Numerical Modelling and Measurement of Spatial Coherence in a Lithographic System

by Arlene Smith

Supervisor: Prof. Chris Dainty

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy,

Applied Optics, School of Physics,
National University of Ireland Galway

September 2011

Abstract

For the last five decades, optical lithography has been used universally for the manufacture of integrated circuits. The drive towards smaller circuit feature size, and the resulting increase in resolution of lithography systems, has been achieved by changing the source type. The original $300 - 400$ *nm* discharge lamps have been replaced with more powerful and monochromatic excimer lasers, operating at user-selectable Ultra-Violet (UV) wavelengths from $157 - 351$ *nm*. The success of UV lithography is due, in part, to the development of industrialized, line narrowed excimer lasers. However, the line narrowing process, along with the intrinsic coherence properties of the laser, results in a source with a considerable amount of spatial coherence. These partially coherent sources are complicated to model. Partially coherent image formation with two-dimensional intensity fields requires evaluating four-dimensional integrals. Thus calculations are complex, slow to process and place demands on system memory. The motivation behind this thesis is an improved understanding of the effects of spatial coherence in optical lithography, including beam homogenization in wafer-stepper systems. In this thesis, we expand the Elementary Function Method, a process similar to coherent mode decomposition, to develop a numerical model of a partially spatially coherent source. We design an experiment to measure the spatial coherence of a partially coherent laser source and present our results. The numerical model is tested on a theoretical Gaussian Schell-model source, and then applied to the excimer source. We examine the traditional method for beam conditioning, the imaging homogenizer, and develop a model of the physical process involved in beam homogenization. Our results show that the imaging homogenizer is designed for use with spatially incoherent sources, where it performs well. However, if the source incident on the homogenizer has any degree of spatial coherence, the intensity output is highly non-uniform.

Acknowledgements

Firstly I would like to thank my supervisor Chris Dainty for his guidance and support during my PhD studies. I have benefited hugely from his enthusiasm, advice, and trust during my time as his student. I would also like to thank everyone in the Applied Optics Group, especially Conor, Matt, and Ruth, for valuable discussions, advice and good humour over the years. I have been very lucky to be surrounded by such a great circle of people, and I will never forget my time in Applied Optics. I am especially indebted to Anna Burvall for her contributions to this project, and for her advice and encouragement even after she moved on from the group. I also wish to thank Graeme Scott at Hewlett Packard for his technical support, advice and patience.

Several parts of this project would not have been possible without technical support from the School of Physics, especially Conor McBrierty. Thank you. Thanks also to the NCLA for the use of their excimer laser, and to Claire O'Connell and Kevin Conlisk for their time and help in the laboratory.

I also wish to thank my family and friends. My parents have been a constant support to me through my (many!) years of education, and have been unwavering in their encouragement and belief in me; thanks Mam and Dad. Thanks also to Kathleen and Paddy for always being there. A huge thank you to my friends Caroline, Catherine, Elaine, Elizabeth and Pauline; I owe you all a great debt (and alot of wine).

Thank you to all the physics gang, especially Brendan and Caoilfhionn, for tea breaks and lots of good times! Thanks also to Tess for her support, friendship, and constant good cheer.

Finally, thank you, Conor, for being here.

This research was supported by the Irish Research Council for Science, Engineering and Technology and by Hewlett Packard under the Enterprise Partnership Scheme and through Science Foundation Ireland under Grant No SFI/07/IN.1/1906.

Contents

List of Figures

Preface

From the late 1960's, when integrated circuits had linewidths of 5 μm, to the last year when the 45nm node threshold was crossed, optical lithography has been used universally for manufacturing. This dominance of optical lithography in production is the result of a worldwide effort to improve optical imaging systems and resists. The International Technology Roadmap for Semiconductors [1] (ITRS, or The Roadmap) has become a forum for listing and updating lithography requirements. It is sponsored by the five leading chip manufacturing regions in the world: Europe, Japan, Korea, Taiwan, and the United States. Its goal is to ensure cost-effective advancements in the performance of the integrated circuits and the products that use them. Its milestones, the "nodes", are the half-pitch of the smallest dimensions that need to be patterned in every generation of microelectronic devices. The industry has followed an impressive path with node dimensions decreasing steadily for the last 50 years. This uninterrupted improvement is part of "Moore's Law", named after Intel co-founder Gordon Moore, and states that the number of transistors that can be placed inexpensively on an integrated circuit will double every two years [2]. According to the ITRS, nodes are expected to reach 11 nm by 2022.

Although lithography system costs increase as minimum feature size on a semiconductor chip decreases, optical lithography remains attractive because of its high wafer throughput and relatively low cost, relative to alternative technologies such as X-ray lithography and E-beam lithography.

Thesis Motivation

The focus of this project is an improved understanding of the effects of spatial coherence in optical lithography, including beam homogenization in wafer stepper systems. Excimer lasers operating in the Deep Ultra-Violet (DUV) have been used in lithographic systems for over a decade. Their high brightness and limited spatial and temporal coherence make them an attractive source for circuit patterning. However, the wide bandwidth of excimer lasers causes problems in designing DUV photolithographic optics. Line-narrowing processes reduce chromatic aberration in refractive optical systems, but have the disadvantage of increasing coherence and causing problems with interference and speckle. Increased spatial coherence brings problems with interference in conventional beam homogenization systems. Optical modelling of such partially coherent sources would be beneficial in the design of efficient beam homogenization techniques.

Thesis Organisation

The primary focus of this work is the development of a method to simplify the equations used to propagate partially coherent light through imaging systems. In **Chapter 1**, the typical components of a lithographic imaging system are examined. In **Chapter 2**, we look at the origin of coherence, arriving at a definition for temporal and spatial coherence. Mathematical methods for the propagation of spatially partially coherent light are discussed in **Chapter 3**, where we arrive at a modal approach of decomposing the beam into elementary functions which are then propagated. In order to make this beam model system-specific, spatial coherence measurements are necessary. In **Chapter 4**, we present a method to measure the spatial coherence of a laser source and the experimental results. In **Chapter 5**, the Elementary Function method is verified numerically using the example of a Gaussian Schell-model beam, and the model is then applied to a real excimer laser beam. In **Chapter 6**, the Elmentary Function Method is extended to include beam homogenization optics, and the consquences of imaging with a partially coherent source are discussed. A brief summary of results including the conclusions drawn from this work are presented in **Chapter 7**.

Publications and Presentations

- A. Smith, A. Burvall, and C. Dainty. Numerical partially coherent imaging using elementary functions. *Proc. SPIE Optical Design and Engineering III*, 7100:710004, 2008.
- A. Smith, A. Burvall, and C. Dainty. Partially coherent image computation using elementary functions. *Proc. SPIE Optical Microlithography XXII*, 7274:727434, 2009.
- A. Burvall, A. Smith, and C. Dainty. Elementary functions: propagation of partially coherent light. *J. Opt. Soc. Am. A*, 26:1721-1729, 2009.
- A. Smith, A. Burvall, and C. Dainty. Partial spatial coherence in an excimer-laser lithographic imaging system. *Proc. SPIE Optical Microlithography XXIII*, 7640:764029, 2010.

Chapter 1

Introduction to Lithographic Systems

1.1 Source

Until the late 1980s, optical lithography was performed with high-pressure Hg discharge lamps operating at 436 *nm* (G line) and 365 *nm* (I line). The transition to shorter wavelengths was first performed with Hg discharge lamps operating at ~ 250 *nm*, but these sources were soon replaced with more powerful and monochromatic excimer lasers. Excimer lasers are a class of very efficient and powerful pulsed ultraviolet lasers operating in a gas-filled optical cavity. They are typically used in high resolution material processing and operate at user-selectable UV wavelengths from $157 - 351$ *nm* depending on the lasing molecule (e.g. KrF, ArF, XeCl). In fact, the term *excimer* is outdated; modern lasers of this type are strictly speaking *exciplex* lasers. Excimer, originating from *excited dimer*, refers to the compound of two identical species which exists only in the excited state. Excited diatomic molecules such as Xenon (Xe_2) were originally used as the laser gas. Modern excimer lasers use an excited complex of rare gas and halogen - such as argon and flouride - and are exciplex lasers. Currently, high resolution lithographic systems are in place, operating at wavelengths of 248 *nm*, 193 *nm* and 157 *nm*, in the Deep-Ultraviolet (DUV) [3–7].

The success of 248 *nm* lithography is due, in part, to the development of reliable, industralized, line narrowed excimer laser sources. Their most important properties, which set an excimer laser apart from most other lasers and which makes them most suitable for lithographic systems, are their high brightness and limited temporal and spatial coherence. The system in this study incorporates a 248 *nm* KrF excimer laser system. KrF excimer lasers are often line-narrowed for use in lithographic systems. Line-narrowing is carried out to reduce the spectral range of the laser, minimising chromatic aberration. Freely running KrF lasers have a natural bandwidth of approximately 300 *pm*, which is too high for high-resolution wafer steppers. All-refractive optics require bandwidths of < 1.0 *pm*, and even catadioptric, moderate-NA systems require bandwidths < 100 *pm* [8]. Also, building an illumination system at 248 *nm* is tricky. The choice of materials with adequate transmission is limited; fused silica and calcium fluoride are commonly used. They are close in refractive index and dispersion, thus chromatic correction requires high optical power lenses that are difficult to fabricate and align. The solution to this is to narrow the spectral bandwidth of the laser, rather than achromatize the lens.

Line-narrowing is most commonly achieved by inserting intra-cavity wavelength dispersive optical elements into the laser cavity. The narrowband output can be tuned over a region which is typically close to the free running bandwidth of the laser. The 248 *nm* laser used in this system has been line-narrowed to 10 *pm*. While line-narrowing processes reduce chromatic aberrations in optical systems, they also increase coherence, which introduces problems with interference and speckle.

An analysis of lithography systems with excimer lasers can be found in [3] and [9].

