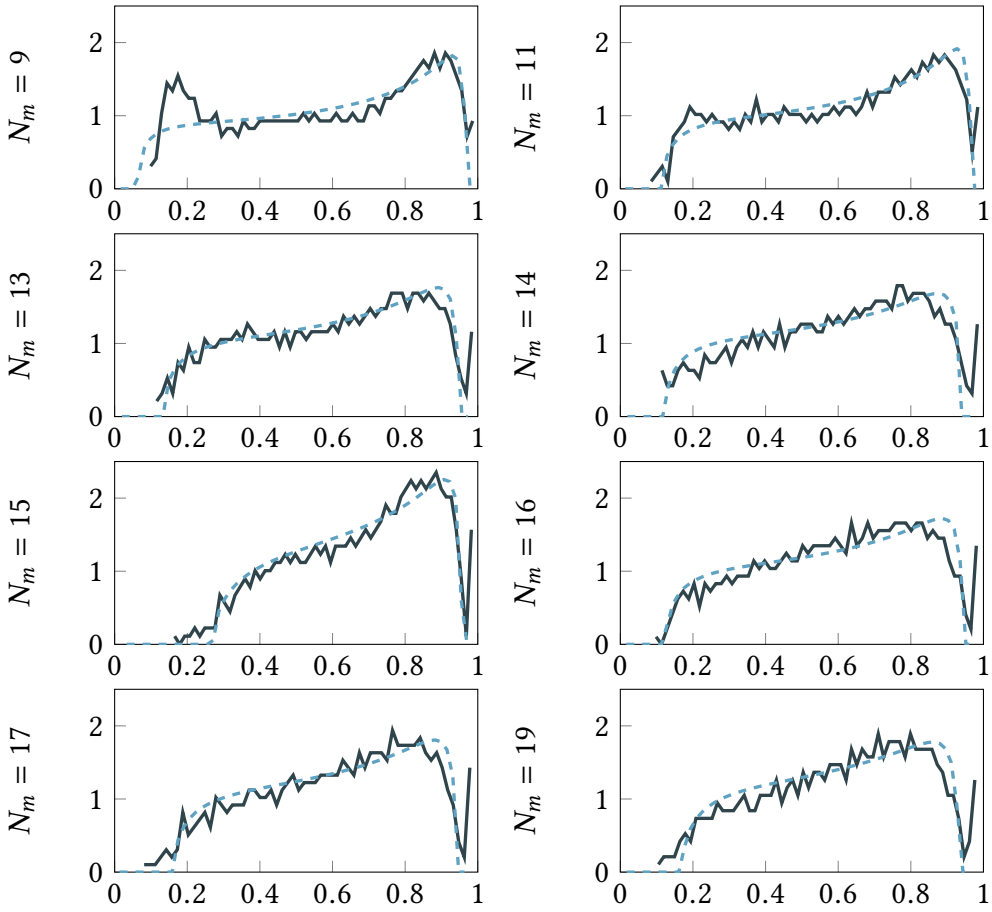

Figure 4.16: Reference model $N_m = 15$ ($N_p = 20$). Full line: distribution of the singular values of t (right). Dashed: fit of model.

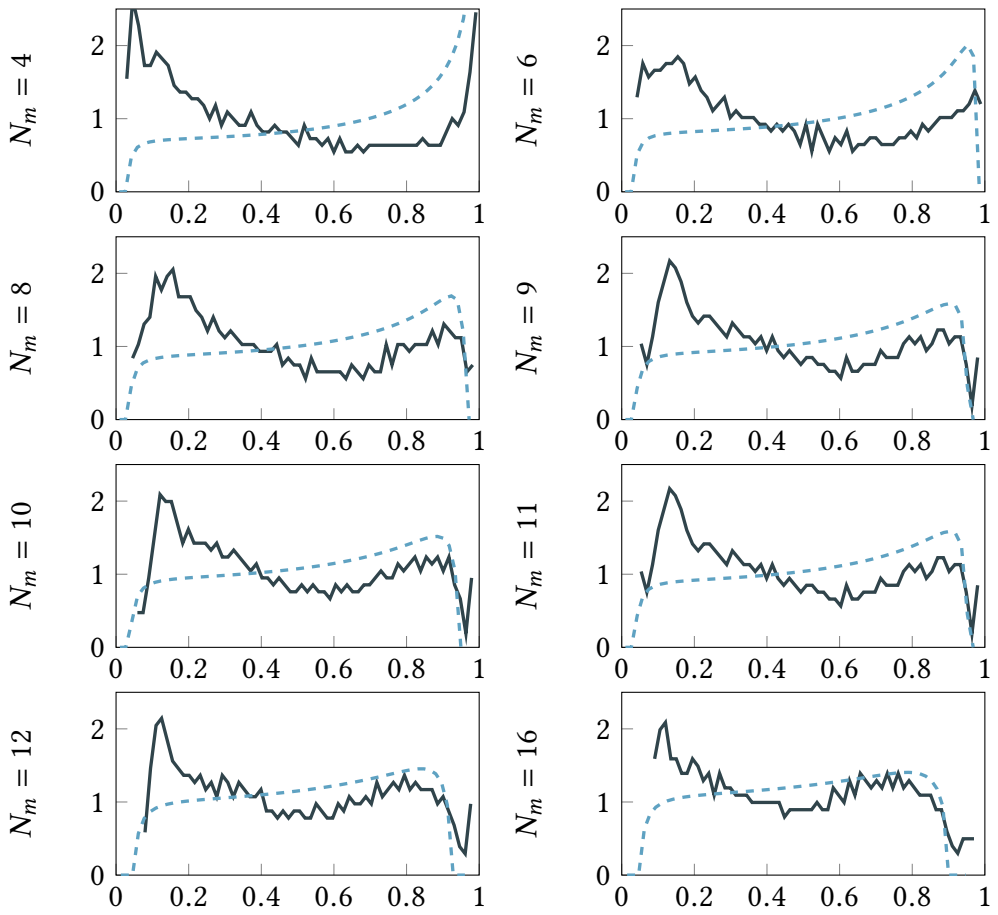

Figure 4.17: Reference model $N_m = 10$ ($N_p = 14$). Full line: distribution of the singular values of t (right). Dashed: fit of model.

but with varying number of phase plates N_p. We calculated the corresponding singular values distributions and compared these with our model (see figures 4.19). The corresponding fitting parameters are presented in figures 4.18. It seems that our model fits best when the number of phase plates is not too large compared with the number of shaped modes. However, in this case the fits are not overall very good. This can probably be explained by the fact that the reference number of shaped modes $N_m = 6$ is small.

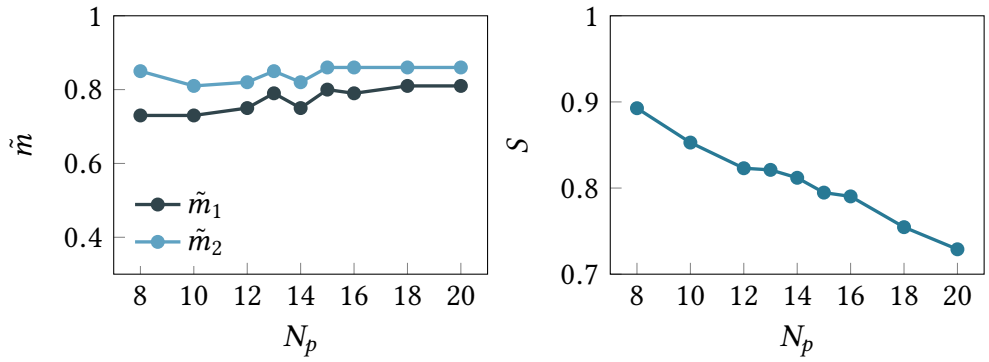

Figure 4.18: Fit coefficients between model 4.69 and eigenvalues distributions of actual MPLC system with $N_m = 6$. The corresponding plots are shown in 4.19.

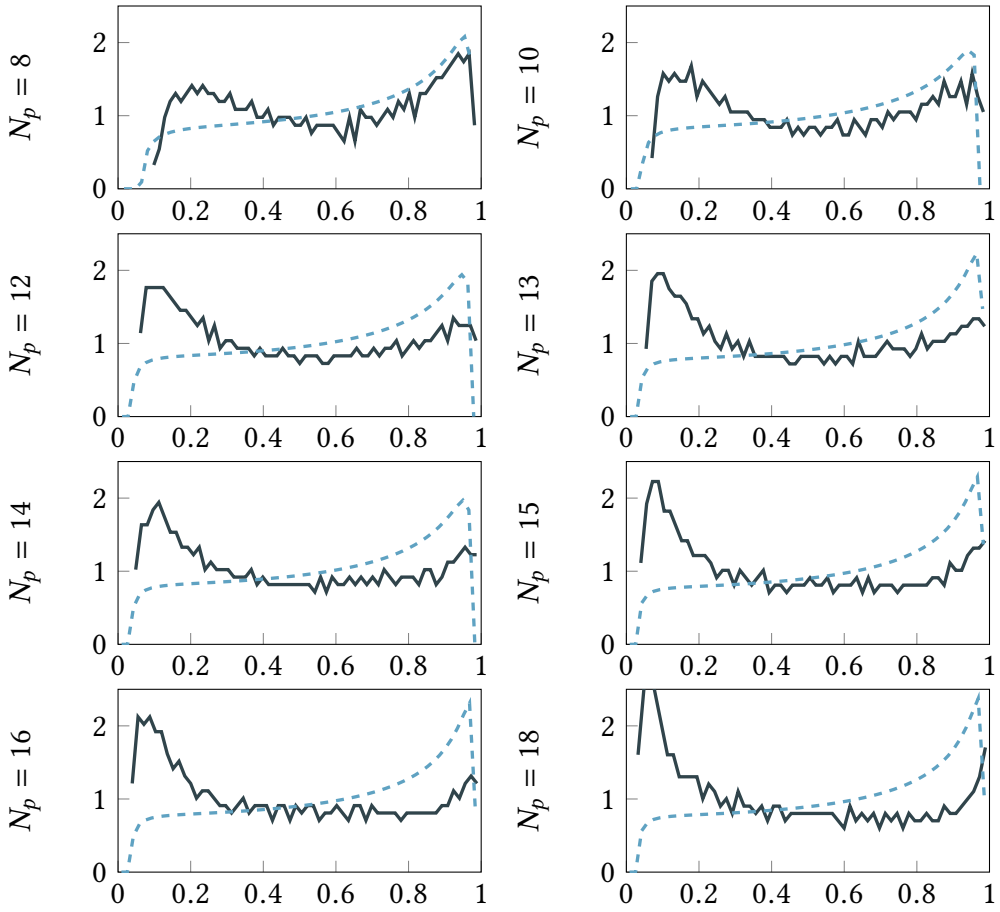

Figure 4.19: Full line: distribution of the singular values of t (right) of an MPLC system with $N_m = 6$. Dashed: fit of model.

4.4 Conclusion

In this chapter our aim was to study the transmission properties of the MPLC system at large and to compare this object to models of scattering in complex media. We have constructed a numerical model of the MPLC system which allows to construct a numerical matrix describing its transmission properties. We observe significant statistical differences in the singular value distribution of the transmission matrix when the ratio between the number of phase plates and that of shaped modes is changed. We derive a model for the transmission matrices singular value distributions using a filtered matrix approach. This model gives good results when the number of shaped modes N_m is sufficiently large, but does not describe the system well for small values. The number of phase plates N_p seems to be linked to the amount of losses in the system: more reflections induce more losses. The fact that the values of the parameters which describe actual MPLC systems N_p and N_m are relatively small ($\leq$ 20) hinders the capacity of random matrix theory to fully describe these systems. Indeed, the distributions derived in the context of random matrix theory are asymptotical ones. Moreover, random matrix theory usually studies objects which are averaged over large numbers of realization. As the algorithm which calculates a given MPLC system is deterministic, we cannot perform such ensemble averaging on the transmission matrices we calculate. However, the modelisation we devised to calculate the transmission matrix t is in its own a useful tool for the study of the MPLC properties. It carries a lot of phenomenological information about the MPLC system it models and is a useful tool to study the influence of the different construction parameters.

A second approach which we did not have time to complete was to model the MPLC systems as a concatenation of unit cells (as is described in 4.53) and to study the properties of one unit cell (phase plate + free-space propagation). As we have seen, using free probability theory, the derivation of the distribution of the product of matrices can easily be derived from that of the distribution of each part of the product. This approach may hold interesting results for the analysis of the transmission properties of MPLC systems.

Appendices

Appendix A

Bessel beam generation

We attach in this appendix the published review paper describing the generation and characterization of a record angle Bessel beam.

Research Article Vol. 57, No. 23 / 10 August 2018 / Applied Optics 6725

applied optics

Generation of high conical angle Bessel–Gauss beams with reflective axicons

Pauline Boucher,[1,2] Jesus Del Hoyo,[3] Cyril Billet,[3] Olivier Pinel,[1] Guillaume Labroille,[1] and François Courvoisier[3,*]

[1]Laboratoire Kastler Brossel, Sorbonne Université, CNRS, PSL Research University, Collège de France, 4 Place Jussieu, 75005 Paris, France
[2]CAILabs, 38 Boulevard Albert 1er, 35200 Rennes, France
[3]FEMTO-ST Institute, Univ. Bourgogne Franche-Comté, CNRS, 15B Avenue des Montboucons, 25030 Besançon, Cedex, France
*Corresponding author: francois.courvoisier@femto-st.fr

Received 25 April 2018; revised 6 July 2018; accepted 13 July 2018; posted 16 July 2018 (Doc. ID 330339); published 6 August 2018

We report the generation of Bessel–Gauss beams of high conical angle, up to 35 deg, using reflective off-axis axicons and a magnification optical system. We experimentally characterize the beams with three-dimensional scans. The high precision of fabrication of the axicons in the vicinity of the axicon singularity allows us to generate a beam with intensity distribution close to analytical description.

OCIS codes: (070.7345) Wave propagation; (070.3185) Invariant optical fields; (080.4228) Nonspherical mirror surfaces.

https://doi.org/10.1364/AO.57.006725

1. INTRODUCTION

Diffraction-free Bessel beams were discovered by Durnin *et al.* in 1987 [1]. They are formed out of a conical interference of an infinite number of plane waves, crossing the optical axis at the same angle, here referred to as the *conical angle*. The constructive interference on the optical axis produces an intense central spot surrounded by several cylindrically symmetric lobes with lower intensity. This beam structure is diffraction-free since its intensity profile is invariant with propagation. Obviously, finite energy beams preserve this structure only over a finite distance [2]. Bessel–Gauss beams are one of those finite energy realizations of Bessel beams. In a Bessel–Gauss beam, the interference is formed after a conical phase is applied on a Gaussian beam [3,4]. In this realization, the on-axis intensity profile evolves smoothly with a bell-shaped profile [2,4,5], in contrast with other forms of finite energy Bessel beams where the on-axis intensity oscillates, such as in the first realization by Durnin *et al.*

Bessel beams have found a variety of applications in optics. They are used as optical traps [6], for light-sheet microscopy [7,8], nonlinear optics and filamentation [9–11], for the guiding of electric discharges [12,13], material processing [14–17], or even for cellular transfection [18].

Bessel–Gauss beams are either produced from transmissive axicons (discovered well before the diffraction-free property [19]) or reflective spatial light modulators. While the first enable shaping of higher-power beams, the latter enable higher-quality beams but at the cost of very low cone angles. Two difficulties arise in this field. On the one hand, the low conical angle values are compensated by telescopic magnification, but at the cost of a reduction in beam length by the square of the magnification if high angles are needed for applications. On the other hand, beam shaping with bulk glass axicons usually encounters another difficulty which is the processing of the tip, which cannot be infinitely sharp. The roundness produces a spherical wave which interferes with the Bessel beam. This generates deleterious oscillations of the on-axis intensity [20]. Several techniques have been developed to circumvent the impact of blunt tips, such as Fourier filtering [20] or using liquid immersion [21], but they are not energy efficient or easily implementable.