1.2 Illumination System

The illumination system contains the optics used to homogenize the beam and focus it on the reticle holding the desired pattern. These systems comprise many elements and are expensive to produce (Figure 1.1). Excimer lasers do not have a uniform light intensity distribution in the plane perpendicular to the optic axis. The beam exiting the laser cavity has an approximately Gaussian intensity profile along the short axis; the intensity profile is approximately top-hat in the long axis of the beam. For lithographic applications, it is desirable that the illumination is uniform in the imaging plane, requiring ± 2 % uniformity. For uniformity of this level, it is necessary to condition the beam [10].

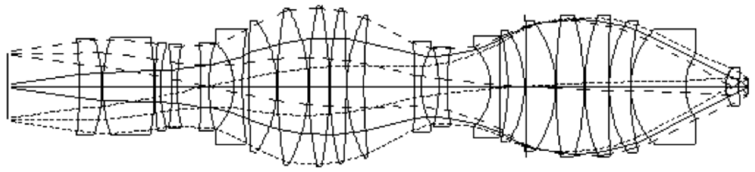

Figure 1.1: Lithographic objective from US Patent 5,805,344 (1998) with image side NA = 0.56, image field size $15 \times 15 mm^2$, working wavelength 248 nm, reduction ratio 0.25.

To date, various methods have been proposed for beam-shaping and homogenization: intensity modification of Gaussian beams by phase filtering or diffraction gratings [11–14], acousto-optical effects [15], saturable absorbers [16], a combination of a phase plate and light-conducting glass rods [17], and diffractive optical elements [18–20].

Most modern beam homogenizers fall into one of two categories: scramblers or beam-folders. Scramblers create a uniform beam by randomly mixing the laser beam, averaging out any intensity variations, e.g. using an incoherent fibre bundle, or a fly's eye lens [21], [22]. Scramblers are relatively inexpensive, but have several drawbacks including high divergence and low throughput. Beam-folding homogenizers use a different technique of overlapping multiple beam segments to average out intensity fluctuations [23–26]. This means that a loss in power density in one area of the beam does not affect the uniformity of the mask exposure plane. The advantages of these shapers are the independence from entrance intensity profile and wide spectrum of wavelengths. However, the periodic structure and the overlapping of beamlets produce interference effects when used with spatially highly coherent light. Nevertheless, in some cases, successful homogenization with these elements can be achieved with the consideration of physical optics, with the usage of additional elements like random diffusers, and, in certain cases, averaging over many exposures.

The beam homogenizer in this system is a Suss MicroOptics imaging homogenizer, incorporating two microlens arrays and a spherical lens (Figure 1.2). The square microlens arrays divide the incident beam into beamlets. These beamlets are then passed through a lens to be overlapped in the homogenization plane located at the back focal plane of the spherical lens. This lens, which Suss MicroOptics refer to as a "Fourier lens", causes the parallel bundles of rays to converge in the homogenization plane. An overview of beam homogenizing techniques can be found in [27]. A more detailed analysis of beam homogenization in lithographic applications can be found

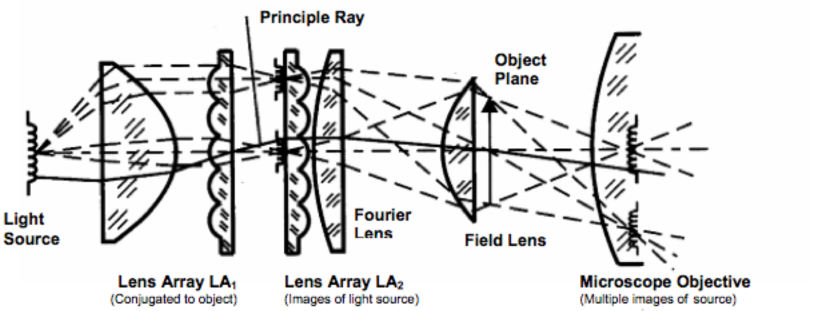

Figure 1.2: Layout of an imaging microlens beam homogenizer. Image courtesy of Suss MicroOptics.

in [28]. The effects of beam homogenization on partially spatially coherent sources will be investigated in Chapter 6.

1.3 Reticle

This is the mask containing information on the desired pattern to be projected onto the silicon layer. The mask (or reticle) is made up of lines, spaces and contacts. A line is an opaque section which light cannot pass through. A space is a clear section which is transparent to light. A contact is an clear section, transparent to light, surrounded on all sides by an opaque region. A combination of these features make up a standard reticle. Photomasks are fabricated with techniques similar to those used in wafer processing. A photomask blank, consisting of an opaque film (usually chromium or chromium-containing compounds) deposited on a glass or quartz substrate, is covered with resist. The resist is exposed according to the circuit pattern, is then developed, and the exposed opaque material is then etched. As many as thirty masks are used to produce the various layers on an Integrated Circuit (IC). The pattern creates dark and light regions that correspond to physical elements - typically part of transistors or connections between them - to be reproduced on each chip on the wafer. In early lithography systems, the transfer was 1–1 between the mask and the wafer for every layer, and the pattern on the mask was often produced through a lithography process - the reticle containing the circuit layout is patterned onto a mask which in turn is projected onto the wafer surface. The invention of the wafer stepper made this process more efficient: the circuit patterns are now stepped directly onto the wafer with a 4–1 or 5–1 reduction. The most common size for reticles today is the

6″/152.4 mm (0.25″/6.35 mm thick) format [8].

1.4 Projection Lens

The light is directed onto the silicon wafer using either a stepper or scanner system. The stepper moves or "steps" from individual areas on the wafer, exposing each one in turn. Previous generations of photolithographic equipment exposed the entire wafer at once. A stepper, working on a limited area, is capable of higher resolution. A scanner is a type of stepper which moves the wafer and reticle with respect to each other during the exposure, as a way of increasing the size of the exposed area. In a scanner system, instead of exposing the entire field at once, the exposure is made through an "exposure slit", that is as wide as the exposure field but only a fraction of its length. The image from the exposure slit is then scanned across the exposure area. The advantage of the scanning technique is that the optical properties of the projection lens can be optimized in the area through which the image of the projection slit passes, while optical aberrations outside this area can be ignored [29], [30].

1.5 Photoresist

In lithographic systems, the reticle is illuminated and the wafer is exposed to UV light through the mask. The wafer is coated with a layer of photosensitive chemical, or *photoresist*, which is sensitive to light. Photoresists are broadly classified as positive or negative. A positive resist has low solubility in developer and becomes soluble by exposure to light. Conversely, negative resists are soluble in developer, and lose their solubility upon exposure to light. Prior to the invention of the wafer stepper, negative resists, based on poly-isopropene with azide additives, were predominately used. Since the late 1970′s, novolak positive resists have been used for most critical applications. In general, the advantages of the novolak positive resist are lack of swelling during the developer stage, and the use of aqueous developers. Negative resists require organic solvent developer and the resist can swell during the developer stage.

Resists are usually optimized for application over specific wavelength ranges. For cost-effective manufacturing it is desirable to have efficient photoresist photochemistry: low exposure dosage increases exposure-tool productivity. For high resolution patterning, resist films must be exposed with a high degree of uniformity from the top of the resist film to the bottom. Light will not penetrate the photoresist properly or evenly if the resist film is too optically absorbent.

1.6 System Resolution

The resolution of an optical projection lithography system is given by a modification of the Rayleigh resolution criterion

$$R = \frac{k_1 \lambda}{NA} \tag{1.1}$$

where R is the resolution, λ and NA are the exposure wavelength and numerical aperture (typically 0.8 or higher) of the lithographic system respectively. NA is defined as $nsin\theta$, where n is the refractive index of the medium above the photoresist, and θ is the largest semi-angle of converging rays subtended at the photoresist. The k_1 factor (replacing the 0.61 of the classical Rayleigh formula for incoherent imaging) is a system-specific proportionality factor which depends on the coherence of illumination, the use of anti-reflection coatings above or below the photoresist, and also on resist parameters, e.g. baking times and diffusion characteristics. In the last 20 years, all three of these quantities have been modified in an effort to reduce the R, or the *critical dimension*. NA has grown from 0.5 to 1.35 (increased due to liquid immersion), λ has been reduced from 365 *nm* to 193 *nm*, and k_1 has been reduced from 0.7 to 0.27. This has led to a dramatic reduction in the critical dimension: R has shrunk from 500 to 45 *nm*.

As the resolution of lithographic systems increases, there is a trade-off with the depth of focus (DOF) of the system. The DOF of an incoherent optical system is given by

$$DOF = \pm 2\lambda F^2 \tag{1.2}$$

where F is the system f-number, $F = 1/2NA$. Rewritten in terms of numerical aperture, we get

$$DOF = \pm \frac{\lambda}{2NA^2} \tag{1.3}$$

Over this range of focus, the peak intensity of a point object focused by a lens remains within 20 % of the peak value for best focus. Typical values for 180 *nm* lithography are $\lambda = 248$ *nm* and $NA = 0.6$, which results in a Rayleigh depth of focus of ± 0.34 *μm*, i.e. the distance between the lens and the wafer needs to be controlled to this amount.

Just as the Rayleigh criterion for resolution was modified for lithographic systems, Equation 1.3 can be modified to give

$$DOF = k_2 \frac{\lambda}{NA^2}, \tag{1.4}$$

where k_2 is a system-specific, normalizing factor.

A DOF greater than 500nm is highly desired, since wafers are not perfectly flat and the photoresist usually has a thickness of several hundred nanometers. As DOF is inversely proportional to the square of the numerical aperture, increasing the NA to improve resolution quickly decreases the depth of focus. In 1979, the state-of-the-art lens had a resolution of 1.25 μm, a ± 0.75 μm depth of focus, a numerical aperture of 0.28, and imaged at the mercury g-line [8]. These figures produced values of 0.8 and 0.13 for k_1 and k_2 respectively. Lithography today routinely operates at values of k_1 between 0.5 and 0.6. Depths of focus requirements have decreased to ± 0.2 μm, mainly due to the use of chemical-mechanical polishing. The combination of focus and exposure allowances is known as the process window. Modern lithographic systems must find a balance between the need for higher resolution production and the difficulties involved with operating a system with a small depth of focus [31].