Here we report high-quality Bessel–Gauss beam generation with reflective off-axis axicons where spatial filtering is unnecessary. These are compatible with high powers and produce cone angles sufficiently high that a magnification factor of only ×55 is required to generate the highest conical angles (35°) produced to the best of our knowledge. This relatively low value of magnification also enables and generates Bessel beams with length exceeding 100 μm.

2. REFLECTIVE AXICONS

Our experimental setup is based on reflective axicons in oblique illumination to enable further use with high-power lasers. In this case, it is necessary to engineer the surface to generate off-axis beams and avoid potential distortion due to the oblique illumination [22]. The theoretical phase profile for the axicons we generated follows [23]: $\Phi(x,y) = (2\pi/\lambda)\tan\beta\sqrt{\cos^2(\alpha)x^2 + y^2}$, where α is the incidence angle of the beam on the axicon, and $\beta = \theta/2$, where θ is the conical

1559-128X/18/236725-04 Journal © 2018 Optical Society of America

6726 Vol. 57, No. 23 / 10 August 2018 / Applied Optics Research Article

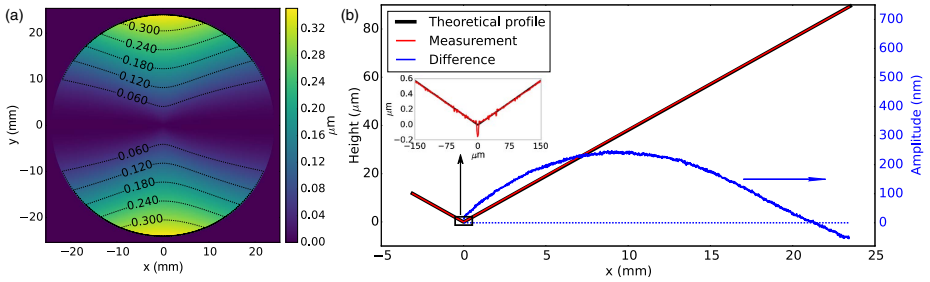

Fig. 1. Characterization of axicon $\theta_1' = 3.842$ mrad. (a) Difference between an "on-axis" axicon and the off-axis axicon produced. (b) Comparison between theoretical and measured axicon profiles.

angle of the Bessel beam produced. The factor 1/2 simply arises from the reflection, which doubles the optical path. The axicons were produced by lithography from a metallic substrate, which has been coated by a broadband reflective dielectric coating at a central wavelength of 800 nm for further use with high-power femtosecond Ti:sapphire lasers. Before coating, the RMS surface roughness of the processed metallic substrate is 4.4 nm.

Three different axicons with different conical angles were produced and tested, with angles $\theta_1' = 3.842$ mrad, $\theta_2' = 4.546$ mrad, and $\theta_3' = 5.215$ mrad. After magnification of 1/55, as we will see later, these axicons allowed us to produce Bessel–Gauss beams with conical angles of, respectively, $\theta_1 = 25°$, $\theta_2 = 30°$, and $\theta_3 = 35°$. Figure 1(a) shows the difference between the theoretical profile of the "on-" and "off-" axis axicons corresponding to the cone angle of θ_1 ($\alpha = 5°$).

In Fig. 1(b), we show the surface profile characterization by interferometry and its comparison with the theoretical profile. The shape error on the profile is less than 250 nm. In particular, we plot in the inset a magnified view around the tip. We observe that the size of the region exhibiting a significant profile difference from the theoretical profile is smaller than 50 μm. This was also the case for the other two axicons produced.

3. EXPERIMENTAL SETUP

We tested the axicons one by one using the same experimental setup, which is shown in Fig. 2. The optical source is a laser diode with central wavelength of $\lambda = 780$ nm. The input beam is a single-mode Gaussian beam with a waist of (radius at $1/e^2$) $w_0 \sim 3$ mm. The optical path was carefully adjusted so that the axicon illumination angle is precisely $\alpha = 5°$. After incidence on the axicon, the beam is magnified by two successive confocal telescopes, with magnifications factors respectively of $\gamma_1 = 2$ and $\gamma_2 = 1/110$ in order to produce a total magnification of $\gamma = 1/55$. The produced beam was then imaged onto a CCD camera with an imaging system of magnification ×56 combining a ×50 microscope objective with 3.6 mm focal length and a lens of focal length 200 mm. The imaging microscope objective was mounted on a motorized translation stage so as to scan across the Bessel beam we generated.

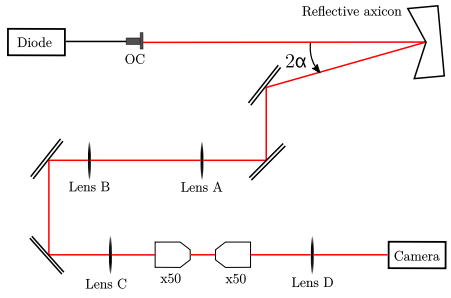

Fig. 2. Experimental setup (OC, output coupler). Lens A and lens B make up the first telescope and have focal lengths of $f_A = 200$ mm and $f_B = 400$ mm. Lens C and lens D have focal lengths of $f_C = 400$ mm and $f_D = 200$ mm. The pixel pitch of the camera is 4.65 μm.

4. RESULTS

In Fig. 3, we plot our results for the three different axicons used. For each Bessel–Gauss beam, we represented the cross section (x, y) of the transverse intensity profile at the propagation distance corresponding to the peak of intensity and a cross section (x, z) of the longitudinal intensity profile. In all three cases, the high quality of the beam is apparent. Only a minor deviation from perfect circular symmetry of the transverse profile is observed and attributed to a small misalignment. Table 1 summarizes the experimental beam parameters.

The evolution of the on-axis intensity with propagation distance is an important parameter qualifying the quality of the Bessel–Gauss beam. In Fig. 4, we compare the on-axis intensity $I(r = 0, z)$ of the beams, shown as a black curve with the theoretical profile (red curve) obtained from stationary phase approximation of the diffraction integral [4,15]: $I(z) = 8\pi P_0 z \sin^2\theta/(\lambda w^2) \exp[-2(z \sin\theta/w)^2]$. θ is the conical angle, $w = 3187$ μm is the waist of the incident Gaussian beam (value obtained by fitting the experimental results), and P_0 is the input peak power. We compared the

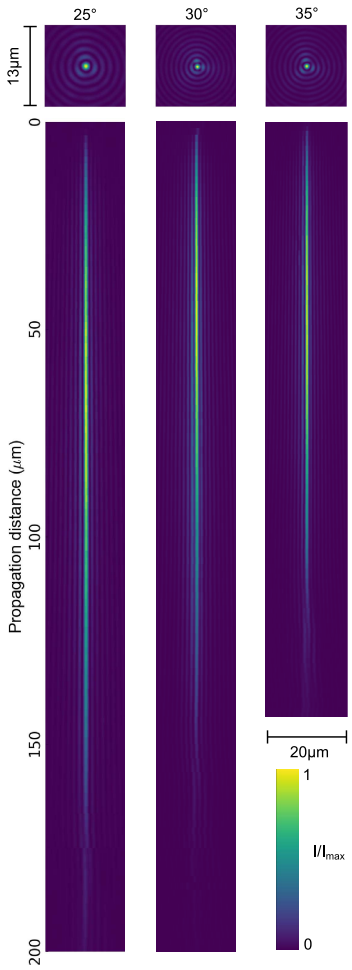

Fig. 3. Experimental images of transverse energy distribution and cross section of beams of 25°, 30°, and 35° conical angles.

Table 1. Theoretical Beam Parameters

Conical Angle (°)	Central Lobe Diameter (FWHM) (μm)	Bessel Zone Length (measured at FWHM) (μm)
25	0.67	130
30	0.58	110
35	0.51	95

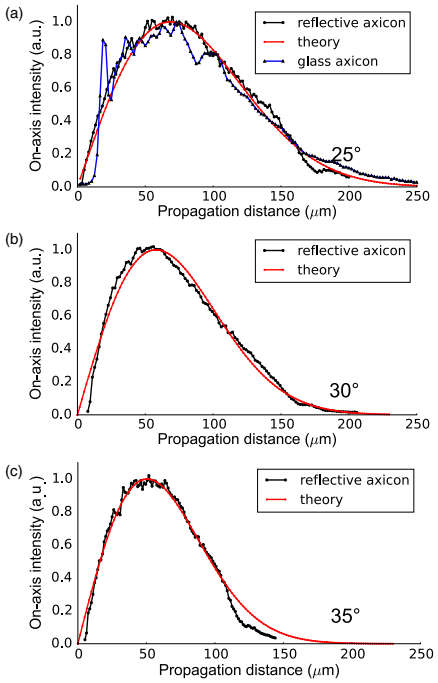

Fig. 4. (a) 25° conical angle: comparison between the on-axis intensity evolution along z for a glass axicon in transmission (blue, triangles), a reflective axicon (black, disks), and the theoretical profile for a Bessel–Gauss beam of the same conical angle (red). (b) 30° and (c) 35° conical angles: comparison between the measured on-axis intensity evolution and the analytical description of the on-axis intensity associated with such conical angles.

normalized profiles. An excellent agreement is found between the experimental data and the analytical description of Bessel–Gauss beams.

We note that commercially available transmission axicons present a larger deviation from the perfect axicon because of the manufacturing of the tip of the axicon. In Fig. 4(a), we compare the profile obtained with a glass transmissive convex axicon with angle 0.5° (from Thorlabs) associated with a magnification system of 1/110 with the same input Gaussian beam. The generated conical angle is $\theta = 26°$, thus comparable to the first axicon. We observe that the on-axis intensity profile exhibits much larger deviations to the Bessel–Gauss profile expression, which is attributed to its blunt tip.

5. CONCLUSION

In conclusion, we have demonstrated experimental generation of high-quality Bessel–Gauss beams with high conical angles, up to 35°. We used reflective off-axis axicons exhibiting high

6728 Vol. 57, No. 23 / 10 August 2018 / Applied Optics Research Article

surface quality in excellent agreement with the target surface profile and with a very reduced imperfect zone around the axicon singularity. We anticipate that these results will enable novel applications for high-power ultrafast optics, such as nonlinear optics or laser micromachining.

Funding. H2020 European Research Council (ERC) (682032-PULSAR); Agence Nationale de la Recherche (ANR) (ANR-11-LABX-0001-01).

REFERENCES

1. J. Durnin, J. J. Miceli, and J. H. Eberly, "Diffraction-free beams," Phys. Rev. Lett. **58**, 1499–1501 (1987).
2. D. McGloin and K. Dholakia, "Bessel beams: diffraction in a new light," Contemp. Phys. **46**, 15–28 (2005).
3. F. Gori, G. Guattari, and C. Padovani, "Bessel-Gauss beams," Opt. Commun. **64**, 491–495 (1987).
4. V. Jarutis, R. Paškauskas, and A. Stabinis, "Focusing of Laguerre-Gaussian beams by axicon," Opt. Commun. **184**, 105–112 (2000).
5. T. Čižmár and K. Dholakia, "Tunable Bessel light modes: engineering the axial propagation," Opt. Express **17**, 15558–15570 (2009).
6. E. Mcleod and C. B. Arnold, "Subwavelength direct-write nanopatterning using optically trapped microspheres," Nat. Nanotechnol. **3**, 413–417 (2008).
7. F. O. Fahrbach, P. Simon, and A. Rohrbach, "Microscopy with self-reconstructing beams," Nat. Photonics **4**, 780–785 (2010).
8. T. A. Planchon, L. Gao, D. E. Milkie, M. W. Davidson, J. A. Galbraith, C. G. Galbraith, and E. Betzig, "Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination," Nat. Methods **8**, 417–423 (2011).
9. E. Gaizauskas, E. Vanagas, V. Jarutis, S. Juodkazis, V. Mizeikis, and H. Misawa, "Discrete damage traces from filamentation of Gauss-Bessel pulses," Opt. Lett. **31**, 80–82 (2006).
10. D. E. Roskey, M. Kolesik, J. V. Moloney, and E. M. Wright, "Self-action and regularized self-guiding of pulsed Bessel-like beams in air," Opt. Express **15**, 9893–9907 (2007).
11. D. Faccio, E. Rubino, A. Lotti, A. Couairon, A. Dubietis, G. Tamošauskas, D. G. Papazoglou, and S. Tzortzakis, "Nonlinear light-matter interaction with femtosecond high-angle Bessel beams," Phys. Rev. A **85**, 033829 (2012).
12. M. Scheller, N. Born, W. Cheng, and P. Polynkin, "Channeling the electrical breakdown of air by optically heated plasma filaments," Optica **1**, 125–128 (2014).
13. M. Clerici, Y. Hu, P. Lassonde, C. Milian, A. Couairon, D. N. Christodoulides, Z. Chen, L. Razzari, F. Vidal, F. Legare, D. Faccio, and R. Morandotti, "Laser-assisted guiding of electric discharges around objects," Sci. Adv. **1**, e1400111 (2015).
14. A. Marcinkevičius, S. Juodkazis, S. Matsuo, V. Mizeikis, and H. Misawa, "Application of Bessel beams for microfabrication of dielectrics by femtosecond laser," Jpn. J. Appl. Phys. **40**, L1197–L1199 (2001).
15. M. K. Bhuyan, F. Courvoisier, P. A. Lacourt, M. Jacquot, R. Salut, L. Furfaro, and J. M. Dudley, "High aspect ratio nanochannel machining using single shot femtosecond Bessel beams," Appl. Phys. Lett. **97**, 081102 (2010).
16. M. Duocastella and C. Arnold, "Bessel and annular beams for materials processing," Laser Photon. Rev. **6**, 607–621 (2012).
17. F. Courvoisier, R. Stoian, and A. Couairon, "Ultrafast laser micro- and nano-processing with nondiffracting and curved beams. Invited paper for the section: hot topics in ultrafast lasers," Opt. Laser Technol. **80**, 125–137 (2016).
18. X. Tsampoula, V. Garceés-Chavez, M. Comrie, D. J. Stevenson, B. Agate, C. T. A. Brown, F. Gunn-Moore, and K. Dholakia, "Femtosecond cellular transfection using a nondiffracting light beam," Appl. Phys. Lett. **91**, 053902 (2007).
19. J. H. McLeod, "The axicon: a new type of optical element," J. Opt. Soc. Am. **44**, 592–597 (1954).
20. O. Brzobohatý, T. Cižmár, and P. Zemánek, "High quality quasi-Bessel beam generated by round-tip axicon," Opt. Express **16**, 12688–12700 (2008).
21. S. Akturk, C. L. Arnold, B. Prade, and A. Mysyrowicz, "Generation of high quality tunable Bessel beams using a liquid-immersion axicon," Opt. Commun. **282**, 3206–3209 (2009).
22. Z. Bin and L. Zhu, "Diffraction property of an axicon in oblique illumination," Appl. Opt. **37**, 2563–2568 (1998).
23. A. Thaning, Z. Jaroszewicz, and A. T. Friberg, "Diffractive axicons in oblique illumination: analysis and experiments and comparison with elliptical axicons," Appl. Opt. **42**, 9–17 (2003).

Appendix B

Continuous axial scanning of a Gaussian beam via beam steering.

We attach in this appendix the published review paper describing the results of chapter 3.

Research Article | Vol. 25, No. 19 | 18 Sep 2017 | OPTICS EXPRESS 23060
Optics EXPRESS

Continuous axial scanning of a Gaussian beam via beam steering

PAULINE BOUCHER,[1,2] NICOLAS BARRÉ,[2] OLIVIER PINEL,[2] GUILLAUME LABROILLE,[2] AND NICOLAS TREPS[1]

[1] *Laboratoire Kastler Brossel, UPMC - Sorbonne Universités, CNRS, ENS-PSL Research University, Collège de France - 4 place Jussieu, Paris, France*
[2] *CAILabs - 38 Boulevard Albert 1er, Rennes, France*
[*] *pauline.boucher@lkb.upmc.fr*

Abstract: We propose and demonstrate experimentally the transfer of one spatial degree of freedom of a laser beam onto another one. Using a multi-plane light conversion device (MPLC) and a modal analysis, we designed a passive setup with immediate response which couples a displacement and tilt in the transverse plane to a longitudinal shift of the focus point of a beam. With this design, we demonstrated a shift of the focal point of the output beam by 4 z_R along the propagation axis.

OCIS codes: (230.0230) Optical devices; (110.0110) Imaging systems.

References and links

1. D. Reddy and P. Saggau, "Development of a random access multiphoton microscope for fast three-dimensional functional recording of neuronal activity,", Proc. SPIE **6443**, 64430U (2007).
2. G. Sancataldo, L. Scipioni, T. Ravasenga, L. Lanzanò, A. Diaspro, A. Barberis, and M. Duocastella, "Three-dimensional multiple-particle tracking with nanometric precision over tunable axial range," Optica **4**, 367-373 (2017).
3. G. Druart, J. Taboury, N. Guérineau, R. Haïdar, H. Sauer, A. Kattnig, and J. Primot, "Demonstration of image-zooming capability for diffractive axicons," Opt. Lett. **33**, 366-–368 (2008).
4. V. X. D. Yang, Y. Mao, B. A. Standish, N. R. Munce, S. Chiu, D. Burnes, B. C. Wilson, I. A. Vitkin, P. A. Himmer, and D. L. Dickensheets, "Doppler optical coherence tomography with a micro-electro-mechanical membrane mirror for high-speed dynamic focus tracking," Opt. Lett. **31**, 1262–1264 (2006).
5. T. Shibaguchi and H. Funato, "Lead-lanthanum zirconate-titanate (PLZT) electrooptic variable focal-length lens with stripe electrodes," Japanese J. Appl. Phys. **31**, 3196–3200 (1992).
6. A. Kaplan, N. Friedman, and N. Davidson, "Acousto-optic lens with very fast focus scanning," Opt. Lett. **26**, 1078–1080 (2001).
7. H.-C. Lin and Y.-H. Lin, "A fast response and large electrically tunable-focusing imaging system based on switching of two modes of a liquid crystal lens," Appl. Phys. Lett. **97**(6), 063505 (2010).
8. S. Sato, A. Sugiyama, and R. Sato, "Variable-Focus Liquid-Crystal Fresnel Lens," Japanese J. Appl. Phys. **24**(8A), L626 (1985).
9. G. Zhu, J. van Howe, M. Durst, W. Zipfel, and C. Xu, "Simultaneous spatial and temporal focusing of femtosecond pulses," Opt. Express **13**, 2153–2159 (2005).
10. M. E. Durst, G. Zhu, and C. Xu, "Simultaneous spatial and temporal focusing for axial scanning," Opt. Express **14**, 12243–12254 (2006).
11. R. Du, K. Bi, S. Zeng, D. Li, S. Xue, and Q. Luo, "Analysis of fast axial scanning scheme using temporal focusing with acousto-optic deflectors," J. Mod. Opt. **56**, 81–84 (2009).
12. D. Oron, E. Tal, and Y. Silberberg, "Scanningless depth-resolved microscopy," Opt. Express **13**, 1468–1476 (2005).
13. B. Berge and J. Peseux, "Variable focal lens controlled by an external voltage: An application of electrowetting," Eur. Phys. J. E. **3**, 159–163 (2000).
14. M. Duocastella and C. B. Arnold, "Enhanced depth of field laser processing using an ultra-high-speed axial scanner," Appl. Phys. Lett. **102**(6), 061113 (2013).
15. A. Mermillod-Blondin, E. McLeod, and C. B. Arnold, "High-speed varifocal imaging with a tunable acoustic gradient index of refraction lens," Opt. Lett. **33**, 2146–2148 (2008).
16. M. Duocastella, G. Vicidomini, and A. Diaspro, "Simultaneous multiplane confocal microscopy using acoustic tunable lenses," Opt. Express **22**, 19293–19301 (2014).
17. G. R. B. E. Römer, P. Bechtold, P. J. Harshman, "Electro-optic and Acousto-optic Laser Beam Scanners," Phy. Procedia **56**, 29–39 (2014).
18. L. Garcia, O. Pinel, P. Jian, N. Barré, L. Jaffrès, J.-F. Morizur, and G. Labroille, "Fast adaptive laser shaping based on multiple laser incoherent combining," Proc. SPIE **10097**, 1009705 (2017).

#302677 https://doi.org/10.1364/OE.25.023060
Journal © 2017 Received 18 Jul 2017; revised 31 Aug 2017; accepted 31 Aug 2017; published 12 Sep 2017

Research Article | Vol. 25, No. 19 | 18 Sep 2017 | OPTICS EXPRESS 23061
Optics EXPRESS

19. J.-F. Morizur, L. Nicholls, P. Jian, S. Armstrong, N. Treps, B. Hage, M. Hsu, W. Bowen, J. Janousek, and H.-A. Bachor, "Programmable unitary spatial mode manipulation," J. Opt. Soc. Am. A **27**, 2524-2531 (2010).
20. G. Labroille, B. Denolle, P. Jian, P. Genevaux, N. Treps, and J.-F. Morizur, "Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion," Opt. Express **22**, 15599-15607 (2014).
21. D. Z. Anderson, "Alignment of resonant optical cavities," Appl. Opt. **23**, 2944–2949 (1984).
22. M. T. L. Hsu, V. Delaubert, P. K. Lam, and W. P. Bowen, "Optimal optical measurement of small displacements," J. Opt. B. Quantum Semiclassical Opt. **12**(6), 495 (2004).
23. V. Delaubert, N. Treps, C. C. Harb, P. K. Lam, and H.-A. Bachor, "Quantum displacement of spatial conjugate variables: displacement and tilt of a Gaussian beam," Opt. Lett. **31**, 1537-1539 (2006).
24. M. Lassen, V. Delaubert, J. Janousek, K. Wagner, H.-A. Bachor, P. K. Lam, N. Treps, P. Buchhave, C. Fabre, and C. C. Harb, "Tools for Multimode Quantum Information: Modulation, Detection, and Spatial Quantum Correlations," Phys. Rev. Lett. **98**, 083602 (2007).
25. A. E. Siegman, *Lasers*, (University Science Books, 1986), Chap. 16.
26. V. Delaubert, "Quantum imaging with a small number of transverse modes," Australian National University and Université Pierre et Marie Curie, Ph.D thesis, January 2007.

1. Introduction

Fast, precise and continuous focal scanning has applications in many different fields, among which are autofocusing systems, imaging, adaptive optics, laser processing, optical inspection, and several types of microscopy (multiphoton, confocal, lightsheet microscopy and endoscopy) [1, 2].

In all fields cited above, an optical system which can precisely scan the focal point of a beam, at high frequency, in a continuous way, and with a high transmission efficiency is desirable. Several techniques already exist to scan the focal point of a beam. The most common approach is to scan a mechanical optical element, for example, a lens after a collimated beam, an axicon [3]. However, inertia imposes a limit to the modulation frequency which can be achieved with a mechanical system. Microelectromechanical systems (MEMs) can be displayed to form a mirror membrane, and electrostatically actuated to adjust a beam's focus [4] but their scanning speed remains also limited by inertia. Electro-optic [5] and acousto-optic [6] systems offer faster alternatives, which can reach several hundred of kHz. However, they may not be used in all the field cited above. Indeed, electro-optic lenses are polarization dependent. In the case of acousto-optic systems, two synchronized acousto-optic crystals are needed to mimic a cylindrical lens which leads to the need of four synchronized crystals to perform a true scan of the focal point and a weak transmission is to be expected at the output of multiple consecutive crystals. Liquid crystal lenses can also be used [7, 8], requiring the use of polarized light. Moreover, in order to achieve long focal lengths, the thickness of the liquid crystal must be increased, which reduces the response time of the system, introducing a trade-off between long scanning lengths and high scanning frequencies. A more sophisticated technique which implements simultaneous spatial and temporal focusing [9–12] has been developed. It relies on operating in the pulsed regime in order to have a sufficient bandwidth to manipulate. A lens of variable focal length - containing two non-miscible liquids, one insulating and the other conducting, such as oil and water - can be fabricated using electrowetting: electric charges at an interface allow the modification of the contact angle of a drop on a metal surface [13]. Finally, acoustically driven liquid lenses - TAG lenses - can be used [14–16]. These lenses continuously vary their focal length, producing periodically varying axial displacement, and show promising results.

In the present work, we do not try to mimic the working principle of a lens but instead, introduce a novel modal approach to this problem. We propose to directly couple the transverse degrees of freedom of a laser beam to its longitudinal ones. Fast and efficient technologies allowing to scan a beam in the transverse plane are already available [17]. The combination of these technologies with the system presented here would allow overcoming many of the limitations listed above. We present a system capable of scanning a beam with a hundred micrometer size waist for up to 4 Rayleigh length. The proposed system being passive, the

scanning rates achievable will be set by those allowed by the transverse scanning technology chosen. The physical implementation of the MPLC device (reflective design compatible with high-reflectivity coatings) makes it compatible with high powers (several kW [18]) for material processing, while being also suitable for microscopy.

2. Theory

The purpose of this work is to improve the control over one degree of freedom of a laser beam by coupling this degree of freedom to another with the help of the Multi-Plane Light Conversion (MPLC) technology [19].

The position and tilt of a beam can be very well controlled, with high precision and at high frequency. These will be the so-called "input" degree of freedom for the MPLC. We are interested in controlling the position of the focal point of a Gaussian beam along its propagation axis. This movement will define the "output" degree of freedom of the MPLC. We aim at exploiting the good control one can have on the position of a beam and use this precise control to defocus the beam coming out of the MPLC with comparable precision. Figure 1 illustrates schematically this principle.

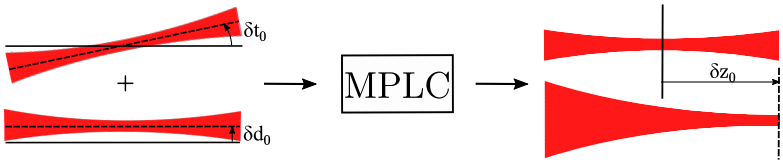

Fig. 1. The MPLC system couples the displacement (δd_0) and tilt (δt_0) of a beam to the scanning of its focal point (δz_0).

MPLC [19, 20] is a novel technology capable of performing, without limitations on resources, any spatial unitary transform. A unitary transform is a transformation which implements a mode basis change without losses. An MPLC device performs unitary transforms on spatial mode basis of the electrical field: it shapes an orthonormal spatial mode basis into another one. The physical transform consists of a succession of phase profiles and is implemented using several consecutive phase plates. To produce an MPLC device, one first has to set the orthonormal input and output mode bases necessary to the definition of the unitary transform.

2.1. Modes of displacement, tilt and defocus

References [21] and [22–24] introduce the idea that, by using a Taylor expansion of a displaced Gaussian beam, one can find the modes which allow to describe this displacement. We write the complex amplitude of the electric field for a fundamental Gaussian beam as $E_0 = \mathcal{E}_0 u_0$ where u_0 is the normalized spatial mode describing the field and $\mathcal{E}_0$ is its amplitude. We will use the following convention regarding notations: p will denote a physical quantity, p_0 will refer to the value of this physical quantity at which we evaluate the reference beam and δp_0 will refer to the amplitude of the perturbation of said physical quantity. We consider a beam E_0 for which p has the value p_0. We write at the first order, the Taylor expansion of E_0 for a small variation δp_0 of p:

$$E_0|_{p_0+\delta p_0} \simeq \mathcal{E}_0 \left(u_0|_{p_0} + \delta p_0 \times \left.\frac{\partial u_0}{\partial p}\right|_{p_0} \right) = E_0|_{p_0} + \frac{\delta p_0}{p_C} \times E_1|_{p_0} \tag{1}$$

where $p_C = \left\| \frac{\partial u_0}{\partial p} \right\|^{-1}$ is a characteristic parameter value for the displacement ($||u||^2 = \int uu^*$), $u_1 = p_C \times \frac{\partial u_0}{\partial p}$ and $E_1 = \mathcal{E}_0 u_1$. Displacing $E_0|_{p_0}$ by a quantity δp_0 transfers energy into mode

Research Article | Vol. 25, No. 19 | 18 Sep 2017 | OPTICS EXPRESS 23063
Optics EXPRESS

$u_1|_{p_0}$, proportionally to δp_0. In a reverse manner, if one can control the relative quantities of energy one puts into the complex amplitudes $u_0|_{p_0}$ and $u_1|_{p_0}$, one can tune the parameter δp_0 for the total beam. This is of course only true at first order, ie. for $\frac{\delta p_0}{p_C} \ll 1$.

In the case of a defocus of a Gaussian beam, focused at z_0 (z being the longitudinal direction), which has been studied in one dimension in [26], we denote the complex amplitude of the field as (see [25], p.627): $E_0(\rho, z)|_{z_0} = \mathcal{E}_0\, u_0(\rho, z)|_{z_0} \exp(-ik(z - z_0))$. The field u_0 is the transverse optical mode of the field $\mathcal{E}_0$, k the wave vector and z_0 is the focus position of the beam. In this case, taking into account the rotational symmetry of the problem, we use the expression (see [25], p.647) of a Laguerre-Gauss mode LG_{pm} and take $u_0(\rho, z)|_{z_0} = \mathrm{LG}_{00}(\rho, z)|_{z_0}$. We use the first order Taylor expansion to write:

$$E_0(\rho, z)|_{z_0+\delta z_0} = \mathcal{E}_0 \left[\mathrm{LG}_{00}(\rho, z)|_{z_0} + \frac{i\delta z_0}{z_R} \left(\alpha \times \mathrm{LG}_{00}(\rho, z)|_{z_0} + \mathrm{LG}_{01}(\rho, z)|_{z_0} \right) \right] \exp(-ik(z - z_0)) \tag{2}$$

where $\alpha = 1 - (kw_0)^2$ and w_0 is the waist. We stress the fact that in this particular case, $E_1(\rho, z)|_{z_0} = \mathcal{E}_0\, u_1(\rho, z)|_{z_0} \exp(-ik(z - z_0))$ and $u_1|_{z_0} = A\left(\alpha \times \mathrm{LG}_{00}|_{z_0} + \mathrm{LG}_{01}|_{z_0}\right)$ (where A is a normalization factor), which means that $u_0|_{z_0}$ and $u_1|_{z_0}$ are not orthogonal.