Further development and performance analysis of lithographic systems incorporating excimer lasers has been carried out by Jain [9], Liu [32], Welford [33], Chen [34], Ito [35], and Brunner [36].

1.7 Vector Effects in High NA Lithography

In early models of lithographic systems, diffraction theory has been treated mathematically with a scalar approach based on the Fresnel-Kirchoff formulation [37]. Scalar theory is valid for a system numerical aperture (NA) of 0.50 or less; NA values higher than this, and certainly > 0.70, require a vectorial approach. Early lithographic systems were modelled using scalar mathematics producing an "aerial image": an image of the photomask projected directly onto the plane of the wafer. The aerial image is a crucially important quantity in lithography for governing how well a developed photoresist structure replicates a mask design. It is also important in optical design as it gives an indication of how the intensity of light is distributed in the image space of a projection lens system. The aerial image is usually generated using scalar diffraction theory. Today, the system NA has grown from 0.5 to as high as 1.35 as a result of liquid immersion lithography. Here, the space between the optical system and the wafer is filled with a transparent liquid of index n, reducing the effective wavelength, or alternatively increasing the numerical aperture. Today's immersion systems use water, which at the operating wavelength of 193 nm, has $n_{193nm} = 1.44$, leading to $NA = 1.35$.

Scalar theory does not properly take into account the effects of oblique directions of propagation of light in high NA systems. As the system pupil diameter approaches the pupil-to-image distance, or the NA is greater than 0.50, propagation angles of the electric field become significant and traditional scalar models require correction. Since scalar methods treat all illumination as having the same polarization, with all polarization vectors perpendicular to the plane of incidence, the significance of propagation angles of the electric field cannot be accounted for. In a vector model the electromagnetic field is treated as a vector quantity. However, scalar models can be corrected for increased NA. In 1992, Cole et. al [38] presented an amended scalar theory, neglecting paraxial approximations. The paraxial approximation assumes that the angles of the light rays to the optic axis are sufficiently small that small angles approximations can be made. By not imposing this approximation, scalar theory can be extended to the limit of its capabilities and present quite a good representation of a higher NA lithographic system, without the need for vectorial diffraction.

In 2003, Adam et. al. [39] presented an adaptation of the Hopkins' Method (see Section 3.1), based on previous work by Mansuripur [40], Yeung et. al. [41] and Flagello [42], to include polarization effects at high NA. The Transmission Cross Coefficient (TCC), a scalar quantity used in the Hopkins formulation to describe light propagation and image formation in the projection printing system, is replaced by a generalized vector TCC, which includes vectorial addition of the electromagnetic fields inside a thin film.

In the interest of simplicity, and recognising the complexity of dealing with partially coherent fields, in this study we will ignore vectorial effects and approximate our theory using scalar diffraction methods.

Chapter 2

Coherence Theory

Coherence can be defined as the degree to which electromagnetic radiation or any other oscillating quantity maintains a near-constant phase relationship, both temporally and spatially. Here we are dealing with optical fields, i.e. electromagnetic fields oscillating in the optical frequency range.

The earliest investigations of coherence appear in the mid to late 1800s by Verdet and Lord Rayleigh in their studies of “coronas” or Fraunhofer rings [43], [44]. In the 1890s, Michelson established the connection between the visibility of interference fringes and intensity distribution which later shaped our understanding of the theory of partial coherence. The work of von Laue, Berek, Lakeman and Groosmuller, Schrödinger, and Wiener, though not solely focused on the area, contributed to early theories of coherence. Van Cittert defined a ‘degree of coherence’ and calculated functions expressing the correlation between the (complex) amplitudes at any two points. Zernike [45] generalized these proofs by employing a simpler definition of partial coherence. The van Cittert-Zernike theorem provided the basis for all subsequent work on spatial coherence theory, and will be discussed further in Section 2.5. The theory of partially coherent fields received its modern formulation through several fundamental papers by Emil Wolf and Leonard Mandel. A summary of these can be found in reference [46].

In order to understand coherence, its causes and its effects, we must first be familiar with stochastic processes. A stochastic process is used to describe changes that are non-deterministic and random in character. In these situations, instead of assigning deterministic values to the optical field for each point in space and time, statistical measurements such as averages, standard deviations and correlations are used to describe the field. First order characteristics of optical fields refer to the statistical behaviour of the field observed in one point in space and time, and can be characterized by the probability density function of the electric field component, providing the probability for the electric field component to be within a certain interval.

In statistical optics, the second order statistical properties of a field are often used to characterise partially coherent fields. By 'second order' we mean that the field is evaluated through correlation between the field in two separate coordinates. By using correlations it is possible to specify the randomness of the field. In the following sections we will investigate correlations of optical fields over time, known as the temporal coherence, and correlations over spatial coordinates, referred to as the spatial coherence. To completely understand these concepts, we look at two fundamental theorems: the Wiener-Khintchine theorem, which relates temporal coherence to the spectrum of the optical field, and the van Cittert-Zernike theorem relating spatial coherence to illumination using incoherent sources.

2.1 The Analytic Signal

The theory of partial coherence, whether based on a classical or a quantum description of the optical field, is generally formulated in terms of space-time correlation functions [37]. The simplest correlation function, the *mutual coherence function*, is the correlation of the space and time-varying electric field, or, more conveniently, the analytic signal (i.e. the positive part of the field). The mathematics of the space-time approach become unwieldy, however, when dealing with the interaction of light with matter. The response of matter to an incident field is most naturally described by frequency-dependent response functions such as the dielectric constant, the refractive index, or the magnetic susceptibility. The dominant effects frequently arise from the statistical features of the medium itself rather than of the incident light, for example, when a laser beam propagates through the atmosphere or when it is scattered from a rough surface (giving rise to speckle). In this case, it is more useful to take a single-frequency description. This space-frequency approach was introduced by Emil Wolf [47] and developed in collaboration with Leonard Mandel [48]. For the purpose of this study, we adopt the original space-time approach [37].

In the scalar theory of light, the electromagnetic field is fully described by a real function of space and time, $V^{(r)}(\mathbf{x},t)$. For any practical light beam, $V^{(r)}(\mathbf{x},t)$, the real disturbance, is a random function of space $\mathbf{x}$ and time t and any one $V^{(r)}$ may be regarded as a typical realization of an ensemble consisting of all realizations. The fluctuation is a result of the atomic processes that give rise to the emission of light and is present even in a laser. Each member of the ensemble is a possible realization of, and shares all the attributes of, the optical field. Figure 2.1 shows a typical realisation (a) as a function of time for a fixed point in space and (b) as a function of space for one point in time. Although the field is purely real, it is mathematically more convenient to describe it by a complex function, $V(\mathbf{x},t)$, called the analytic signal, which can be found from $V^{(r)}(\mathbf{x},t)$ in the following way.

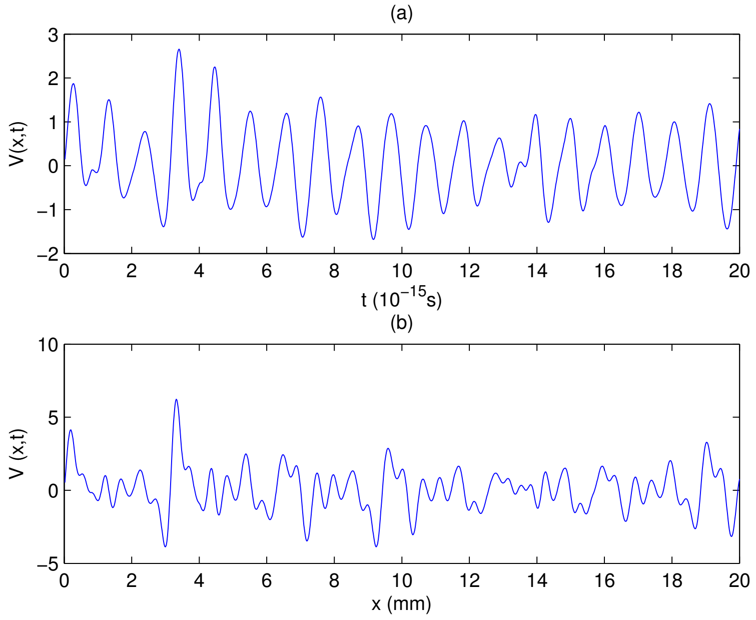

Figure 2.1: The analytic signal: a typical realisation (a) as a function of time for a fixed point in space, and (b) as a function of space for one point in time.

Let $v(\mathbf{x},\nu)$ be the Fourier transform of $V^{(r)}(\mathbf{x},t)$:

$$v(\mathbf{x},\nu) = \int_{-\infty}^{\infty} V^{(r)}(\mathbf{x},t)exp(2\pi i\nu t)dt \tag{2.1}$$

The analytic signal $V(\mathbf{x},t)$ is defined by

$$V(\mathbf{x},t) = \int_0^\infty v(\mathbf{x},\nu) exp(-2\pi i \nu t) d\nu \tag{2.2}$$

$V(\mathbf{x},t)$ contains only positive frequencies and this gives rise to its analytical properties. For our purposes it is sufficient to say that $V(\mathbf{x},t)$ is a complex description of a real disturbance [37].

If the light is narrowband, i.e. it has significant amplitude only for frequencies in the region of the mean frequency $\bar{\nu}$, then the analytic signal can be written as the product of a slowly varying complex function called the time-varying complex amplitude $A(\mathbf{x},t)$ and a rapidly varying term

$$V(\mathbf{x},t) = A(\mathbf{x},t) exp(-2\pi i \bar{\nu} t) \tag{2.3}$$

Therefore, for narrowband fields, the time-varying complex amplitude, which represents the envelope and phase of the oscillatory term, is an adequate descriptor of the field. The signal randomly fluctuates in space and time and so it must be described in statistical terms. The *mutual coherence function* is defined as

$$\Gamma(\mathbf{x_1},\mathbf{x_2};t_1,t_2) = \langle V^*(\mathbf{x_1},t_1) V(\mathbf{x_2},t_2) \rangle \tag{2.4}$$

where $\langle \ldots \rangle$ denotes an average over the ensemble of possible realizations. This function describes the correlation of the light at point $\mathbf{x_1}$ at time t_1 with that at point $\mathbf{x_2}$ at time t_2. It has a relatively large value when the two signals are statistically similar and is zero if they are statistically independent, (assuming $\langle V \rangle = 0$). If $V(\mathbf{x},t)$ is statistically stationary in time (i.e. its probability distributions are unaffected by a change in the time origin), then Γ depends only on the difference $t_2 - t_1 = \tau$, thus

$$\Gamma(\mathbf{x_1},\mathbf{x_2},\tau) = \langle V^*(\mathbf{x_1},t) V(\mathbf{x_2},t+\tau) \rangle \equiv \Gamma_{12}(\tau). \tag{2.5}$$

It is often convenient to use a normalized version of the mutual coherence function. We define the *complex degree of coherence*, $\gamma(\mathbf{x_1},\mathbf{x_2},\tau)$ or simply $\gamma_{12}(\tau)$ as

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1 I_2}} \tag{2.6}$$

This normalized function has the property that $0 \leq |\gamma_{12}(\tau)| \leq 1$. A value of zero indicates an absence of coherence, or *incoherence*, between the field at points 1 and 2, and a value of 1 indicates a completely *coherent* field. Any value of γ between these limits indicates a level of *partial coherence*.

The quantity commonly used to describe the intensity profile and coherence of a source is known as the *cross-spectral density*, $W(\mathbf{x_1},\mathbf{x_2},\nu)$ or simply $W_{12}(\nu)$, which is a correlation of the Fourier components of the analytic signal. The cross-spectral density is an important quantity when imaging with light that is partially spatially coherent as it contains information on the correlation between pairs of points in such sources. This will be explored further in Section 2.6.

2.2 Temporal Coherence

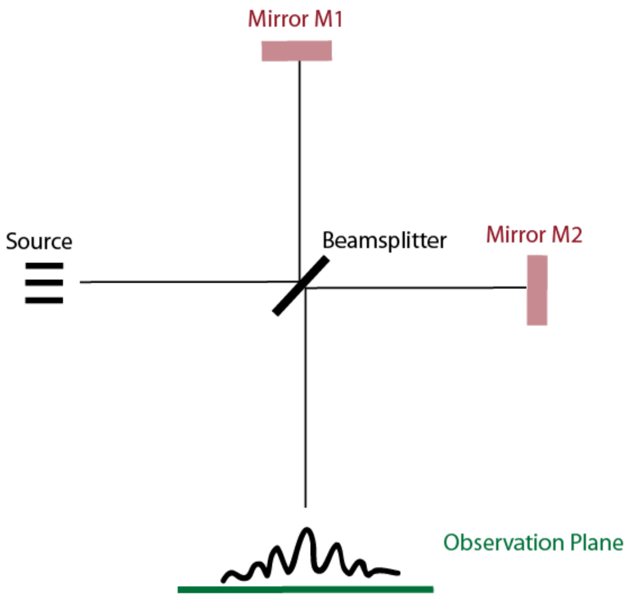

Figure 2.2: The Michelson interferometer: At least one of the mirrors is moveable, introducing a tunable path length difference.

To describe temporal coherence, we consider a beam from a point source σ, divided into two beams in a Michelson interferometer as in Figure 2.2. The two beams are united after a path delay Δs has been introduced between them such that

$$\Delta s = c\Delta t \tag{2.7}$$

where $c =$ velocity of light. If Δs is sufficiently small, interference fringes are formed in the observation plane. These fringes are observed due to the temporal coherence occurring between the two beams, since the fringe contrast depends on the time delay Δt introduced between them. In general, interference fringes will be observed if

$$\Delta t \Delta \nu \leq 1 \tag{2.8}$$

where $\Delta \nu$ is the effective bandwidth of the light. The time delay

$$\Delta t \sim \frac{1}{\Delta \nu} \tag{2.9}$$

is known as the coherence time of the light and the corresponding path $c\Delta t$ is the coherence length, or more precisely, the longitudinal coherence length of the light. Since $v = c/\lambda$, where λ is the wavelength, $\Delta \nu \sim c\Delta\lambda/\bar{\lambda}^2$, the expression for the coherence length may also be written as

$$\Delta l \sim (\frac{\bar{\lambda}}{\Delta \lambda})\bar{\lambda} \tag{2.10}$$

where $\bar{\lambda}$ is the mean wavelength. The coherence time of the light, Δt, is a measure of the time interval in which appreciable amplitude correlations and phase correlations of the light vibrations at a particular point in an optical field will persist.

This phenomenon can be explained by using the concept of correlations. We can consider the source as emitting a succession of slowly modulated wave trains, the mean frequency of which coincides with the mean frequency of the light, and whose duration is of the order of the reciprocal bandwidth of the light, i.e. of the order of the coherence time in Equation 2.9. Each wave train incident on the beamsplitter splits into two wave trains of the same form but with a reduced amplitude. In the observation plane, the wave trains of the two partial beams become superposed after a time delay has been introduced between them. There will be a strong correlation between the fluctuations in the two beams arriving at the observation plane if the time delay introduced is short compared to the coherence time of the light, Δt. There is effectively no correlation between them when the introduced time delay is much greater than the coherence time. Thus we can assume that the formation or absence of interference fringes in the observation plane is a result of the correlation or lack of correlation between the fluctuations of the two partial beams arriving at the observation plane [48].

2.3 The Wiener-Khintchine Theorem

The Wiener-Khintchine theorem expresses the power density of a stationary random process in terms of the autocorrelation function. One of the reasons for its importance is that for random signals the Fourier transform may not exist, and the autocorrelation is directly measurable.

Let

$$V_T^{(r)}(P,t) = \int_{-\infty}^{\infty} v_T(P,\nu) e^{-2\pi i \nu t} d\nu \tag{2.11}$$

be the Fourier integral representation of the truncated real function $V_T^{(r)}$, which is a function of position P and time t. Applying the inverse Fourier transform we get

$$v_T(P,\nu) = \int_{-\infty}^{\infty} V_T^{(r)}(P,t) e^{2\pi i \nu t} dt, \tag{2.12}$$

and it follows that

$$\begin{aligned}
\int_{-\infty}^{\infty} V_T^{(r)}(P_1,t+\tau) V_T^{(r)}(P_2,t) dt \quad &= \\
&= \int_{-\infty}^{\infty} V_T^{(r)}(P_2,t) \left[\int_{-\infty}^{\infty} v_T(P_1,\nu) e^{-2\pi i \nu (t+\tau)} d\nu \right] dt \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} V_T^{(r)}(P_2,t) e^{-2\pi i \nu t} dt \right] v_T(P_1,\nu) e^{-2\pi i \nu \tau} d\nu \\
&= \int_{-\infty}^{\infty} v_T(P_1,\nu) v_t^*(P_2,\nu) e^{-2\pi i \nu \tau} d\nu. \qquad (2.13)
\end{aligned}$$

If we now divide both sides of 2.13 by $2T$, take the average (denoted by a bar) of $v_T(P_1,\nu)v_t^*(P_2,\nu)/2T$ over the ensemble of random functions $V^{(r)}$, and proceed to the limit $T \to \infty$, we find

$$\langle V^{(r)}(P_1,t+\tau) V^{(r)}(P_2,t) \rangle = \int_{-\infty}^{\infty} W_{12}(\nu) e^{-2\pi i \nu \tau} d\nu, \tag{2.14}$$

where

$$W_{12}(\nu) = \lim_{T\to\infty} \left[\frac{\overline{v_T(P_1,\nu) v_T^*(P_2,\nu)}}{2T} \right] \tag{2.15}$$

The function $W_{12}(\nu)$ may be called the *mutual spectral density*, or the *cross-spectral density*, of the light vibrations at P_1 and P_2. Equation 2.14 shows that the real correlation function $\langle V^{(r)}(P_1,t+\tau)V^{(r)}(P,t)\rangle$ and the mutual spectral density $W_{12}(\nu)$ form a Fourier transform pair.

Moving now to the complex representation. Let

$$V_T(P,t) = 2\int_0^{\infty} v_T(\nu)e^{-2\pi i\nu\tau}d\nu \tag{2.16}$$

be the analytic signal associated with $V_T^{(r)}(P,t)$. Following an analysis similar to that above, again letting $T \to \infty$, we arrive at

$$\Gamma_{12}(\tau) = \langle V(P_1,t+\tau)V^*(P_2,t)\rangle = 4\int_0^{\infty} W_{12}(\nu)e^{-2\pi i\nu\tau}d\nu. \tag{2.17}$$

In the case when $P_1 = P_2 = P$, Equations 2.14 and 2.17 imply that

$$\langle V^{(r)}(P,t+\tau)V^{(r)}(P,t)\rangle = \int_{-\infty}^{\infty} S(P,\nu)e^{-2\pi\nu\tau}d\nu, \tag{2.18}$$

and

$$\langle V(P,t+\tau)V^*(P,t)\rangle = 4\int_0^{\infty} S(P,\nu)e^{-2\pi\nu\tau}d\nu, \tag{2.19}$$

where $S(P,\nu)$ is the spectral density at the point P. Equation 2.18 shows that the real correlation function $\langle V^{(r)}(P,t+\tau)V^{(r)}(P,t)\rangle$ and the spectral density $S(P,\nu)$ form a Fourier transform pair. This result is the optical equivalent of the *Wiener-Khintchine Theorem* of the theory of stationary random processes.

In Section 2.6, we will see that the cross-spectral density and its inverse are generalized forms of the Wiener-Khintchine theorem. A more detailed analysis of this relation can be found in [48].