We are also interested in the displacement and tilt of a beam in the y (vertical) direction of the transverse plane, which are the degrees of freedom we aim at coupling to the defocus. These displacements were studied in [21, 22] and it was shown that the mode basis which is relevant to describe a fundamental Gaussian beam displaced or tilted along one direction of the transverse plane is the Hermite-Gaussian mode basis. This basis is referred to as the $\mathrm{HG}_{n_x n_y}$ basis. We work in the (y, z) plane and set $z_0 = 0$. At the first order, for a beam displaced along d by an amount δd_0 with respect to a reference position d_0, we can write:

$$E_0(\rho, z)|_{d_0+\delta d_0} = \mathcal{E}_0 \left[\mathrm{HG}_{00}(\rho, z)|_{d_0} + \frac{\delta d_0}{w_0} \times \mathrm{HG}_{01}(\rho, z)|_{d_0} \right] \exp(-ikz). \tag{3}$$

For a beam tilted in the $(y - z)$ plane by a quantity $\delta t_0 = 2\pi\theta_y/\lambda$, with respect to a reference propagation axis defined by θ_0 and $t_0 = 2\pi\theta_{y0}/\lambda$, we can write at the first order:

$$E_0(\rho, z)|_{t_0+\delta t_0} = \mathcal{E}_0 \left[\mathrm{HG}_{00}(\rho, z)|_{t_0} + i\frac{\delta t_0 w_0}{2} \times \mathrm{HG}_{01}(\rho, z)|_{t_0} \right] \exp(-ikz). \tag{4}$$

We observe that in both cases, $u_1|_{p_0} = \mathrm{HG}_{01}|_{p_0}$ and $u_1|_{p_0}$ is orthogonal to $u_0|_{p_0}$.
The first order Taylor expansion is only valid for small variations of a parameter. However, we aim at producing displacements of the focal point larger than z_R (characteristic length for this displacement), which means we need to take into account several higher orders of the Taylor expansion to describe such a displacement. We recall that in order to define an MPLC device, one needs orthonormal input and output mode bases. We introduced mode bases which are described by linear combinations of the Hermite-Gauss and Laguerre-Gauss mode bases. But the first order derivative in the case of a longitudinal waist displacement is not orthogonal to the Gaussian fundamental mode. This non-orthogonality appears for the displacement and tilt starting at the second order derivative and means that these mode bases cannot be used as such to define an MPLC device. Suitable mode bases can be defined using these bases and an orthonormalization algorithm.

However, we decided instead to develop a numerical method - inspired by the above theoretical approach - to generate orthonormal input and output mode bases fitted to this "large displacement" problem.

2.2. Numerical approach for the definition of the input-output mode bases

The complexity (number of phase plates) of an MPLC device scales with the number of modes. In order for our device to have a standard complexity (10 phase plates) while producing a significant defocus (4 z_R), we found it necessary through an iterative process linking the mode basis and the defocus range to use a 5 modes mode basis (with which we can produce an overlap > 88% with a perfect defocused Gaussian beam).

The process defining the output mode basis takes as an input the range of defocus we aim at. We first define the modes v_i, as depicted in Fig. 2: we take beams which are regularly defocused along the 4 z_R defocus range and define them on a common transverse plane $\mathcal{P}$. This set of modes v_i does not constitute an orthogonal mode basis. We orthonormalize this set of modes, using the Gram-Schmidt algorithm, defining the output mode basis we will use.

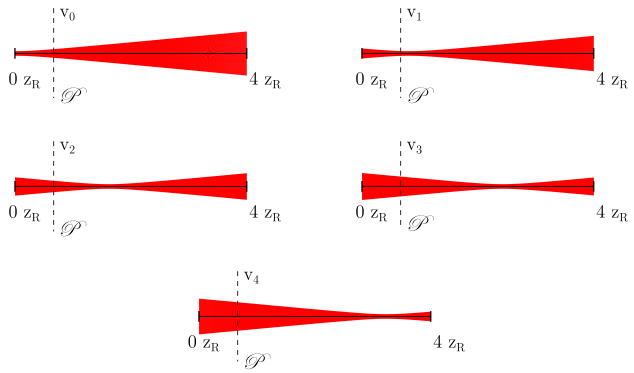

Fig. 2. Target modes before orthonormalization (output mode basis).

The input mode basis is defined in a similar fashion, replacing the role of defocus by that of lateral displacement. We give their representation on Fig. 3. As inferred by the analytical approach, they bear a close resemblance to the Hermite-Gauss and Laguerre-Gauss mode bases.

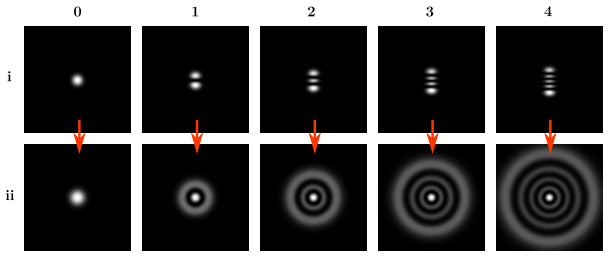

Fig. 3. Intensity profiles of i) Input mode basis, ii) Output mode basis. The MPLC matches mode (i, n) with mode (ii, n) with $n \in [\![0, 4]\!]$.

With the given output mode basis, we computed the projection of the desired output modes

Research Article | Vol. 25, No. 19 | 18 Sep 2017 | OPTICS EXPRESS 23065
Optics EXPRESS

(defocused Gaussian beam between 0 and $4\,z_R$) on this basis. This led to the definition of a projection function f_i for each mode of the basis (the functions f_i take their values in $\mathbb{C}$). Indeed, since we can only address the 5 modes of the output basis, we can only perform an approximation (E_{out}) of the desired output mode (E_{out}^{obj}). For $\delta z_0^{(k)} \in [0\,z_R, 4\,z_R]$, we project the desired mode onto the output basis $(v_i)_{i\in[\![0,4]\!]}$:

$$E_{out}^{obj}\Big|_{\delta z_0^{(k)}} \simeq E_{out}|_{\delta z_0^{(k)}} = \sum_{i=0}^{4} f_i\left(\delta z_0^{(k)}\right) \times v_i \tag{5}$$

If we suppose a perfect MPLC device and using equation 5, we can also infer the following decomposition:

$$E_{in}\left(\delta z_0^{(k)}\right) = \sum_{i=0}^{4} f_i\left(\delta z_0^{(k)}\right) \times u_i \tag{6}$$

Knowing the f_i functions, we can compute the input modes: indeed, the MPLC device shapes the different modes of the input basis while leaving their respective weight coefficients unchanged. We have subsequently calculated for each focusing position $\delta z_0^{(k)}$, the displacement and tilt couple $(\delta d_0^{(k)}, \delta t_0^{(k)})$ which maximized the following function F - we set $d_0 = 0$ and $t_0 = 0$. F represents the overlap between the field we aim at producing and the field we actually produce at the input of the MPLC device:

$$F(\delta d_0^{(k)}, \delta t_0^{(k)}) = \left|\langle E_{in}\left(\delta d_0^{(k)}, \delta t_0^{(k)}\right) | \sum_{i=0}^{4} f_i\left(\delta z_0^{(k)}\right) \times u_i\rangle\right|^2 \tag{7}$$

where the input basis is the basis $(u_i)_{i\in[\![0,4]\!]}$. We used a least-square numerical method to perform this optimization. In Fig. 4 (left) we have plotted, in normalized displacement and tilt units, the trajectory which needs to be described to induce a continuous focal shift from $0\,z_R$ to $4\,z_R$ at the output of the MPLC. We note that this trajectory is continuous, meaning that we can describe any defocusing between $0\,z_R$ and $4\,z_R$ following this curve. The amplitudes of displacement and tilt which needs to be imposed on the beam are functions of the size of the waist w_0. For instance, in order to shift the focus point of the beam by $\delta z_0 = 4\,z_R$, one needs to impose a displacement of $\delta d_0/w_0 = 1.09$ and a tilt of $\delta t_0 w_0/2 = -0.15$ to the beam that is launched in MPLC device.

In Fig. 4 (right), we show the overlap between a perfectly defocused Gaussian beam and the beam produced by a perfect unitary operation, when following the trajectory displayed on the same figure. As a mean of comparison, we plotted the overlap between the perfectly defocused Gaussian beam and a constant Gaussian beam.

2.3. Implementation of the MPLC

The MPLC technology can, in theory, implement perfectly any unitary transformation. However, since the experimental resources we use are limited (size and number of phase plates, propagation length between two phase plates), an actual MPLC only performs an approximation of the desired unitary transform. Table 1 shows the overlap between the theoretical output modes (ideal output mode basis displayed on Fig. 3) and the actual mode produced by the MPLC we calculated. These overlaps are theoretical ones, and in an experimental configuration, misalignments of the physical system with respect to the simulated one will introduce an even greater difference from the initial "ideal" output mode basis.

On figure 4 (right), we plotted the overlap between the perfectly defocused Gaussian beam and the optimized output mode calculated with the effective output mode basis of the MPLC. We observe a drop in efficiency of maximum 2% compared to the performances of an ideal unitary operation.

Table 1. Overlap between theoretical modes and simulated modes at the output of the MPLC

Mode	0	1	2	3	4
Overlap	0.9893	0.9902	0.9920	0.9933	0.9928

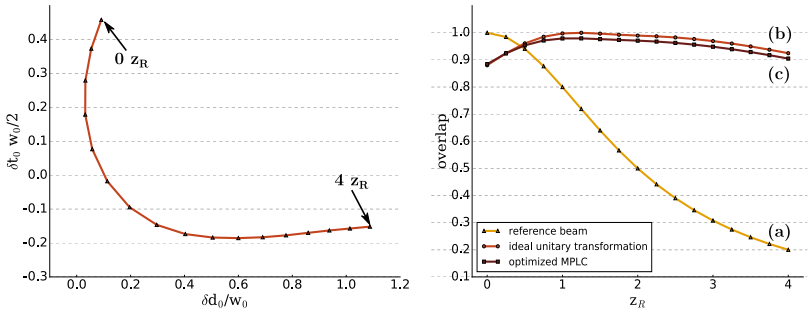

Fig. 4. Left: Calculated trajectory in normalized displacement. δd_0 represents the lateral displacement imposed on the beam and w_0 is the waist of the beam, δt_0 represents the tilt imposed on the beam. This trajectory needs to be impressed on the input beam for the focus point of the beam at the output of the MPLC to see its focus point vary between $0\,z_R$ and $4\,z_R$, z_R being the Rayleigh length of the beam. Right: Overlap between the desired output (perfectly defocused Gaussian beam) and the reference beam ie. a Gaussian beam focused at $0\,z_R$ (yellow color or (a)), the beam which would be produced by an ideal unitary operation (red color or (b)) and the beam which can be produced by the actual MPLC we calculated (brown color or (c)).

3. Experiment and results

We tested the validity of our numerical simulations with the following proof of principle experiment (see Fig. 5).

To precisely control the displacement and tilt, we used a liquid-crystal spatial light modulator (SLM) to control the light beam. Acousto-optic modulators or galvanometers have better performances at high frequency and can achieve comparable steering precision. In this experiment, the displacement and tilt being produced by the SLM, the maximum operating frequency is set by the maximum operating frequency of the SLM - in this case 203 Hz. The MPLC being a passive phase element, the frequency of the achieved defocus depends exclusively on the operating frequency of the technological solution chosen to steer the beam, as long as the associated timescale is small compared to the timescale associated with light propagation in the system.

We use a continuous wave laser at $\lambda = 1.064$ nm coupled to a single-mode fiber. At the output of the fiber, the beam is collimated, and we select the polarization compatible with the SLM liquid-crystals with a polarizing beam splitter. We use a lens to mode-match the beam to produce a Gaussian beam with a waist $w_0 = 126\,\mu$m, which is the waist size we used to define the input mode basis of the MPLC device. The beam then reflects twice on the SLM. The SLM is used as two independent mirrors. We divide its active surface in two and each side plays the role of an independent mirror which can be tilted: on the two halves, we display linear ramps of angle α_1 and α_2. Knowing the lengths L_1, L_2 and L_3, we can compute the displacement and tilt we impose on the beam, and therefore precisely control the parameters δd_0 and δt_0.

Due to the footprint of the optical elements, a 4-f telescope is used to launch the light into the MPLC fabricated by CAILabs. The MPLC is composed of a phase plate (printed using

Research Article | Vol. 25, No. 19 | 18 Sep 2017 | OPTICS EXPRESS 23067
Optics EXPRESS

lithography) and a plane mirror between which the light bounces 10 times before exiting the system. It is represented on Fig. 6 (left). At the output of this system, we used a lens in a 2f-2f configuration to image the output mode on a scanning slit beam profiler. The MPLC was designed for the output beam to have a waist of $w_0 = 113\ \mu$m (with and associated $z_R = 37.7$ mm).

The spatial modes are defined in the plane normal to the propagation direction or z-axis. The vertical and horizontal direction with respect to the optical table are the y and x-directions. The footprint of the optical elements led to a small clip of one side of the beam - it is a particular problem of the experimental setup and not a fundamental limit. For this reason, the quality of the beam was degraded on the horizontal or x-direction. The results we obtained on the x-dimension were consequently degraded compared to those of the y-dimension, and we present here only the results on the y-dimension. In the present experiment, mirror holders with a smaller footprint may have allowed to avoid this problem.

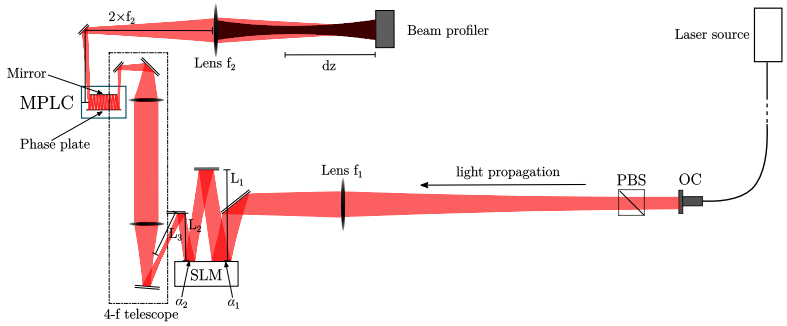

Fig. 5. Experimental setup built to implement the defocus of a beam with an MPLC device. The output beam of a single mode fiber, collimated by an output coupler (OC), is mode-matched to the correct waist size with lens f_1. We select the correct polarization with a polarizing beam splitter (PBS) in order for the beam to be compatible with the spatial light modulator (SLM). Two reflections on the SLM allow to impress on the beam two successive angles α_1, α_2. The focus point of the beam is imaged with a 4-f telescope and launched into the MPLC. At the output of the MPLC, lens f_2 performs a 2-f imaging of the defocused beam. The beam transverse profile is monitored by a beam profiler.

The experimental procedure we carried out was the following. For a given displacement δz_0, we positioned lens f_2 at a distance $\delta z_0 + 2{\times}f_2$ (f_2 being the focal length of lens f_2) and the beam profiler at a distance $2{\times}f_2$ of the lens. We adjusted the parameters (δd_0, δt_0) through the angles (α_1, α_2) to minimize the waist size measured by the beam profiler. The correct parameters (α_1, α_2) were then fixed and the position of the beam profiler was scanned along the propagation axis to record the evolution of the beam waist size. The results of those measurements are presented on Fig. 6 (right). A Gaussian fit to the recorded data is performed in order to calculate the waist size (dotted lines on Fig. 6 (right)). For every position, the experimental data is plotted as well as the corresponding numerical fit (superimposed solid lines) with the waist formula: $w(z) = w_0[1 + \left(M^2(z - z_0)/z_R\right)^2]^{1/2}$.

The theoretical waist size which we aimed at producing is plotted as well. We observe a clear defocus on a $4\ z_R$ amplitude range, confirming the validity of our approach. The measured waist size varies between 113 μm and 154 μm, while the M^2 factor of the fit to the measurements varies between 1.17 and 1.36. The experimental values, while in good agreement with the theory, do not exactly match with our model. We believe this to be mainly due to two effects. First, as mentioned before, the beam was clipped in the x-direction: this introduces a mismatch with

respect to a perfect Gaussian beam. Furthermore, we suspect a waist size mismatch between the input beam experimentally produced and the beam used to produce the MPLC device. Those two elements both reduce the overlap between the input beam and the input mode basis used to define the MPLC device.