2.4 Spatial Coherence

When considering spatial coherence, it is useful to consider an interference experiment of the Young's double slit type, with quasi-monochromatic light from an extended thermal source. The term quasi-monochromatic describes light that is (a) narrowband, i.e.

$$\Gamma_{12}(\tau) = \langle A^*(\mathbf{x_1},t)A(\mathbf{x_2},t+\tau)\rangle exp(-2\pi i\bar{\nu}\tau) \tag{2.20}$$

and (b) has a long temporal coherence length compared with the path differences τ in the experimental arrangement in Figure 2.2, i.e.,

$$\langle A^*((\mathbf{x_1},t)A(\mathbf{x_2},t+\tau)\rangle = \langle A^*((\mathbf{x_1},t)A(\mathbf{x_2},t)\rangle = J(\mathbf{x_1},\mathbf{x_2}) \tag{2.21}$$

where $J(\mathbf{x_1},\mathbf{x_2})$ or simply J_{12} is called the *mutual intensity*. This is a special case, $\tau = 0$, of the mutual coherence function, $J_{12} = \Gamma_{12}(0)$. Note that light that is considered quasi-monochromatic in this experiment may not be in another arrangement involving larger path differences.

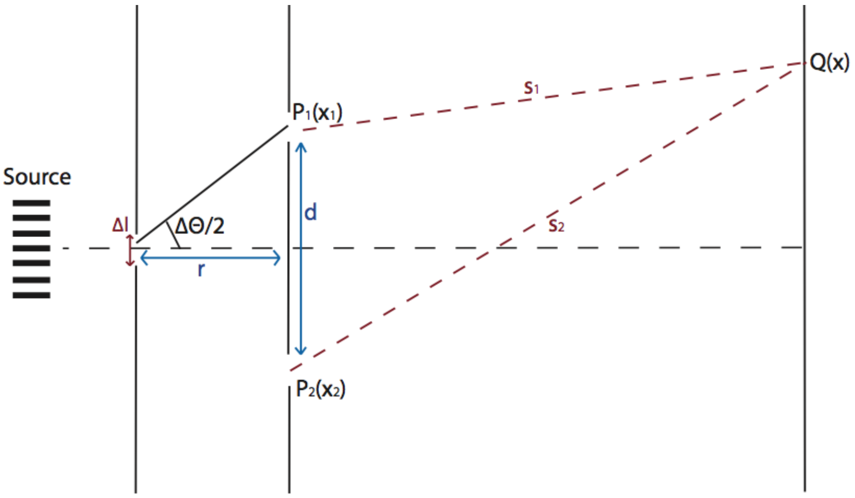

Figure 2.3: Spatial coherence illustrated by Young's double-slit interference experiment

Consider the set-up in Figure 2.3. If the pinholes P_1 and P_2 are sufficiently close to each other, interference fringes will be observed at the point $Q(x)$ on the screen. These fringes are a manifestation of spatial coherence between the light reaching $Q(x)$ from the two pinholes, since the fringe contrast depends on the spatial separation of the pinholes. Interference will be observed near $Q(x)$ if

$$\Delta\theta\Delta l \leq \bar{\lambda} \tag{2.22}$$

where $\Delta\theta$ is the angle that the distance P_1P_2 of separation between the pinholes subtends at the source and

$$\bar{\lambda} = \frac{c}{\bar{\nu}} \tag{2.23}$$

is the effective or mean wavelength of the light. In order to observe interference fringes at or near $Q(x)$, the pinholes must lie within a certain region known as the *coherence area*, ΔA:

$$\Delta A \approx (r\Delta\Theta)^2 \approx \frac{r^2\bar{\lambda}^2}{(\Delta l)^2} = \frac{c^2 r^2}{\bar{\nu}^2 S}, \tag{2.24}$$

where r denotes the distance between the plane of the source and the plane containing the pinholes (see Figure 2.3), and $S = \Delta l^2$ is the area of the source. The solid angle $\Delta\Omega$ which the area subtends at the source is given by

$$\Delta\Omega \approx \frac{\Delta A}{r^2} \approx \frac{c^2}{\bar{\nu}^2 S} \tag{2.25}$$

The intensity $I(\mathbf{x})$ at the point $Q(x)$ is

$$I(\mathbf{x}) = \langle |K_1 V(\mathbf{x}_1, t - \frac{s_1}{c}) + K_2 V(\mathbf{x}_2, t - \frac{s_2}{c})|^2 \rangle \tag{2.26}$$

where K_1 and K_2 are imaginary propagation constants and the medium between P and Q is non-dispersive. This expression can be simplified to give

$$I(\mathbf{x}) = I_1(\mathbf{x}) + I_2(\mathbf{x}) + 2|K_1 K_2^*| \; Re\{J_{12} \, exp(-2\pi i \bar{\nu} \, \frac{s_2 - s_1}{c})\} \tag{2.27}$$

where I_1 and I_2 are the intensities at $\mathbf{x}$ due to slit P_1 only and slit P_2 only, respectively. A normalized form of J_{12} is defined as

$$\mu'_{12} = \frac{J_{12}}{\sqrt{J_{11} J_{22}}} = \gamma_{12}(0) \tag{2.28}$$

where μ'_{12} is called the *complex coherence factor*. This is essentially the same as the complex degree of coherence, $\gamma_{12}(\tau)$. Using the complex coherence factor, we can rewrite Equation 2.27, the intensity in the observation plane, as

$$I(\mathbf{x}) = I_1(\mathbf{x}) + I_2(\mathbf{x}) + 2\sqrt{I_1(\mathbf{x}) I_2(\mathbf{x})} \; |\mu'_{12}| \; Cos\,[2\pi\bar{\nu}(\frac{s_2 - s_1}{c}) + \beta_{12}] \tag{2.29}$$

where $\beta_{12} = arg\ \mu'_{12}$. The fringe visibility V is defined as

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{2.30}$$

and can also be written as

$$V = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}\,|\mu'_{12}| \tag{2.31}$$

From this we see that the fringe visibility equals the complex coherence factor if $I_1 = I_2$. The quantity μ'_{12} describes the degree of spatial coherence of the light at points P_1 and P_2, separated by a distance d. If $|\mu'_{12}| = 1$, then the light is spatially coherent and the fringes have unit visibility; if $|\mu'_{12}| = 0$, the light is said to be spatially incoherent and no fringes at all are seen. It follows that Young's experiment can be used to measure the spatial coherence of a light source. Methods based on this principle will be discussed further in Section 4.1.

2.5 The van Cittert-Zernike Theorem

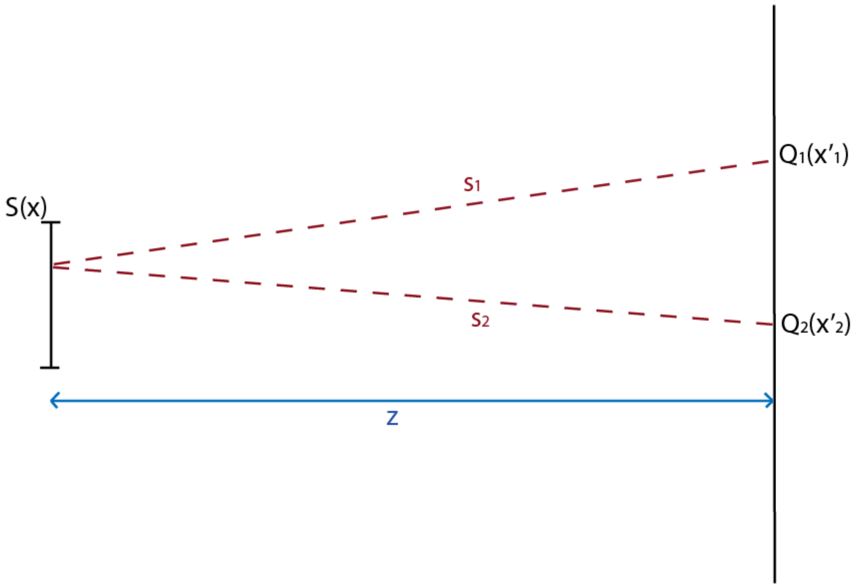

Figure 2.4: The van Cittert-Zernike Theorem

The van Cittert-Zernike theorem can be viewed as the spatial equivalent of the Wiener-Khintchine theorem. It relates the field correlations at two points in the field to the properties of a spatially incoherent, quasi-monochromatic, planar source. The most common use of the theorem is to estimate the spatial coherence from an incoherent source of a specific dimension.

If we know the mutual intensity J_{12} for all points on some surface Σ_1, then we can find the mutual intensity for all pairs of points on another surface separated from the first by free space using the Huygens-Fresnel formula.

$$J(\mathbf{x}_1', \mathbf{x}_2') = \int_{\Sigma_1} \int_{\Sigma_1} J(\mathbf{x}_1, \mathbf{x}_2)\ exp[i\bar{k}(s_2 - s_1)] \frac{\Lambda_1^* \Lambda_2}{s_1 s_2} d\mathbf{x}_1 d\mathbf{x}_2 \tag{2.32}$$

where Λ_1 and Λ_2 denote the obliquity factors, which are proportional to the amplitudes of secondary waves propagating in various directions according to Huygens' principle. Consider the layout in Figure 2.4 in which an incoherent primary source illuminates a screen. We wish to find the mutual intensity and complex coherence factor at the screen. Once again, the light is assumed to be quasi-monochromatic. The source is described by a mutual intensity

$$J(\mathbf{x}_1', \mathbf{x}_2') = \delta(\mathbf{x}_1 - \mathbf{x}_2) S(\mathbf{x}_1) \tag{2.33}$$

where δ is the Dirac delta function and $S(\mathbf{x}_1)$ is the intensity of the source. Substituting Equation 2.33 into 2.32 gives a general statement of the van Cittert-Zernike theorem:

$$J(\mathbf{x}_1', \mathbf{x}_2') = \int_{\Sigma_1} S(\mathbf{x}_1)\ exp[i\bar{k}(s_2 - s_1)] \frac{\Lambda_1^* \Lambda_2}{s_1 s_2} d\mathbf{x}_1 \tag{2.34}$$

with distances s_1 and s_2 representing the distance from source to observation plane, as in Figure 2.4. The complex coherence factor μ_{12}' is simply a normalized version of this:

$$\mu_{12}' = \frac{J(\mathbf{x}_1', \mathbf{x}_2')}{\sqrt{I(\mathbf{x}_1') I(\mathbf{x}_2')}} \tag{2.35}$$

If the observation plane lies in the far field of the source, μ_{12}' depends only on the coordinate differences $\mathbf{x}' = \mathbf{x}_1' - \mathbf{x}_2'$ and is essentially the normalized Fourier transform of the source intensity

$$\mu_{12}'(\mathbf{x}') = \frac{e^{i\psi}\ \int_{\Sigma} S(\mathbf{x})\ exp(\frac{-i\bar{k}}{z} \cdot \mathbf{x} \cdot \mathbf{x}') d\mathbf{x}}{\int_{\Sigma} S(\mathbf{x}) d\mathbf{x}} \tag{2.36}$$

where ψ is a phase term.