In this experimental setup, the transmission efficiency is limited by two elements: the efficiency of the steering system and that of the MPLC system. The transmission losses for the MPLC take their origin in the fact that we use finite-size phase plates, discrepancies between the theoretically calculated phase plates and the printed phase plates, as well as from the losses due to the quality of the coating used on the phase plate and mirror. On a typical telecommunication system manufactured by CAILabs, with standard quality evaporated gold reflective coatings and 14 reflections, the overall transmission efficiency of the system is of the order of 70 %. To improve this number, the envisioned technological solutions are first to improve the quality of the coatings used, and the accuracy of the phase plate fabrication. Also, the transmitted power does not theoretically depend on the chosen point on the defocusing trajectory.

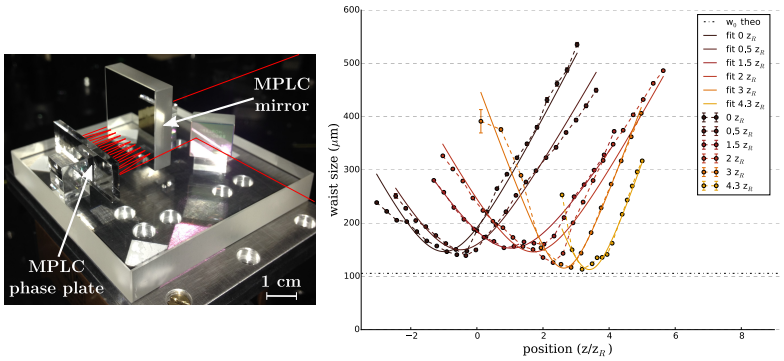

Fig. 6. Left: Photograph of the MPLC device used in the experimental setup. The MPLC consists of a plane mirror facing a planar phase plate on which 10 successive phase profiles are written to implement the desired transformation. Right: Experimental demonstration of defocusing over $4\,z_R$ range using the displacement and tilt of a beam combined with an MPLC. The waist size evolution in the y direction along the propagation direction of the output beam is shown for 6 different focusing positions. The dotted black line represents the theoretical waist size.

4. Conclusion

We demonstrate that we can couple different spatial parameters of a laser beam. Since some parameters of the transverse profile of a laser beam can be tuned with ease and precision at high frequency, while some others cannot be as easily modified, this coupling can be used to improve the control quality one can have on certain parameters. In this work, we first recalled the expression of modes which describes the small variations of one parameter, using a perturbative treatment. This theoretical analysis seeded the numerical method we developed to address large displacements. We defined two orthogonal mode bases, one associated with a transverse beam displacement, the other with a defocus. Using an MPLC device coupling those two mode bases, we were able to numerically and experimentally couple the angle and displacement of an input beam to the defocus of an output beam over a $4\,z_R$ range. This range was chosen in order to ensure a theoretical overlap superior to 88% between the produced mode and the perfect

Research Article Vol. 25, No. 19 | 18 Sep 2017 | OPTICS EXPRESS 23069
Optics EXPRESS

defocused Gaussian beam over the whole scanning range. Experimentally, we observed that outside this range, the quality of the beam deteriorates. We demonstrated that with this system, we can control the defocus of a beam with the control quality one can have on the displacement and tilt of a beam. Using the most precise techniques available for the displacement and tilt of a beam, the scan of the focus point of a Gaussian beam in a nearly continuous fashion in a given range is made available: in order to control the position of the beam with a $z_R/10$ precision (in this case 3.7mm) along a 4 z_R defocusing trajectory, a precision of $\delta d_0/w_0 = 0.14\%$ and $\delta t_0 w_0/2 = 0.032\%$ (in this experiment $\delta d_0 = 0.15\,\mu$m and $\delta\theta_y = 1\,\mu$rad) is necessary. This scanning can be done at high-frequency since no mechanical element is involved in this system. The theoretical and numerical approaches we put forward in this work are not limited to displacement/tilt and defocus and can be adapted to other spatial degrees of freedom.

5. Acknowledgments

The authors thank Claude Fabre and Valentina Parigi for helpful discussions. N.Treps is a member of the Institut Universitaire de France.

Appendix C

Transforms in free probability

We give here the definitions of two types of transforms introduced in 4.

The R-transform is defined as

$$\mathrm{R}(z) = g^{-1}(z) - \frac{1}{z}. \tag{C.1}$$

This transform allows to express some spectrum laws in a concise way [Tulino 04]. The R-transform of the semicircle law is

$$\mathrm{R}(z) = z \tag{C.2}$$

and the R-transform of the Marčenko-Pastur law is

$$\mathrm{R}(z) = \frac{1}{1-\beta z}. \tag{C.3}$$

Another type of transform is called the S-transform. If we define $\chi(z)$ as a solution of the following equation

$$\frac{1}{\chi(z)} g\left(\frac{1}{\chi(z)}\right) - 1 = z, \tag{C.4}$$

the S-transform is given by [Goetschy 11]

$$\mathrm{S}(z) = \frac{1+z}{z}\chi(z). \tag{C.5}$$

References

[Abbe 73] E. Abbe. *Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung.* Archiv für mikroskopische Anatomie, vol. 9, no. 1, pages 413–418, Dec 1873. Online URL. (Cited on page 44.)

[Abbott 16] B. P. Abbott & *et al. Observation of Gravitational Waves from a Binary Black Hole Merger.* Phys. Rev. Lett., vol. 116, page 061102, Feb 2016. Online URL. (Cited on page 2.)

[Akemann 11] Gernot Akemann, Jinho Baik & Philippe Di Francesco. The oxford handbook of random matrix theory. Oxford Handbooks, 2011. (Cited on page 120.)

[Akkermans 07] Eric Akkermans & Gilles Montambaux. Mesoscopic physics of electrons and photons. Cambridge University Press, 2007. (Cited on pages 120 and 128.)

[Alekseev 98] A. N. Alekseev, K. N. Alekseev, O. S. Borodavka, A. V. Volyar & Yu. A. Fridman. *Conversion of Hermite-Gaussian and Laguerre-Gaussian beams in an astigmatic optical system. 1. Experiment.* Technical Physics Letters, vol. 24, no. 9, pages 694–696, Sep 1998. Online URL. (Cited on page 23.)

[Anderson 84] Dana Z. Anderson. *Alignment of resonant optical cavities.* Appl. Opt., vol. 23, no. 17, pages 2944–2949, Sep 1984. Online URL. (Cited on page 100.)

[Ang 17] Shan Zheng Ang, Ranjith Nair & Mankei Tsang. *Quantum limit for two-dimensional resolution of two incoherent optical point sources.* Phys. Rev. A, vol. 95, page 063847, Jun 2017. Online URL. (Cited on pages 66 and 67.)

[Annoni 17] Andrea Annoni, Emanuele Guglielmi, Marco Carminati, Giorgio Ferrari, Marco Sampietro, David AB Miller & Francesco Morichetti. *Unscrambling light—automatically undoing strong mixing between modes.* Light: Science & Applications, vol. 6, 2017. Online URL. (Cited on page 34.)

[Ashkin 86] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm & Steven Chu. *Observation of a single-beam gradient force optical trap for dielectric particles.* Opt. Lett., vol. 11, no. 5, pages 288–290, May 1986. Online URL. (Cited on page 3.)

[Aubry 09] Alexandre Aubry & Arnaud Derode. *Random Matrix Theory Applied to Acoustic Backscattering and Imaging In Complex Media.* Phys. Rev. Lett., vol. 102, page 084301, Feb 2009. Online URL. (Cited on page 120.)

[Aubry 10] Alexandre Aubry & Arnaud Derode. *Singular value distribution of the propagation matrix in random scattering media.* Waves in Random and Complex Media, vol. 20, no. 3, pages 333–363, 2010. Online URL. (Cited on page 135.)

[Bade 18] S. Bade, B. Denolle, G. Trunet, N. Riguet, P. Jian, O. Pinel & G. Labroille. *Fabrication and Characterization of a Mode-selective 45-Mode Spatial Multiplexer based on Multi-Plane Light Conversion.* In 2018 Optical Fiber Communications Conference and Exposition (OFC), pages 1–3, March 2018. (Cited on page 96.)

[Barnett 03] S.M. Barnett, C. Fabre & A. Maıtre. *Ultimate quantum limits for resolution of beam displacements.* The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics, vol. 22, no. 3, pages 513–519, Mar 2003. Online URL. (Cited on page 61.)

[Barnett 09] Stephen M. Barnett. Quantum information. Oxford Master Series in Physics, 2009. (Cited on page 57.)

[Barredo 18] Daniel Barredo, Vincent Lienhard, Sylvain de Léséleuc, Thierry Lahaye & Antoine Browaeys. *Synthetic three-dimensional atomic structures assembled atom by atom.* Na-

ture, vol. 561, pages 79–82, Sep 2018. Online URL. (Cited on page 3.)

[Beckers 93] Jacques M. Beckers. *Adaptive Optics for Astronomy: Principles, Performance, and Applications.* Annual Review of Astronomy and Astrophysics, vol. 31, no. 1, pages 13–62, 1993. Online URL. (Cited on pages 2 and 25.)

[Beenakker 97] C. W. J. Beenakker. *Random-matrix theory of quantum transport.* Rev. Mod. Phys., vol. 69, pages 731–808, Jul 1997. Online URL. (Cited on pages 120, 128, and 133.)

[Beenakker 11] C. W. J. Beenakker. *Applications of random matrix theory to condensed matter and optical physics.* In Gernot Akemann, Baik & Philippe Di Francesco, editeurs, The Oxford Handbook of Innovation of Random Matrix Theory, chapitre 35–36. Oxford University Press, Oxford, (Cited on pages 131 and 132.)

[Beijersbergen 93] M.W. Beijersbergen, L. Allen, H.E.L.O. van der Veen & J.P. Woerdman. *Astigmatic laser mode converters and transfer of orbital angular momentum.* Optics Communications, vol. 96, no. 1, pages 123 – 132, 1993. Online URL. (Cited on page 23.)

[Berdagué 82] S. Berdagué & P. Facq. *Mode division multiplexing in optical fibers.* Appl. Opt., vol. 21, no. 11, pages 1950–1955, Jun 1982. Online URL. (Cited on page 26.)

[Berge 00] B. Berge & J. Peseux. *Variable focal lens controlled by an external voltage: An application of electrowetting.* The European Physical Journal E, vol. 3, no. 2, pages 159–163, Oct 2000. Online URL. (Cited on page 107.)

[Berry 79] M. V. Berry & N. L. Balazs. *Nonspreading wave packets.* American Journal of Physics, vol. 47, no. 3, pages 264–267, 1979. Online URL. (Cited on page 18.)

[Bertero 82a] M. Bertero, P. Boccacci & E.R. Pike. *Resolution in Diffraction-limited Imaging, a Singular Value Analysis — II. The case of incoherent illumination.* Optica Acta: International Journal of Optics, vol. 29, no. 12, pages 1599–1611, 1982. Online URL. (Cited on pages 46 and 47.)

[Bertero 82b] M. Bertero & E.R. Pike. *Resolution in Diffraction-limited Imaging, a Singular Value Analysis — I. The case of coherent illumination.* Optica Acta: International Journal of Optics, vol. 29, no. 6, pages 727–746, 1982. Online URL. (Cited on page 46.)

[Betzig 06] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz & Harald F. Hess. *Imaging Intracellular Fluorescent Proteins at Nanometer Resolution.* Science, vol. 313, no. 5793, pages 1642–1645, 2006. Online URL. (Cited on page 51.)

[Birks 15] T. A. Birks, I. Gris-Sánchez, S. Yerolatsitis, S. G. Leon-Saval & R. R. Thomson. *The photonic lantern.* Adv. Opt. Photon., vol. 7, no. 2, pages 107–167, Jun 2015. Online URL. (Cited on page 28.)

[Bobin 08] J. Bobin, J.-L. Starck & R. Ottensamer. *Compressed Sensing in Astronomy.* IEEE Journal of Selected Topics in Signal Processing, vol. 2, pages 718–726, (Cited on page 88.)

[Bohigas 84] O. Bohigas, M. J. Giannoni & C. Schmit. *Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws.* Phys. Rev. Lett., vol. 52, pages 1–4, Jan 1984. Online URL. (Cited on page 119.)

[Borevich 81] Z. I. Borevich & S. L. Krupetskii. *Subgroups of the unitary group that contain the group of diagonal matrices.* Journal of Soviet Mathematics, vol. 17, no. 4, pages 1951–1959, Nov 1981. Online URL. (Cited on pages 30 and 31.)

[Boucher 17] Pauline Boucher, Nicolas Barré, Olivier Pinel, Guillaume Labroille & Nicolas Treps. *Continuous axial scanning of a Gaussian beam via beam steering.* Opt. Express, vol. 25, no. 19, pages 23060–23069, Sep 2017. Online URL. (Cited on page 100.)

[Brouwer 96] P. W. Brouwer & C. W. J. Beenakker. *Diagrammatic method of integration over the unitary group, with applications to quantum transport in mesoscopic systems.* Journal of Mathematical Physics, vol. 37, no. 10, pages 4904–4934, 1996. Online URL. (Cited on page 146.)

[Brouwer 97] P.W. Brouwer. On the random-matrix theory of quantum transport. 1997. (Cited on pages 128 and 130.)

[Candès 05] Emmauel Candès & Terence Tao. *Decoding by linear programming*. IEEE Transactions on Information Theory, vol. 51, no. 12, pages 4203–4215, (Cited on page 95.)

[Candès 06a] Emmanuel Candès. *Compressive sampling*. Proc Int Congre Math, (Cited on page 89.)

[Candès 06b] Emmanuel Candès, Justin Romberg & Terence Tao. *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*. IEEE Transactions on Information Theory, vol. 52, no. 2, pages 489–509, (Cited on page 89.)

[Candès 06c] Emmanuel Candès, Justin Romberg & Terence Tao. *Stable signal recovery from incomplete and inaccurate measurements*. Communications on Pure and Applied Mathematics, vol. 59, no. 8, pages 1207–1223, 2006. Online URL. (Cited on page 89.)

[Candès 06d] Emmanuel Candès & Terence Tao. *Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?* IEEE Transactions on Information Theory, vol. 52, no. 12, pages 5406–5425, (Cited on page 89.)

[Candès 07a] Emmanuel Candès & Justin Romberg. *Sparsity and incoherence in compressive sampling*. Inverse Problems, vol. 23, no. 3, page 969, 2007. Online URL. (Cited on page 92.)

[Candès 07b] Emmanuel Candès & Michael B. Wakin. *"People Hearing Without Listening": An Introduction To Compressive Sampling*. 2007. (Cited on pages 89, 90, 92, and 94.)

[Caves 81] Carlton M. Caves. *Quantum-mechanical noise in an interferometer*. Phys. Rev. D, vol. 23, pages 1693–1708, Apr 1981. Online URL. (Cited on pages 53, 57, and 61.)

[Chang 94] A. M. Chang, H. U. Baranger, L. N. Pfeiffer & K. W. West. *Weak Localization in Chaotic versus Nonchaotic Cavities: A*

Striking Difference in the Line Shape. Phys. Rev. Lett., vol. 73, pages 2111–2114, Oct 1994. Online URL. (Cited on page 131.)