According to the van Cittert-Zernike theorem, the relation between the coherence function and the spatial frequency power spectrum is a Fourier transform. A more detailed proof of the van Cittert-Zernike theorem can be found in [37] and [49].

2.6 The Cross-Spectral Density

The cross-correlation of two real random processes $x_1(t)$ and $x_2(t)$ is defined by the average product $\langle x_1(t_1)x_2(t_2)\rangle$ at two different times t_1, t_2, similar to the definition of autocorrelation. We can define the cross-correlation of two complex processes as

$$\Gamma_{12}(t, t+\tau) = \langle z_1^*(t) z_2(t+\tau)\rangle \tag{2.37}$$

If $z_1(t)$ and $z_2(t)$ are jointly stationary, then $\Gamma_{12}(t, t+\tau)$, like $\Gamma(t, t+\tau)$, is a function of τ only, and may be written as $\Gamma_{12}(\tau)$. If we then let $z_1(t), z_2(t), \ldots, z_N(t)$ be a set of N different, jointly stationary random processes, then we can say

$$\Gamma_{ij}(\tau) \equiv \langle z_i^*(t) z_j(t+\tau)\rangle, \quad (i, j = 1, 2, \ldots, N) \tag{2.38}$$

is an $N \times N$ matrix known as the *cross-correlation matrix*. We can define the cross-spectral density (or cross-power spectrum), $W_{ij}(\omega)$, of the jointly stationary random processes $z_i(t)$ and $z_j(t)$ by the formula

$$\langle \tilde{z}_i^*(\omega)\tilde{z}_j(\omega')\rangle = W_{ij}(\omega)\delta(\omega - \omega'), \tag{2.39}$$

where $\tilde{z}_i(\omega)$ is the Fourier transform of $z_i(t)$. This expression for the cross-spectral density shows that Fourier components that belong to different frequencies are unrelated. Thus, the cross-spectral density is a measure of the correlations between the fluctuations of different components at the same frequency. The cross-spectral density can be rewritten as

$$W_{ij}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Gamma_{ij}(\tau) e^{i\omega\tau} d\tau. \tag{2.40}$$

This expression is a Fourier transform of the cross-correlation of $z_i(t)$ and $z_j(t)$, also known as the mutual coherence function. This is analogous to the previously defined function $W_{12}(\nu)$ (see Equation 2.14), and thus, with its inverse, is known as a generalized form of the Wiener-Khintchine theorem.

We can extend this theory into the context of optical coherence. The analytic signal $V(\mathbf{r},t)$, assumed to be stationary and ergodic, can be represented as a Fourier integral

$$V(\mathbf{r},t)=\int_0^\infty \widetilde{V}(\mathbf{r},\nu)e^{-2\pi i\nu t}d\nu \tag{2.41}$$

Using Equation 2.39, we can express the cross-spectral density as

$$\langle \widetilde{V}^*(\mathbf{r}_1,\nu)\widetilde{V}(\mathbf{r}_2,\nu')\rangle = W(\mathbf{r}_1,\mathbf{r}_2,\nu)\delta(\nu-\nu'), \tag{2.42}$$

where the ensemble average is taken over the different realisations of the field and δ is the Dirac delta function. As before, we can also normalise the cross-spectral density function

$$\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)=\frac{W(\mathbf{r}_1,\mathbf{r}_2,\nu)}{[W(\mathbf{r}_1,\mathbf{r}_1,\nu)]^{1/2}[W(\mathbf{r}_2,\mathbf{r}_2,\nu)]^{1/2}} \tag{2.43}$$

$$=\frac{W(\mathbf{r}_1,\mathbf{r}_2,\nu)}{[S(\mathbf{r}_1,\nu)]^{1/2}[S(\mathbf{r}_2,\nu)]^{1/2}} \tag{2.44}$$

where $S(\mathbf{r},\nu)\geq 0$ is the non-negative spectral density (power spectrum) of the light. $\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)$ is referred to as the *spectral degree of coherence at frequency* ν (or the complex degree of spatial coherence) [48], [50]. This is not to be confused with the previously defined *complex degree of coherence* $\gamma_{12}(\tau)$ (Equation 2.6); $\gamma_{12}(\tau)$ may be said to characterise field correlations in the space-time domain, whereas $\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)$ characterises them in the space-frequency domain. Thus $\gamma_{12}(\tau)$ and $\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)$ are measures of different characteristics of the field. They are related, however, as shown by Friberg and Wolf [51]:

$$\gamma_{12}(\tau)=\int_0^\infty \sqrt{s_1(\omega)}\sqrt{s_2(\omega)}\mu_{12}(\omega)exp(-i\omega\tau)d\omega, \tag{2.45}$$

where $s_j(\omega),(j=1,2)$ are the normalised spectra of the field at the two points.

2.6.1 Cross-Spectral Density and the Schell-Model Beam

The cross-spectral density is a key quantity to describe partially coherent light. The Schell-model beam was first postulated by Schell in 1967 [52]. He separated the mutual coherence function into the product of functions "s_1,s_2, and s_1-s_2", i.e.,

$$\Gamma(s_1,s_2,0)=\langle V(s_1)V^*(s_2)\rangle = g(s_1)g^*(s_2)\gamma(s_1-s_2) \tag{2.46}$$

where $g(s)$ is a description of the rms average amplitude across the aperture and $\gamma(s_1 - s_2)$ is the normalised correlation. His area of application was in radio astronomy and the detection of radio signals, but the theory is applicable to quasi-monochromatic signals that are stationary in time and space.

The concept of Schell-model sources has been widely used to describe the structure of spatially partially coherent sources and the far zone representation of the respective optical field [53–57]. In particular, Gaussian Schell-model (GSM) sources characterized by Gaussian distributions of both the optical intensity and the complex degree of spatial coherence have been extensively analysed [58–62]. This approach is popular mainly because the mode sources of this type represent many actual sources to a good approximation and are mathematically convenient to work with.

A source is said to be Schell type if its cross-spectral density has the form

$$W(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{I(\mathbf{r}_1) I(\mathbf{r}_2)} \gamma(\mathbf{r}_1 - \mathbf{r}_2) \tag{2.47}$$

If its cross-spectral density and its intensity envelope have a Gaussian profile, the source is said to be a Gaussian Schell-model (GSM). The quasi-monochromatic GSM beam is one of a very limited number of partially coherent fields that can be propagated analytically [48]. The fact that they can be propagated analytically make them very useful in verifying numerical propagation algorithms. The essential point is that the degree of spectral coherence depends on $\mathbf{r}_1$ and $\mathbf{r}_2$ only through their difference. The Schell-model beam will be explored further in Chapters 3 and 4.

Chapter 3

Propagation of Partially Coherent Optical Fields

The theory of propagation of partially coherent light is well established, basically taking the form of diffraction integrals that are extensions of those used for coherent or incoherent light [37], [48], [63]. When it comes to performing these integrations numerically, a problem arises: the integrals are four-dimensional (4D) rather than two-dimensional (2D) as for purely coherent or incoherent light. The reason is that for partially coherent light, the second-order correlation between fields at different points in space must be taken into account, so integration must be performed not just over all points of a 2D field distribution but over all pairs of points. Methods of avoiding these extended calculations have been developed, all with their distinct advantages and disadvantages. Some of these, such as Hopkins' Method and coherent-mode decomposition are discussed in Sections 3.1 and 3.2 respectively. In Section 3.4, the Elementary Function Method, previously presented by Wald et al [64] and Burvall et al [65], is summarized and discussed.

3.1 Hopkins' Method

In general, Fourier optics can be used to represent imaging systems by means of convolution [66]. If we wish to find the image intensity, given the amplitude of the object and the amplitude point spread function (psf), for an object illuminated by *coherent* light, we apply a simple convolution formula:

$$I_i = |h \otimes U_g|^2 \tag{3.1}$$

where h represents the amplitude point spread function and U_g is the object amplitude. A coherent imaging system is linear in complex amplitude (assumed to be the same everywhere in the field, i.e. isoplanatic), as shown in the amplitude convolution equation above. A non-coherent, or *incoherent*, imaging system is linear in intensity. For this case, the corresponding convolution equation involves the intensity impulse response and the ideal image intensity:

$$I_i = |h|^2 \otimes |U_g|^2 \tag{3.2}$$

However, if the object illumination is partially spatially coherent then, in general, no convolution formula, and therefore no linear transfer function exists.

Consider an optical system (see Figure 3.1) in which an "incoherent" source (the condenser exit pupil) illuminates an object which is imaged using the objective lens. The condenser exit pupil may not act as an incoherent source if the original primary source is extremely small, but under practical conditions it is a sufficiently good approximation. Given this optical system, our aim is to find the image intensity in terms of the amplitude transmittance of the object and other functions. Our basic method is to propagate the two coordinates of the mutual intensity through the system and then find the image intensity by letting the two coordinates of the mutual intensity in the image plane be equal; this basic method is used in solving virtually all spatially partially coherent imaging problems. We assume the light is quasi-monochromatic.

Let $S(f,g)$ be the source intensity distribution in the exit pupil of the condenser; the coordinates (f,g) are related to the real distance coordinates (ξ,ζ) in the exit pupil by

$$f = \frac{\xi}{\lambda f_c}, \quad g = \frac{\zeta}{\lambda f_c} \tag{3.3}$$

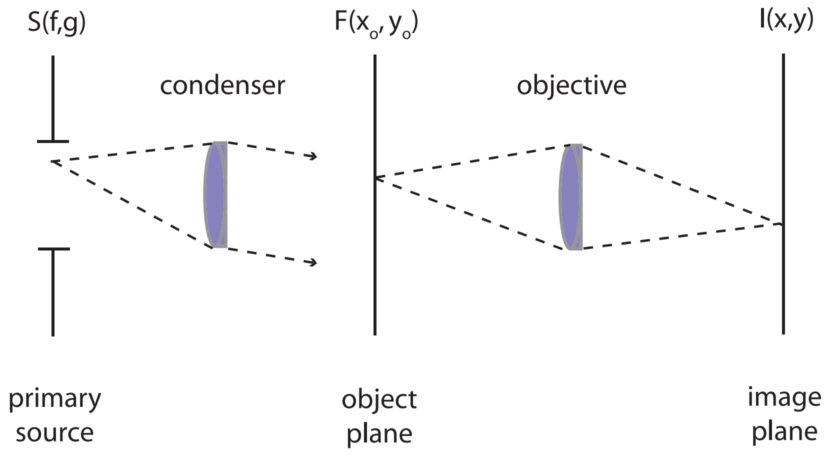

Figure 3.1: Optical system in a microscope.

where f_c is the focal length of the condenser. Using the van Cittert-Zernike theorem, the mutual intensity just before the object is therefore given by the Fourier transform of $S(f,g)$:

$$J_0^-(x_0 - x_0', y_0 - y_0') = \iint_{-\infty}^{\infty} S(f,g) exp[+2\pi i f(x_0 - x_0') + g(y_0 - y_0')] df dg \tag{3.4}$$

The mutual intensity just after the object is

$$J_0(x_0, y_0, x_0', y_0') = J_0^-(x_0 - x_0', y_0 - y_0') F(x_0, y_0) F^*(x_0', y_0') \tag{3.5}$$

where $F(x_0, y_0)$ is the complex amplitude transmittance (TCC) of the object. Assuming an isoplanatic optical system, the mutual intensity in the image plane is

$$\begin{aligned} J_I(x_1, y_1, x_1', y_1') \quad = \quad & \iiiint_{-\infty}^{\infty} J_0^-(x_0 - x_0', y_0 - y_0') \\ & \times F(x_0, y_0) F^*(x_0', y_0') \\ & \times K(x_1 - x_0, y_1 - y_0) K^*(x_1' - x_0', y_1' - y_0') dx_0 dy_0 dx_0' dy_0' \end{aligned} \tag{3.6}$$

where $K(x,y)$ is the amplitude point spread function of the objective lens. The image intensity is found from Equation 3.6 by letting $x_1 = x_1'$ and $y_1 = y_1'$. We avoid this lengthy expression and rewrite the functions in terms of their Fourier transforms:

$$\tilde{F}(f,g) = \iint_{-\infty}^{\infty} F(x,y) exp[+2\pi i(fx+gy)]dxdy \tag{3.7}$$

and

$$\tilde{K}(f,g) = \iint_{-\infty}^{\infty} K(x,y) exp[+2\pi i(fx+gy)]dxdy, \tag{3.8}$$

where $\tilde{F}(f,g)$ is the Fourier transform of the complex amplitude transmittance, $F(x,y)$. The function $\tilde{K}(f,g)$ is the coherent transfer function and is equal to the pupil transmission function of the objective lens if distances in the pupil (ξ',ζ') are related to (f,g) by

$$f = \frac{\xi'}{\lambda f_o}, \quad g = \frac{\zeta'}{\lambda f_o}, \tag{3.9}$$

where f_o is the focal length of the objective. Using Equations 3.7 and 3.8 and rearranging, the expression for the image intensity becomes

$$\begin{aligned} I(x_1,y_1) &= \iiiint_{-\infty}^{\infty} T(f',g',f'',g'')\tilde{F}(f',g')\tilde{F}^*(f'',g'') \\ &= \times exp[-2\pi i\{(f'-f'')x_1 + (g'-g'')y_1\}]df'dg'df''dg'' \end{aligned} \tag{3.10}$$

where $T(f',g',f'',g'')$ is called the *transmission cross coefficient* and is defined by

$$T(f',g',f'',g'') == \iint_{-\infty}^{\infty} S(f,g)\tilde{K}(f+f',g+g')\tilde{K}^*(f+f'',g+g'')dfdg \tag{3.11}$$

Finally, the Fourier transform of the image intensity is given by the relatively simple expression

$$\tilde{I}(f,g) = \iint_{-\infty}^{\infty} F(f'+f,g'+g,f',g')\tilde{F}(f'+f,g'+g)\tilde{F}^*(f',g')df'dg' \tag{3.12}$$

Hopkins was the original pioneer for the use of transfer function methods for the assessment of the quality of optical imaging systems [67], [68]. The previous work-through is a generalised method, and Hopkins' method follows, making the right substitutions. Hopkins introduced a "phase coherence factor" [67] which details the correlation of phase between wave disturbances at any two points.

The phase-coherence factor in the object plane may be built up step by step: from the source to the entrance pupil of the first condenser, to the exit pupil and so on. For two points P_1, P_2 of an illuminated surface the phase-coherence factor is defined as

$$\gamma_{12} = \frac{1}{\sqrt{I_1 I_2}} \int_{\Sigma} U_1 U_2^* d\sigma \tag{3.13}$$

where I_1, I_2 are the intensities at P_1, P_2 produced by the source Σ, and U_1, U_2 are the complex amplitudes at these points associated with an element $d\sigma$ of the source. The factor γ_{12} is in general complex; its modulus defines the the coherence of the vibrations and its argument denotes their difference in phase. It is analogous to the complex degree of coherence defined in Chapter 2 (see Equation 2.6). Hopkins postulated an 'effective source', found by calculating the Fourier transform of the phase-coherence factor,

$$\Gamma(x,y) = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} \gamma(u,v) e^{-i(ux+vy)} du dv \tag{3.14}$$

which can replace the illuminating system, provided $\Gamma(x,y)$ is real. This quantity is analogous to the mutual coherence function defined in Chapter 2. To specify the conditions of illumination of any surface, both the phase-coherence between all pairs of points on the surface and the intensity of the illumination at all points of the surface must be specified. We also assume the original illumination is uniform in intensity; even if this is not the case, the non-uniformity can be incorporated into the effective source as a non-uniform transparency. The complex transmission of the optical system is given as

$$F(u'-u_1, v'-v_1) = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} f(x,y) e^{i\{(u'-u_1)x + (v'-v_1)y\}} dx dy \tag{3.15}$$

where $f(x,y)$ is the Fourier transform of $F(u,v)$.

Let a complex function $E(u,v)$ represent the complex transmission of the object. If an element $d\sigma$ of the source produces a complex amplitude U_1 at (u_1, v_1), the amplitude in the image plane due to light from $d\sigma$ is given by

$$\iint_{-\infty}^{+\infty} U_1 E(u_1, v_1) F(u'-u_1, v'-v_1) du_1 dv_1 \tag{3.16}$$

An equivalent expression can be obtained using an independent current point (u_2, v_2). The partial intensity in the image which is associated with light from $d\sigma$ is then given by

$$\begin{aligned} dI'(u',v') &= d\sigma \iiiint_{-\infty}^{+\infty} U_1 E(u_1,v_1) F(u'-u_1,v'-v_1) U_2^* E^*(u_2,v_2) \\ &\quad \times F^*(u'-u_2,v'-v_2) du_1 du_2 dv_1 dv_2 \end{aligned} \tag{3.17}$$

Integrating over the whole source, the total intensity in the image plane is given as

$$\begin{aligned} I'(u',v') &= \iiiint_{-\infty}^{+\infty} \Gamma(u_1-u_2,v_1-v_2) E(u_1,v_1) \\ &\quad \times F(u'-u_1,v'-v_1) E^*(u_2,v_2) \\ &\quad \times F^*(u'-u_2,v'-v_2) du_1 du_2 dv_1 dv_2 \end{aligned} \tag{3.18}$$

where $\Gamma(u_1,u_2,v_1,v_2)$ is the mutual coherence function, $F(u',u_1,u_2,v',v_1,v_2)$ is the object amplitude and $E(u_1,u_2,v_1,v_2)$ represents the system response. The intensity in the image plane is computed for a single point (x,y) of an 'effective source'. Integration over the area (A) of the effective source gives the intensity distribution due to the whole source.

This approach reduces the complexity of the calculations involved in modelling systems where the illuminating source is partially spatially coherent. However, Hopkins' Method applies only in certain situations: systems containing a perfectly incoherent primary source.

3.2 Coherent-Mode Decomposition

When dealing with the modelling and propagation of partially coherent light fields, the coherent-mode decomposition approach is well established. Although the mutual coherence function (and the associated cross-spectral density) is sufficient for analysis of interference and diffraction effects of light at various degrees of coherence, the calculations quickly become complicated and expensive. The coherent mode method was introduced as an alternative to direct propagation of the cross-spectral density, reducing the complexity of the calculations as only two-dimensional integrals are required. The method was first proposed by Gori [69], followed by some key publications by Wolf [70], [47] and Starikov and Wolf [58]. The basic feature of the theory is a rigorous decomposition of the cross-spectral density function of the source into coherent modes. The modes are propagated and added coherently in the image plane. The expansion uses a basis of orthogonal functions which are eigenfunctions of the mutual coherence function $\Gamma(\mathbf{r}_1, \mathbf{r}_2)$. We will proceed with the approach of Wolf, including notation from [48].