[Chang 14] Sun Hyok Chang, Hwan Seok Chung, Nicolas K. Fontaine, Roland Ryf, Kyung Jun Park, Kwangjoon Kim, Jyung Chan Lee, Jong Hyun Lee, Byoung Yoon Kim & Young Kie Kim. *Mode division multiplexed optical transmission enabled by all–fiber mode multiplexer.* Opt. Express, vol. 22, no. 12, pages 14229–14236, Jun 2014. Online URL. (Cited on page 28.)

[Cisco 17] Cisco. *Cisco Visual Networking Index: Forecast and Methodology, 2016–2021.* 2017. Online URL. (Cited on page 3.)

[Clements 16] William R. Clements, Peter C. Humphreys, Benjamin J. Metcalf, W. Steven Kolthammer & Ian A. Walmsley. *Optimal design for universal multiport interferometers.* Optica, vol. 3, no. 12, pages 1460–1465, Dec 2016. Online URL. (Cited on page 34.)

[Cohen-Tannoudji 68] Claude Cohen-Tannoudji. "optical pumping and interaction of atoms with the electromagnetic fieldïn cargèse lectures in physics. Gordon and Breach, 1968. (Cited on page 126.)

[Cramér 99] Harald Cramér. Mathematical methods of statistics (pms-9). Princeton University Press, 1999. (Cited on page 57.)

[Davenport 12] Mark A. Davenport, Marco F. Duarte, Yonina C. Eldar & Gitta Kutyniok. Introduction to compressed sensing, page 1–64. Cambridge University Press, 2012. (Cited on page 89.)

[Davidson 95] Nir Davidson, Heun Jin Lee, Charles S. Adams, Mark Kasevich & Steven Chu. *Long Atomic Coherence Times in an Optical Dipole Trap.* Phys. Rev. Lett., vol. 74, pages 1311–1314, Feb 1995. Online URL. (Cited on page 3.)

[De Cock 99] M. De Cock, M. Fannes & P. Spincemaille. *On quantum dynamics and statistics of vectors.* Journal of Physics A: Mathematical and General, vol. 32, no. 37, page 6547, 1999. Online URL. (Cited on page 144.)

[Defienne 16] Hugo Defienne, Marco Barbieri, Ian A. Walmsley, Brian J. Smith & Sylvain Gigan. *Two-photon quantum walk in a multimode fiber*. Science Advances, vol. 2, no. 1, 2016. Online URL. (Cited on page 2.)

[Delaubert 06a] V. Delaubert, N. Treps, M. Lassen, C. C. Harb, C. Fabre, P. K. Lam & H-A. Bachor. TEM_{10} *homodyne detection as an optimal small-displacement and tilt-measurement scheme*. Phys. Rev. A, vol. 74, page 053823, Nov 2006. Online URL. (Cited on page 65.)

[Delaubert 06b] Vincent Delaubert, Nicolas Treps, Charles C. Harb, Ping Koy Lam & Hans-A. Bachor. *Quantum measurements of spatial conjugate variables: displacement and tilt of a Gaussian beam*. Opt. Lett., vol. 31, no. 10, pages 1537–1539, May 2006. Online URL. (Cited on pages 65 and 100.)

[Delaubert 07] Vincent Delaubert. *Quantum imaging with a small number of transverse modes*. PhD thesis, UPMC & ANU, 2007. (Cited on pages 59 and 61.)

[Delaubert 08] V. Delaubert, N. Treps, C. Fabre, H. A. Bachor & P. Réfrégier. *Quantum limits in image processing*. EPL, vol. 81, no. 4, page 44001, 2008. Online URL. (Cited on page 65.)

[di Francia 69] G. Toraldo di Francia. *Degrees of Freedom of an Image*. J. Opt. Soc. Am., vol. 59, no. 7, pages 799–804, Jul 1969. Online URL. (Cited on pages 46, 47, and 48.)

[Dickey 05] Fred M. Dickey, Todd E. Lizotte, Scott C. Holswade & David L. Shealy. Laser beam shaping applications. CRC Press, 2005. (Cited on page 5.)

[Donoho 06] D. L. Donoho. *Compressed sensing*. IEEE Transactions on Information Theory, vol. 52, no. 4, pages 1289–1306, (Cited on page 89.)

[Dorokhov 82] O.N. Dorokhov. *Transmission coefficient and the localization length of an electron in N bound disordered chains*. JETP Lett., vol. 36, pages 318–321, (Cited on page 132.)

[Doron 90] E. Doron, U. Smilansky & A. Frenkel. *Experimental demonstration of chaotic scattering of microwaves.* Phys. Rev. Lett., vol. 65, pages 3072–3075, Dec 1990. Online URL. (Cited on page 120.)

[Driscoll 13] Jeffrey B. Driscoll, Richard R. Grote, Brian Souhan, Jerry I. Dadap, Ming Lu & Richard M. Osgood. *Asymmetric Y junctions in silicon waveguides for on-chip mode-division multiplexing.* Opt. Lett., vol. 38, no. 11, pages 1854–1856, Jun 2013. Online URL. (Cited on page 28.)

[Druart 08] Guillaume Druart, Jean Taboury, Nicolas Guérineau, Riad Haïdar, Hervé Sauer, Alain Kattnig & Jérôme Primot. *Demonstration of image-zooming capability for diffractive axicons.* Opt. Lett., vol. 33, no. 4, pages 366–368, Feb 2008. Online URL. (Cited on page 106.)

[Du 09] Rui Du, Kun Bi, Shaoqun Zeng, Derong Li, Songchao Xue & Qingming Luo. *Analysis of fast axial scanning scheme using temporal focusing with acousto-optic deflectors.* Journal of Modern Optics, vol. 56, no. 1, pages 81–84, 2009. Online URL. (Cited on page 107.)

[Duarte 05] M. F. Duarte, Shriram Sarvotham, Dror Baron, M. Wakin & R. G. Baraniuk. *Distributed Compressed Sensing of Jointly Sparse Signals.* Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005., pages 1537–1541, (Cited on page 95.)

[Duocastella 13] M. Duocastella & C. B. Arnold. *Enhanced depth of field laser processing using an ultra-high-speed axial scanner.* Applied Physics Letters, vol. 102, no. 6, page 061113, 2013. Online URL. (Cited on page 107.)

[Duocastella 14] Martí Duocastella, Giuseppe Vicidomini & Diaspro Alberto. *Simultaneous multiplane confocal microscopy using acoustic tunable lenses.* Opt. Express, vol. 22, no. 16, pages 19293–19301, Aug 2014. Online URL. (Cited on page 107.)

[Durnin 87a] J. Durnin. *Exact solutions for nondiffracting beams. I. The scalar theory.* J. Opt. Soc. Am. A, vol. 4, no. 4, pages 651–654, Apr 1987. Online URL. (Cited on page 17.)

[Durnin 87b] J. Durnin, J. J. Miceli & J. H. Eberly. *Diffraction-free beams.* Phys. Rev. Lett., vol. 58, pages 1499–1501, Apr 1987. Online URL. (Cited on page 17.)

[Durst 06] Michael E. Durst, Guanghao Zhu & Chris Xu. *Simultaneous spatial and temporal focusing for axial scanning.* Opt. Express, vol. 14, no. 25, pages 12243–12254, Dec 2006. Online URL. (Cited on page 107.)

[Dyson 62] Freeman J. Dyson. *Statistical Theory of the Energy Levels of Complex Systems. I.* Journal of Mathematical Physics, vol. 3, no. 1, pages 140–156, 1962. Online URL. (Cited on page 132.)

[Efron 94] Uzi Efron. Spatial light modulator technology: Materials, devices and applications. CRC Press, 1994. (Cited on page 24.)

[Fabre 00] C. Fabre, J. B. Fouet & A. Maître. *Quantum limits in the measurement of very small displacements in optical images.* Opt. Lett., vol. 25, no. 1, pages 76–78, Jan 2000. Online URL. (Cited on page 60.)

[Fabre 11] Claude Fabre. Atomes et Lumière: Interaction matière rayonnement. Lecture, November 2011. (Cited on page 8.)

[Froufe-Pérez 17] Luis S. Froufe-Pérez, Michael Engel, Juan José Sáenz & Frank Scheffold. *Band gap formation and Anderson localization in disordered photonic materials with structural correlations.* Proceedings of the National Academy of Sciences, vol. 114, no. 36, pages 9570–9574, 2017. Online URL. (Cited on page 148.)

[Garcia 17] Lionel Garcia, Olivier Pinel, Pu Jian, Nicolas Barré, Lionel Jaffrès, Jean-François Morizur & Guillaume Labroille. *Fast adaptive laser shaping based on multiple laser incoherent combining*, 2017. (Cited on page 108.)

[Gazit 09] Snir Gazit, Alexander Szameit, Yonina C. Eldar & Mordechai Segev. *Super-resolution and reconstruction of sparse sub-wavelength images.* Opt. Express, vol. 17, no. 26, pages 23920–23946, Dec 2009. Online URL. (Cited on page 95.)

[Gérardin 14] Benoît Gérardin, Jérôme Laurent, Arnaud Derode, Claire Prada & Alexandre Aubry. *Full Transmission and Reflection of Waves Propagating through a Maze of Disorder.* Phys. Rev. Lett., vol. 113, page 173901, Oct 2014. Online URL. (Cited on page 133.)

[Goetschy 11] Arthur Goetschy. *Light in disordered atomic systems : Euclidean matrix theory of random lasing*, November 2011. (Cited on pages 120, 126, 127, and 177.)

[Goetschy 13] A. Goetschy & A. D. Stone. *Filtering Random Matrices: The Effect of Incomplete Channel Control in Multiple Scattering.* Phys. Rev. Lett., vol. 111, page 063901, Aug 2013. Online URL. (Cited on pages 136 and 146.)

[Goodman 76] J. W. Goodman. *Some fundamental properties of speckle∗.* J. Opt. Soc. Am., vol. 66, no. 11, pages 1145–1150, Nov 1976. Online URL. (Cited on page 135.)

[Goodman 96] Joseph W. Goodman. Introduction to fourier optics. McGraw-Hill, 1996. (Cited on page 31.)

[Goodman 00] Joseph W. Goodman. Statistical optics. Wiley, 2000. (Cited on pages 36, 39, 40, 43, and 45.)

[Gori 87] F. Gori, G. Guattari & C. Padovani. *Bessel-Gauss beams.* Optics Communications, vol. 64, no. 6, pages 491 – 495, 1987. Online URL. (Cited on page 18.)

[Gouy 88] Gouy. *ALBERT A. MICHELSON et EDWARD MORLEY. - On a method of making the wave-length of sodium light the actual and practical standard of lenght (Sur une méthode pour faire de la longueur d'onde de la lumière du sodium l'étalon actuel et pratique de longueur); Philosophical Magazine, t. XXIV, p. 463; 1887.* J. Phys. Theor. Appl., vol. 7, no. 1, pages 443–444, 1888. Online URL. (Cited on page 2.)

[Grier 03] David G. Grier. *A revolution in optical manipulation.* Nature, vol. 424, 2003. Online URL. (Cited on page 25.)

[Gross 15] Simon Gross, Nicolas Riesen, John D. Love & Michael J. Withford. *C-Band Mode-Selective Couplers Fabricated by the Femtosecond Laser Direct-Write Technique.* In Optical Fiber Communication Conference, page W3B.2. Optical Society of America, 2015. (Cited on page 28.)

[Grynberg 93] G. Grynberg, B. Lounis, P. Verkerk, J.-Y. Courtois & C. Salomon. *Quantized motion of cold cesium atoms in two- and three-dimensional optical potentials.* Phys. Rev. Lett., vol. 70, pages 2249–2252, Apr 1993. Online URL. (Cited on page 3.)

[Grynberg 10] Gilbert Grynberg, Alain Aspect & Claude Fabre. Introduction to quantum optics: from the semi-classical approach to quantized light. Cambridge University Press, 2010. (Cited on pages 22 and 61.)

[Guan 15] B. Guan, B. Ercan, N. K. Fontaine, R. Ryf, H. Chen, R. P. Scott, Y. Zhang & S. J. B. Yoo. *15-Spatial-mode photonic lanterns based on ultrafast laser inscription.* In 2015 European Conference on Optical Communication (ECOC), pages 1–3, Sept 2015. (Cited on page 28.)

[Gustafsson 01] M. G. L. Gustafsson. *Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy.* Journal of Microscopy, vol. 198, no. 2, pages 82–87, 2001. Online URL. (Cited on page 2.)

[Gutiérrez-Vega 00] J. C. Gutiérrez-Vega, M. D. Iturbe-Castillo & S. Chávez-Cerda. *Alternative formulation for invariant optical fields: Mathieu beams.* Opt. Lett., vol. 25, no. 20, pages 1493–1495, Oct 2000. Online URL. (Cited on page 18.)

[Haake 06] Fritz Haake. Quantum signatures of chaos. Springer-Verlag, Berlin, Heidelberg, 2006. (Cited on pages 123, 124, and 125.)

[Hell 94] Stefan W. Hell & Jan Wichmann. *Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-*

depletion fluorescence microscopy. Opt. Lett., vol. 19, no. 11, pages 780–782, Jun 1994. Online URL. (Cited on pages 2 and 51.)

[Helstrom 69] Carl W. Helstrom. *Quantum detection and estimation theory*. Journal of Statistical Physics, vol. 1, no. 2, pages 231–252, Jun 1969. Online URL. (Cited on page 57.)

[Helstrom 73] C. Helstrom. *Resolution of point sources of light as analyzed by quantum detection theory*. IEEE Transactions on Information Theory, vol. 19, no. 4, pages 389–398, (Cited on page 64.)

[Hill 92] P. M. Hill, R. Olshansky & W. K. Burns. *Optical polarization division multiplexing at 4 Gb/s*. IEEE Photonics Technology Letters, vol. 4, no. 5, pages 500–502, (Cited on page 26.)

[Hsu 04] Magnus T L Hsu, Vincent Delaubert, Ping Koy Lam & Warwick P Bowen. *Optimal optical measurement of small displacements*. Journal of Optics B: Quantum and Semiclassical Optics, vol. 6, no. 12, page 495, 2004. Online URL. (Cited on pages 60, 65, and 100.)

[Imry 86] Y. Imry. *Active transmission channels and universal conductance fluctuations*. EPL (Europhysics Letters), vol. 1, page 249, (Cited on page 134.)

[Jian 14] Pu Jian. *Quantum limits in range-finding measurements with optical frequency combs*. PhD thesis, UPMC, 2014. (Cited on pages 54, 57, 58, and 65.)

[Kani 06] J. Kani, K. Iwatsuki & T. Imai. *Optical multiplexing technologies for access-area applications*. IEEE Journal of Selected Topics in Quantum Electronics, vol. 12, no. 4, pages 661–668, (Cited on page 26.)

[Kaplan 01] Ariel Kaplan, Nir Friedman & Nir Davidson. *Acousto-optic lens with very fast focus scanning*. Opt. Lett., vol. 26, no. 14, pages 1078–1080, Jul 2001. Online URL. (Cited on page 107.)

[Keiser 99] Gerd E. Keiser. *A Review of WDM Technology and Applications*. Optical Fiber Technology, vol. 5, no. 1, pages 3 – 39, 1999. Online URL. (Cited on page 26.)

[Klar 00] Thomas A. Klar, Stefan Jakobs, Marcus Dyba, Alexander Egner & Stefan W. Hell. *Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission.* vol. 97, no. 15, pages 8206–8210, (Cited on page 51.)

[Kogan 95] Eugene Kogan & Moshe Kaveh. *Random-matrix-theory approach to the intensity distributions of waves propagating in a random medium.* Phys. Rev. B, vol. 52, pages R3813–R3815, Aug 1995. Online URL. (Cited on page 136.)

[Kolobov 00] Mikhail I. Kolobov & Claude Fabre. *Quantum Limits on Optical Resolution.* Phys. Rev. Lett., vol. 85, pages 3789–3792, Oct 2000. Online URL. (Cited on page 48.)

[Labroille 14] Guillaume Labroille, Bertrand Denolle, Pu Jian, Philippe Genevaux, Nicolas Treps & Jean-François Morizur. *Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion.* Opt. Express, vol. 22, no. 13, pages 15599–15607, Jun 2014. Online URL. (Cited on pages 3, 28, and 100.)

[Landau 61] H. J. Landau & H. O. Pollak. *Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty — II.* Bell System Technical Journal, vol. 40, no. 1, pages 65–84, 1961. Online URL. (Cited on page 18.)

[Lassen 07] M. Lassen, V. Delaubert, J. Janousek, K. Wagner, H.-A. Bachor, P. K. Lam, N. Treps, P. Buchhave, C. Fabre & C. C. Harb. *Tools for Multimode Quantum Information: Modulation, Detection, and Spatial Quantum Correlations.* Phys. Rev. Lett., vol. 98, page 083602, Feb 2007. Online URL. (Cited on pages 65 and 100.)

[Leon-Saval 05] S. G. Leon-Saval, T. A. Birks, J. Bland-Hawthorn & M. Englund. *Multimode fiber devices with single-mode performance.* Opt. Lett., vol. 30, no. 19, pages 2545–2547, Oct 2005. Online URL. (Cited on page 28.)

[Li 15] Dong Li, Lin Shao, Bi-Chang Chen, Xi Zhang, Mingshu Zhang, Brian Moses, Daniel E. Milkie, Jordan R. Beach,

John A. Hammer, Mithun Pasham, Tomas Kirchhausen, Michelle A. Baird, Michael W. Davidson, Pingyong Xu & Eric Betzig. *Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics.* Science, vol. 349, no. 6251, 2015. Online URL. (Cited on page 50.)

[Li 17] Long Li, Runzhou Zhang, Zhe Zhao, Guodong Xie, Peicheng Liao, Kai Pang, Haoqian Song, Cong Liu, Yongxiong Ren, Guillaume Labroille, Pu Jian, Dmitry Starodubov, Brittany Lynn, Robert Bock, Moshe Tur & Alan E. Willner. *High-Capacity Free-Space Optical Communications Between a Ground Transmitter and a Ground Receiver via a UAV Using Multiplexing of Multiple Orbital-Angular-Momentum Beams.* Scientific Reports, vol. 7, pages 2045–2322, Dec 2017. Online URL. (Cited on page 5.)

[Lin 10] Hung-Chun Lin & Yi-Hsin Lin. *A fast response and large electrically tunable-focusing imaging system based on switching of two modes of a liquid crystal lens.* Applied Physics Letters, vol. 97, no. 6, page 063505, 2010. Online URL. (Cited on page 107.)

[Lindinger 04] Albrecht Lindinger, Cosmin Lupulescu, Mateusz Plewicki, Franziska Vetter, Andrea Merli, Stefan M. Weber & Ludger Wöste. *Isotope Selective Ionization by Optimal Control Using Shaped Femtosecond Laser Pulses.* Phys. Rev. Lett., vol. 93, page 033001, Jul 2004. Online URL. (Cited on page 2.)

[Lo 61] Y. T. Lo. *On the Theoretical Limitation of a Radio Telescope in Determining the Sky Temperature Distribution.* Journal of Applied Physics, vol. 32, no. 10, pages 2052–2054, 1961. Online URL. (Cited on page 48.)

[Lohmann 93] Adolf W. Lohmann. *Image rotation, Wigner rotation, and the fractional Fourier transform.* J. Opt. Soc. Am. A, vol. 10, no. 10, pages 2181–2186, Oct 1993. Online URL. (Cited on pages 30 and 31.)

[Lord Rayleigh 79] F.R.S. Lord Rayleigh. *XXXI. Investigations in optics, with special reference to the spectroscope.* The London, Edinburgh, and

Dublin Philosophical Magazine and Journal of Science, vol. 8, no. 49, pages 261–274, 1879. Online URL. (Cited on page 44.)

[Lu 16] X.-M. Lu, R. Nair & M. Tsang. *Quantum-optimal detection of one-versus-two incoherent sources with arbitrary separation.* ArXiv e-prints, (Cited on pages 67 and 68.)

[Luo 14] Lian-Wee Luo, Noam Ophir, Christine P. Chen, Lucas H. Gabrielli, Carl B. Poitras, Keren Bergmen & Michal Lipson. *WDM-compatible mode-division multiplexing on a silicon chip.* Nature Communications, vol. 5, Jan 2014. Online URL. (Cited on page 28.)

[Lupo 16] Cosmo Lupo & Stefano Pirandola. *Ultimate Precision Bound of Quantum and Subwavelength Imaging.* Phys. Rev. Lett., vol. 117, page 190802, Nov 2016. Online URL. (Cited on pages 66 and 67.)

[Lustig 07] Michael Lustig, David Donoho & John M. Pauly. *Sparse MRI: The application of compressed sensing for rapid MR imaging.* Magnetic Resonance in Medicine, vol. 58, no. 6, pages 1182–1195, 2007. Online URL. (Cited on page 95.)

[Mallat 08] Stéphane Mallat. A wavelet tour of signal processing, third edition: The sparse way. Academic Press, Inc., Orlando, FL, USA, 3rd edition, 2008. (Cited on page 90.)

[Mandel 95] Leonard Mandel & Emil Wolf. Optical coherence and quantum optics. Cambridge University Press, 1995. (Cited on pages 36 and 42.)

[Marrucci 06] L. Marrucci, C. Manzo & D. Paparo. *Optical Spin-to-Orbital Angular Momentum Conversion in Inhomogeneous Anisotropic Media.* Phys. Rev. Lett., vol. 96, page 163905, Apr 2006. Online URL. (Cited on page 24.)

[Marrucci 13] Lorenzo Marrucci. *The q-plate and its future.* Journal of Nanophotonics, vol. 7, pages 7 – 7 – 5, 2013. Online URL. (Cited on page 24.)

[Marčenko 67] V A Marčenko & L A Pastur. *Distribution of Eigenvalues for some sets of Random Matrices.* Mathematics of the USSR-Sbornik, vol. 1, no. 4, page 457, 1967. Online URL. (Cited on page 125.)

[Maurer 10] C. Maurer, A. Jesacher, S. Bernet & M. Ritsch-Marte. *What spatial light modulators can do for optical microscopy.* Laser & Photonics Reviews, vol. 5, no. 1, pages 81–101, 2010. Online URL. (Cited on page 25.)

[Mazilu 10] M. Mazilu, D.James Stevenson, F. Gunn-Moore & K. Dholakia. *Light beats the spread: "non-diffracting" beams.* Laser & Photonics Reviews, vol. 4, no. 4, pages 529–547, 2010. Online URL. (Cited on page 18.)

[McGloin 05] D McGloin & K Dholakia. *Bessel beams: Diffraction in a new light.* Contemporary Physics, vol. 46, no. 1, pages 15–28, 2005. Online URL. (Cited on page 18.)

[McLeod 54] John H. McLeod. *The Axicon: A New Type of Optical Element.* J. Opt. Soc. Am., vol. 44, no. 8, pages 592–597, Aug 1954. Online URL. (Cited on page 24.)

[Mehta 91] M. L. Mehta. Random matrices. Academic, New York, 1991. (Cited on page 119.)

[Mello 88] P.A Mello, P. Pereyra & N. Kumar. *Macroscopic approach to multichannel disordered conductors.* Annals of Physics, vol. 181, no. 2, pages 290 – 317, 1988. Online URL. (Cited on page 132.)

[Mendlovic 93] David Mendlovic & Haldun M. Ozaktas. *Fractional Fourier transforms and their optical implementation: I.* J. Opt. Soc. Am. A, vol. 10, no. 9, pages 1875–1881, Sep 1993. Online URL. (Cited on page 31.)

[Mermillod-Blondin 08] Alexandre Mermillod-Blondin, Euan McLeod & Craig B. Arnold. *High-speed varifocal imaging with a tunable acoustic gradient index of refraction lens.* Opt. Lett., vol. 33, no. 18, pages 2146–2148, Sep 2008. Online URL. (Cited on page 107.)

[Michelson 87] A. A. Michelson & E. W. Morley. *On a method of making the wavelength of sodium light the actual and practical standard of length.* American Journal of Science, vol. Series 3 Vol. 34, no. 204, pages 427–430, 1887. Online URL. (Cited on page 2.)

[Michelson 90] Albert A. Michelson. *I. On the application of interference methods to astronomical measurements.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, vol. 30, no. 182, pages 1–21, 1890. Online URL. (Cited on page 2.)

[Mignardi 16] Michael Mignardi. *The pioneering work that led to the DMD.* 2016. Online URL. (Cited on page 25.)

[Moerner 89] W. E. Moerner & L. Kador. *Optical detection and spectroscopy of single molecules in a solid.* Phys. Rev. Lett., vol. 62, pages 2535–2538, May 1989. Online URL. (Cited on page 51.)

[Moffitt 08] Jeffrey R. Moffitt, Yann R. Chemla, Steven B. Smith & Carlos Bustamante. *Recent Advances in Optical Tweezers.* Annual Review of Biochemistry, vol. 77, no. 1, pages 205–228, 2008. Online URL. PMID: 18307407. (Cited on page 3.)

[Montie 91] E. A. Montie, E. C. Cosman, G. W. 't Hooft, M. B. van der Mark & C. W. J. Beenakker. *Observation of the optical analogue of quantized conductance of a point contact.* Nature, vol. 350, (Cited on page 134.)

[Morizur 10] Jean-François Morizur, Lachlan Nicholls, Pu Jian, Seiji Armstrong, Nicolas Treps, Boris Hage, Magnus Hsu, Warwick Bowen, Jiri Janousek & Hans-A. Bachor. *Programmable unitary spatial mode manipulation.* J. Opt. Soc. Am. A, vol. 27, no. 11, pages 2524–2531, Nov 2010. Online URL. (Cited on pages 3 and 28.)

[Morizur 11a] J.-F. Morizur, S. Armstrong, N. Treps, J. Janousek & H.-A. Bachor. *Spatial reshaping of a squeezed state of light.* The European Physical Journal D, vol. 61, no. 1, pages 237–239, Jan 2011. Online URL. (Cited on page 5.)

[Morizur 11b] Jean-François. Morizur. *Quantum protocols with transverse spatial modes*. PhD thesis, UPMC & ANU, 2011. (Cited on pages 3, 5, 28, and 30.)

[Murphy Jr. 08] T. W. Murphy Jr., E. G. Adelberger, J. B. R. Battat, L. N. Carey, C. D. Hoyle, P. LeBlanc, E. L. Michelsen, K. Nordtvedt, A. E. Orin, J. D. Strasburg, C. W. Stubbs, H. E. Swanson & E. Williams. *The Apache Point Observatory Lunar Laser-ranging Operation: Instrument Description and First Detections*. Publications of the Astronomical Society of the Pacific, vol. 120, no. 863, page 20, 2008. Online URL. (Cited on page 2.)

[Murshid 08] Syed Murshid, B. Grossman & P. Narakorn. *Spatial domain multiplexing: A new dimension in fiber optic multiplexing*. Optics & Laser Technology, vol. 40, no. 8, pages 1030 – 1036, 2008. Online URL. (Cited on page 26.)

[Nair 16] Ranjith Nair & Mankei Tsang. *Far-Field Superresolution of Thermal Electromagnetic Sources at the Quantum Limit*. Phys. Rev. Lett., vol. 117, page 190801, Nov 2016. Online URL. (Cited on page 66.)

[Nazarov 94] Yu. V. Nazarov. *Limits of universality in disordered conductors*. Phys. Rev. Lett., vol. 73, pages 134–137, Jul 1994. Online URL. (Cited on page 136.)

[Noordegraaf 09] Danny Noordegraaf, Peter M. W. Skovgaard, Martin D. Nielsen & Joss Bland-Hawthorn. *Efficient multi-mode to single-mode coupling in a photonic lantern*. Opt. Express, vol. 17, no. 3, pages 1988–1994, Feb 2009. Online URL. (Cited on page 27.)

[Oron 05] Dan Oron, Eran Tal & Yaron Silberberg. *Scanningless depth-resolved microscopy*. Opt. Express, vol. 13, no. 5, pages 1468–1476, Mar 2005. Online URL. (Cited on page 107.)

[Ozaktas 93] Haldun M. Ozaktas & David Mendlovic. *Fractional Fourier transforms and their optical implementation. II*. J. Opt. Soc. Am. A, vol. 10, no. 12, pages 2522–2531, Dec 1993. Online URL. (Cited on page 31.)

[Paúr 16] Martin Paúr, Bohumil Stoklasa, Zdenek Hradil, Luis L. Sánchez-Soto & Jaroslav Rehacek. *Achieving the ultimate optical resolution.* Optica, vol. 3, no. 10, pages 1144–1147, Oct 2016. Online URL. (Cited on pages 68 and 69.)

[Pinel 10] Olivier Pinel. *Optique quantique multimode avec des peignes de fréquence.* PhD thesis, UPMC, 2010. (Cited on pages 54, 57, and 65.)

[Pinel 12] Olivier Pinel, Julien Fade, Daniel Braun, Pu Jian, Nicolas Treps & Claude Fabre. *Ultimate sensitivity of precision measurements with intense Gaussian quantum light: A multimodal approach.* Phys. Rev. A, vol. 85, page 010101, Jan 2012. Online URL. (Cited on page 65.)

[Popoff 10] S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara & S. Gigan. *Measuring the Transmission Matrix in Optics: An Approach to the Study and Control of Light Propagation in Disordered Media.* Phys. Rev. Lett., vol. 104, page 100601, Mar 2010. Online URL. (Cited on page 136.)

[Popoff 11] S M Popoff, G Lerosey, M Fink, A C Boccara & S Gigan. *Controlling light through optical disordered media: transmission matrix approach.* New Journal of Physics, vol. 13, no. 12, page 123021, 2011. Online URL. (Cited on page 136.)

[Rao 47] C. Radhakrishna Rao. *Minimum variance and the estimation of several parameters.* Mathematical Proceedings of the Cambridge Philosophical Society, vol. 43, no. 2, page 280–283, (Cited on page 57.)

[Reck 94] Michael Reck, Anton Zeilinger, Herbert J. Bernstein & Philip Bertani. *Experimental realization of any discrete unitary operator.* Phys. Rev. Lett., vol. 73, pages 58–61, Jul 1994. Online URL. (Cited on page 34.)

[Reddy 07] Duemani Reddy & Peter Saggau. *Development of a random access multiphoton microscope for fast three-dimensional functional recording of neuronal activity*, 2007. (Cited on page 107.)

[Rehacek 17a] J. Rehacek, Z. Hradil, B. Stoklasa, M. Paúr, J. Grover, A. Krzic & L. L. Sánchez-Soto. *Multiparameter quantum metrology of incoherent point sources: Towards realistic superresolution.* Phys. Rev. A, vol. 96, page 062107, Dec 2017. Online URL. (Cited on page 67.)

[Rehacek 17b] J. Rehacek, M. Paúr, B. Stoklasa, Z. Hradil & L. L. Sánchez-Soto. *Optimal measurements for resolution beyond the Rayleigh limit.* Opt. Lett., vol. 42, no. 2, pages 231–234, Jan 2017. Online URL. (Cited on page 68.)

[Réfrégier 02] Philippe Réfrégier. Théorie du bruit et applications en physique. Lavoisier, 2002. (Cited on pages 51 and 54.)

[Richardson 13] D. J. Richardson, J. M. Fini & L. E. Nelson. *Space-division multiplexing in optical fibres.* Nature Photonics, vol. 7, page 354–362, 2013. Online URL. (Cited on pages 26 and 27.)

[Römer 14] G.R.B.E. Römer & P. Bechtold. *Electro-optic and Acousto-optic Laser Beam Scanners.* Physics Procedia, vol. 56, pages 29 – 39, 2014. Online URL. 8th International Conference on Laser Assisted Net Shape Engineering LANE 2014. (Cited on page 107.)

[Roddier 99] François Roddier. Adaptive optics in astronomy. Cambridge University Press, 1999. (Cited on pages 2 and 25.)

[Rosales-Guzmán 17] Carmelo Rosales-Guzmán & Andrew Forbes. How to shape light with spatial light modulators. SPIE Press Book, 2017. (Cited on page 24.)

[Rotter 17] Stefan Rotter & Sylvain Gigan. *Light fields in complex media: Mesoscopic scattering meets wave control.* Rev. Mod. Phys., vol. 89, page 015005, Mar 2017. Online URL. (Cited on pages 120, 128, 130, 134, and 135.)

[Ryf 12] Roland Ryf, Sebastian Randel, Alan H. Gnauck, Cristian Bolle, Alberto Sierra, Sami Mumtaz, Mina Esmaeelpour, Ellsworth C. Burrows, René-Jean Essiambre, Peter J. Winzer,

David W. Peckham, Alan H. McCurdy & Robert Lingle. *Mode-Division Multiplexing Over 96 km of Few-Mode Fiber Using Coherent 6 × 6 MIMO Processing.* J. Lightwave Technol., vol. 30, no. 4, pages 521–531, Feb 2012. Online URL. (Cited on page 28.)

[Ryf 14] R. Ryf, N. K. Fontaine, M. Montoliu, S. Randel, B. Ercan, H. Chen, S. Chandrasekhar, A. H. Gnauck, S. G. Leon-Saval, J. Bland-Hawthorn, J. R. Salazar-Gil, Y. Sun & R. Lingle. *Photonic-Lantern-Based Mode Multiplexers for Few-Mode-Fiber Transmission.* In Optical Fiber Communication Conference, page W4J.2. Optical Society of America, 2014. (Cited on page 28.)

[Sancataldo 17] Giuseppe Sancataldo, Lorenzo Scipioni, Tiziana Ravasenga, Luca Lanzanò, Alberto Diaspro, Andrea Barberis & Martí Duocastella. *Three-dimensional multiple-particle tracking with nanometric precision over tunable axial ranges.* Optica, vol. 4, no. 3, pages 367–373, Mar 2017. Online URL. (Cited on page 107.)

[Sarma 16] Raktim Sarma, Alexey G. Yamilov, Sasha Petrenko, Yaron Bromberg & Hui Cao. *Control of Energy Density inside a Disordered Medium by Coupling to Open or Closed Channels.* Phys. Rev. Lett., vol. 117, page 086803, Aug 2016. Online URL. (Cited on page 137.)

[Sato 85] Susumu Sato, Akira Sugiyama & Rumiko Sato. *Variable-Focus Liquid-Crystal Fresnel Lens.* Japanese Journal of Applied Physics, vol. 24, no. 8A, page L626, 1985. Online URL. (Cited on page 107.)

[Schwartz 09] Noah H. Schwartz, Nicolas Védrenne, Vincent Michau, Marie-Thérèse Velluet & Frédéric Chazallet. *Mitigation of atmospheric effects by adaptive optics for free-space optical communications*, 2009. (Cited on page 2.)

[Shechtman 10] Yoav Shechtman, Snir Gazit, Alexander Szameit, Yonina C. Eldar & Mordechai Segev. *Super-resolution and reconstruction of sparse images carried by incoherent light.* Opt. Lett.,

vol. 35, no. 8, pages 1148–1150, Apr 2010. Online URL. (Cited on page 95.)

[Shibaguchi 92] Takashi Shibaguchi & Hiroyoshi Funato. *Lead-Lanthanum Zirconate-Titanate (PLZT) Electrooptic Variable Focal-Length Lens with Stripe Electrodes.* Japanese Journal of Applied Physics, vol. 31, no. 9S, page 3196, 1992. Online URL. (Cited on page 107.)

[Siegman 86] Anthony E. Siegman. Lasers. University Science Books, Mill Valley, CA, 1986. (Cited on pages 1, 5, 11, and 12.)

[Siviloglou 07] G. A. Siviloglou, J. Broky, A. Dogariu & D. N. Christodoulides. *Observation of Accelerating Airy Beams.* Phys. Rev. Lett., vol. 99, page 213901, Nov 2007. Online URL. (Cited on page 18.)

[Slepian 61] D. Slepian & H. O. Pollak. *Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty — I.* Bell System Technical Journal, vol. 40, no. 1, pages 43–63, 1961. Online URL. (Cited on page 18.)

[Slepian 83] D. Slepian. *Some comments on Fourier analysis, uncertainty and modeling.* SIAM Review, vol. 25, no. 3, pages 379–393, 1983. Online URL. (Cited on page 18.)

[Smith 70] Richard L. Smith. *The Velocities of Light.* American Journal of Physics, vol. 38, no. 8, pages 978–984, 1970. Online URL. (Cited on page 21.)

[Snyder 83] Allan W. Snyder & John D. Love. Optical waveguide theory. Springer, 1983. (Cited on pages 20 and 21.)

[Soma 18] Daiki Soma, Yuta Wakayama, Shohei Beppu, Seiya Sumita, Takehiro Tsuritani, Tetsuya Hayashi, Takuji Nagashima, Masato Suzuki, Masato Yoshida, Keisuke Kasai, Masataka Nakazawa, Hidenori Takahashi, Koji Igarashi, Itsuro Morita & Masatoshi Suzuki. *10.16-Peta-B/s Dense SDM/WDM Transmission Over 6-Mode 19-Core Fiber Across the C+L Band.* J. Lightwave Technol., vol. 36, no. 6, pages 1362–1368, (Cited on page 3.)

[Sparrow 16] C. M. Sparrow. *On Spectroscopic Resolving Power.* Astrophysical Journal, vol. 44, page 76, (Cited on page 44.)

[Steen 91] William Steen & Jyotirmoy Mazumder. Laser material processing. Springer, 1991. (Cited on page 5.)

[Stöckmann 90] H.-J. Stöckmann & J. Stein. *"Quantum" chaos in billiards studied by microwave absorption.* Phys. Rev. Lett., vol. 64, pages 2215–2218, May 1990. Online URL. (Cited on page 120.)

[Sun 14] Hengxin Sun, Kui Liu, Zunlong Liu, Pengliang Guo, Junxiang Zhang & Jiangrui Gao. *Small-displacement measurements using high-order Hermite-Gauss modes.* Applied Physics Letters, vol. 104, no. 12, page 121908, 2014. Online URL. (Cited on pages 65 and 69.)

[Tang 16] Zong Sheng Tang, Kadir Durak & Alexander Ling. *Fault-tolerant and finite-error localization for point emitters within the diffraction limit.* Opt. Express, vol. 24, no. 19, pages 22004–22012, Sep 2016. Online URL. (Cited on pages 68 and 69.)

[TeleGeography 18] TeleGeography. *Submarine Cable Map.* 2018. Online URL. (Cited on page 3.)

[Tham 17] Weng-Kian Tham, Hugo Ferretti & Aephraim M. Steinberg. *Beating Rayleigh's Curse by Imaging Using Phase Information.* Phys. Rev. Lett., vol. 118, page 070801, Feb 2017. Online URL. (Cited on pages 68 and 69.)

[Treps 04] N Treps, N Grosse, W P Bowen, M T L Hsu, A Maître, C Fabre, H-A Bachor & P K Lam. *Nano-displacement measurements using spatially multimode squeezed light.* Journal of Optics B: Quantum and Semiclassical Optics, vol. 6, no. 8, page S664, 2004. Online URL. (Cited on page 60.)

[Treps 05] N. Treps, V. Delaubert, a. Maître, J. Courty & C. Fabre. *Quantum noise in multipixel image processing.* Physical Review A, vol. 71, no. 1, page 013820, jan 2005. Online URL. (Cited on pages 62 and 65.)

[Tsang 16a] M. Tsang, R. Nair & X.-M. Lu. *Semiclassical Theory of Superresolution for Two Incoherent Optical Point Sources.* ArXiv e-prints, (Cited on page 67.)

[Tsang 16b] Mankei Tsang, Ranjith Nair & Xiao-Ming Lu. *Quantum Theory of Superresolution for Two Incoherent Optical Point Sources.* Phys. Rev. X, vol. 6, page 031033, Aug 2016. Online URL. (Cited on pages 66, 67, and 68.)

[Tsang 17] Mankei Tsang. *Subdiffraction incoherent optical imaging via spatial-mode demultiplexing.* New Journal of Physics, vol. 19, no. 2, page 023054, 2017. Online URL. (Cited on pages 67 and 68.)

[Tsang 18] Mankei Tsang. *Conservative classical and quantum resolution limits for incoherent imaging.* Journal of Modern Optics, vol. 65, no. 11, pages 1385–1391, 2018. Online URL. (Cited on page 67.)

[Tulino 04] Antonia M. Tulino & Sergio Verdú. *Random Matrix Theory and Wireless Communications.* Foundations and Trends® in Communications and Information Theory, vol. 1, no. 1, pages 1–182, 2004. Online URL. (Cited on pages 120, 125, 127, and 177.)

[Tyson 10] Robert Tyson. Principles of adaptive optics. CRC Press, 2010. (Cited on pages 2 and 25.)

[van Wees 88] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel & C. T. Foxon. *Quantized conductance of point contacts in a two-dimensional electron gas.* Phys. Rev. Lett., vol. 60, pages 848–850, Feb 1988. Online URL. (Cited on page 134.)

[Vellekoop 08a] I. M. Vellekoop & A. P. Mosk. *Universal Optimal Transmission of Light Through Disordered Materials.* Phys. Rev. Lett., vol. 101, page 120601, Sep 2008. Online URL. (Cited on pages 136 and 137.)

[Vellekoop 08b] Ivo Micha Vellekoop. *Controlling the propagation of light in disordered scattering media.* PhD thesis, 4 2008. (Cited on page 120.)

[Vellekoop 15] Ivo M. Vellekoop. *Feedback-based wavefront shaping*. Opt. Express, vol. 23, no. 9, pages 12189–12206, May 2015. Online URL. (Cited on page 2.)

[Velázquez-Benítez 18] Amado M. Velázquez-Benítez, J. Enrique Antonio-López, Juan C. Alvarado-Zacarías, Nicolas K. Fontaine, Roland Ryf, Haoshuo Chen, Juan Hernández-Cordero, Pierre Sillard, Chigo Okonkwo, Sergio G. Leon-Saval & Rodrigo Amezcua-Correa. *Scaling photonic lanterns for space-division multiplexing*. Scientific Reports, vol. 8, (Cited on page 27.)

[Verkerk 92] P. Verkerk, B. Lounis, C. Salomon, C. Cohen-Tannoudji, J.-Y. Courtois & G. Grynberg. *Dynamics and spatial order of cold cesium atoms in a periodic optical potential*. Phys. Rev. Lett., vol. 68, pages 3861–3864, Jun 1992. Online URL. (Cited on page 3.)

[Voiculescu 83] D. Voiculescu. *Asymptotically commuting finite rank unitary operators without commuting approximants*. Acta Sci. Math., vol. 45, pages 429–431, (Cited on page 125.)

[Wang 12] Hui Wang, Shensheng Han & Mikhail I. Kolobov. *Quantum limits of super-resolution of optical sparse objects via sparsity constraint*. Opt. Express, vol. 20, no. 21, pages 23235–23252, Oct 2012. Online URL. (Cited on page 95.)

[Wang 17] Li-Lian Wang. *A Review of Prolate Spheroidal Wave Functions from the Perspective of Spectral Methods*. vol. 50, pages 101–143, (Cited on pages 18 and 20.)

[Weaver 89] R. L. Weaver. *Spectral statistics in elastodynamics*. The Journal of the Acoustical Society of America, vol. 85, no. 3, pages 1005–1013, 1989. Online URL. (Cited on page 120.)

[Weiner 00] A. M. Weiner. *Femtosecond pulse shaping using spatial light modulators*. Review of Scientific Instruments, vol. 71, no. 5, pages 1929–1960, 2000. Online URL. (Cited on page 25.)

[Weisenburger 17] Siegfried Weisenburger, Daniel Boening, Benjamin Schomburg, Karin Giller, Stefan Becker, Christian Griesinger &

Vahid Sandoghdar. *Cryogenic optical localization provides 3D protein structure data with Angstrom resolution.* Nature Methods, vol. 14, page 141, Jan 2017. Online URL. (Cited on page 51.)

[Whittaker 03] E. T. Whittaker. *On the partial differential equations of mathematical physics.* Mathematische Annalen, vol. 57, no. 3, pages 333–355, Sep 1903. Online URL. (Cited on page 17.)

[Willett 11] Rebecca M. Willett, Roummel F. Marcia & Jonathan M. Nichols. *Compressed sensing for practical optical imaging systems: a tutorial.* Optical Engineering, vol. 50, pages 50 – 50 – 14, 2011. Online URL. (Cited on page 95.)

[Wishart 28] John Wishart. *The generalised product moment distribution in samples from a normal multivariate population.* Biometrika, vol. 20A, no. 1-2, pages 32–52, 1928. Online URL. (Cited on page 119.)

[Wolter 61] Hans Wolter. *V On Basic Analogies and Principal Differences Between Optical and Electronic Information.* volume 1 of *Progress in Optics*, pages 155 – 210. Elsevier, 1961. Online URL. (Cited on page 48.)

[Yang 06] Victor X. D. Yang, Youxin Mao, Beau A. Standish, Nigel R. Munce, Stephanie Chiu, Daina Burnes, Brian C. Wilson, I. Alex Vitkin, Phillip A. Himmer & David L. Dickensheets. *Doppler optical coherence tomography with a micro-electromechanical membrane mirror for high-speed dynamic focus tracking.* Opt. Lett., vol. 31, no. 9, pages 1262–1264, May 2006. Online URL. (Cited on page 107.)

[Yang 16] Fan Yang, Arina Tashchilina, E. S. Moiseev, Christoph Simon & A. I. Lvovsky. *Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode.* Optica, vol. 3, no. 10, pages 1148–1152, Oct 2016. Online URL. (Cited on pages 68 and 69.)

[Yang 17] Fan Yang, Ranjith Nair, Mankei Tsang, Christoph Simon & Alexander I. Lvovsky. *Fisher information for far-field linear optical superresolution via homodyne or heterodyne detection*

in a higher-order local oscillator mode. Phys. Rev. A, vol. 96, page 063829, Dec 2017. Online URL. (Cited on page 68.)

[Yu 15] Hyeonseung Yu, Jongchan Park, KyeoReh Lee, Jonghee Yoon, KyungDuk Kim, Shinwha Lee & YongKeun Park. *Recent advances in wavefront shaping techniques for biomedical applications.* Current Applied Physics, vol. 15, no. 5, pages 632 – 641, 2015. Online URL. (Cited on page 2.)

[Zhu 05] Guanghao Zhu, James van Howe, Michael Durst, Warren Zipfel & Chris Xu. *Simultaneous spatial and temporal focusing of femtosecond pulses.* Opt. Express, vol. 13, no. 6, pages 2153–2159, Mar 2005. Online URL. (Cited on page 107.)