Consider a stationary optical field $V(\mathbf{r}, t)$ in some finite closed domain D in free space and let $\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ be its mutual coherence function. We assume $\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ falls off sufficiently rapidly with τ as $|\tau| \to \infty$ to ensure that, for all points $\mathbf{r}_1 \in D$ and $\mathbf{r}_2 \in D$, the mutual coherence function is absolutely integrable with respect to τ, i.e. that

$$\int_{-\infty}^{\infty} |\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| d\tau < \infty \tag{3.19}$$

The Fourier transform of $\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ is

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) e^{i\nu\tau} d\tau \tag{3.20}$$

$W(\mathbf{r}_1, \mathbf{r}_2, \nu)$ is the cross-spectral density of the source distribution, and its "diagonal" element

$$S(\mathbf{r}, \nu) \equiv W(\mathbf{r}, \mathbf{r}, \nu) \tag{3.21}$$

is the spectral density (the spectrum) of the source distribution at $\mathbf{r}$. It is a property of cross-correlation functions that their absolute integrability implies also square integrability, i.e.,

$$\int_{-\infty}^{\infty} |\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)|^2 d\tau < \infty \tag{3.22}$$

According to Mercer's theorem, any continuous, hermitian, nonnegative definite Hilbert-Schmidt kernel that is not identically zero, and hence our cross-spectral density function $W(\mathbf{r}_1, \mathbf{r}_2, \nu)$, may be expressed in the form

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \sum_n \lambda_n(\nu) \psi_n^*(\mathbf{r}_1, \nu) \psi_n(\mathbf{r}_2, \nu). \tag{3.23}$$

The functions $\psi_n(\mathbf{r}, \nu)$ are the eigenfunctions, and the coefficients $\lambda_n(\nu)$ are the eigenvalues of the integral equation

$$\int_D W(\mathbf{r}_1, \mathbf{r}_2, \nu) \psi_n(\mathbf{r}_1, \nu) d^3 r_1 = \lambda_n(\nu) \psi_n(\mathbf{r}_2, \nu) \tag{3.24}$$

which is seen to be a homogeneous Fredholm integral equation of the second kind.

We can choose that the eigenfunctions form an orthonormal set, i.e. that

$$\int_D \psi_n^*(\mathbf{r},\nu)\psi_m(\mathbf{r},\nu)d^3r = \delta_{nm} \tag{3.25}$$

where δ_{nm} is the Kronecker symbol.

Rewriting the Mercer expansion (Equation 3.23):

$$W(\mathbf{r}_1,\mathbf{r}_2,\nu) = \sum_n \lambda_n(\nu)W^{(n)}(\mathbf{r}_1,\mathbf{r}_2,\nu) \tag{3.26}$$

where

$$W^{(n)}(\mathbf{r}_1,\mathbf{r}_2,\nu) = \psi_n^*(\mathbf{r}_1,\nu)\psi_n(\mathbf{r}_2,\nu) \tag{3.27}$$

An expression of this form represents the cross-spectral density of a completely coherent field. The expansion in Equation 3.26 represents the cross-spectral density of the field as a superposition of elementary modes, each of which is spatially completely coherent at each frequency. For this reason, the expansion 3.23 is sometimes referred to as the *coherent-mode representation* of the cross-spectral density. Starikov and Wolf [58] applied this theory to determine the mode structure of a Gaussian Schell-model source. The number of modes required to represent the source depends on β, the ratio of the coherence length of the source to the size of the source. In the spatially coherent case, $\beta >> 1$, and the source can be fully represented by a single mode. When $\beta << 1$, the source is essentially spatially incoherent, and the number of modes to be propagated is of the order of $1/\beta$, and the coherent mode approach becomes unsuitable. Thus, for a relatively coherent source, the coherent mode approach is useful.

Vahimaa and Turunen [71] extended the coherent mode theory based on a new type of coherent mode expansion involving an incoherent superposition of a set of identical, but spatially displaced coherent elementary fields. Their approach is essentially an extension of the decomposition presented by Gori and Palma [72] but for a special class of quasihomogenous Gaussian fields of low coherence. As with Wolf's approach, the modes are identical, and so a two-dimensional integral is sufficient to propagate the functions. They observe that the cross-spectral density function of any field with a Schell-model angular correlation function can be represented in the form of a continuous, incoherent linear superposition of fully coherent elementary fields, each with the same wave form $f(\rho,z_0)$ and cross-spectral density function

$$W_e(\rho_1,\rho_2,z_0) = f^*(\rho_1)f(\rho_2). \tag{3.28}$$

The elementary fields originate from different lateral locations, and are weighted according to their position across the source. In [71], Vahimaa and Turunen state: *'The free-space propagation of every partially coherent field with a Schell-model angular correlation function can be governed by propagating a single coherent field and then linearly combining the contributions from the entire source area'*.

The coherent-mode decomposition method has obvious advantages over expensive 4D integral calculations. When the degree of coherence is high, the number of coherent modes is small, and the calculations are relatively fast. The main disadvantage is that the modes are particular to a specific field, so if the intensity or coherence distribution changes, the modes must be recalculated.

3.3 The Wigner Distribution Function

The Wigner distribution function is another useful tool in yielding a simpler description of partially coherent light and was first presented by Wigner in 1932 [73] in the description of a mechanical phenomena in phase space. In general, the Wigner function describes a signal in space and spatial frequency simultaneously, and can be considered as the local frequency spectrum of the signal (as opposed to a global energy distribution). The function was extended into optics by Walther in 1968 [74] as the *generalized radiance*, to relate partial coherence with radiometry. The distribution is now well established in the area of Fourier optics, with key contributions from Bastiaans.

Bastiaans [75] showed that the Wigner distribution function concept yields a link between geometrical optics and wave optics, and can also be applied to stochastic signals. The notion of a local frequency spectrum is analogous to geometrical optics, where the signal is described by giving the directions of the rays that should be present at a certain position. In the literature, the Wigner distribution function $f(x,u)$ of a signal $\psi(x)$ is defined by

$$F(x,u) = \int \psi(x+\frac{1}{2}x')\psi^*(x-\frac{1}{2}x')exp[-iux']dx' \tag{3.29}$$

where $\psi^*(x)$ denotes the complex conjugate. There exists an equivalent definition in the frequency domain

$$F(x,u) = \frac{1}{2\pi} \int \hat{\psi}(u + \frac{1}{2}u')\hat{\psi}^*(u - \frac{1}{2}u')exp[iu'x]du' \tag{3.30}$$

The function $F(x,u)$ represents the signal in space and frequency simultaneously, and can thus be considered an intermediate signal representation between the pure space description $\psi(x)$ and the pure frequency description $\hat{\psi}(u)$.

The Wigner distribution function can be used to describe partially coherent light and to propagate it through imaging systems. In Section 2.1, we introduced the mutual coherence function $\Gamma(\mathbf{x}_1,\mathbf{x}_2,\tau)$ (see Equation 2.5) as a statistical ensemble of the correlations in the light source. The Fourier transform of this quantity is the cross-spectral density (Equation 2.40). We can rewrite it here in terms of x_1 and x_2.

$$W(x_1,x_2,\omega) = \int \Gamma(x_1,x_2,\tau)e^{i\omega\tau}d\tau \tag{3.31}$$

The cross-spectral density can also be referred to as the mutual power spectrum. The basic property of this power spectrum is that it is a nonnegative definite Hermitian function of x_1 and x_2. This means that

$$W(-(x_1,x_2,\omega)) = W^*(x_1,x_2,\omega) \tag{3.32}$$

and

$$\int\int g(x_1,\omega)W(x_1,x_2,\omega)g^*(x_2,\omega)dx_1dx_2 \geq 0 \tag{3.33}$$

Equation 3.31 expresses the coherence of the light at two different positions, and its Fourier transform $\widehat{W}(u_1,u_2,\omega)$ expresses the coherence in two different directions.

$$\widehat{W}(u_1,u_2,\omega) = \int\int W(x_1,x_2,\omega)e^{-i(u_1x_1-u_2x_2)}dx_1dx_2. \tag{3.34}$$

If we refer to $W(x_1,x_2,\omega)$ as the *positional* power spectrum, then $\widehat{W}(u_1,u_2,\omega)$ can be called the *directional* power spectrum [76], [77]. The Wigner distribution function offers an alternative to the cross-spectral density by describing a stochastic process in space and spatial frequency simultaneously. For a Gaussian Schell-model beam, the positional power spectrum can be written as

$$W(x_1,x_2) = \frac{\sqrt{2\sigma}}{\rho} exp\left[-\left(\frac{\pi}{2\rho^2}\right)\left(\sigma(x_1+x_2)^2 + \frac{1}{\sigma}(x_1-x_2)^2\right)\right] \quad (0 < \sigma \leq 1); \tag{3.35}$$

where σ represents the degree of coherence of the Schell-model source. The corresponding Wigner distribution function is

$$F(x,u) = 2\sigma exp\left[-\sigma\left(\frac{2\pi}{\rho^2}x^2 + \frac{\rho^2}{2\pi}u^2\right)\right] \quad (0 < \sigma \leq 1), \tag{3.36}$$

which is Gaussian in both x and in u. In Fourier optics, the response of an imaging system can be described by its point spread function (or *psf*). Using the system psf, we can derive the double Wigner distribution function of the system, which gives us a relationship between input and output in terms of Wigner distribution functions. Bastiaans [75] determined the zeroth -, first -, and second - order moments of the Wigner distribution function and derived the propagation of these moments through a first-order system.

3.4 The Elementary Function Method

The Elementary Function Method is a modal method based on work recently developed by Wald et al [64], where a transform is introduced that will give us an exact expression for the cross-spectral density. This method is similar to the coherent-mode method but easier to handle numerically. Using what amounts to approximate modes (which we refer to as *elementary functions* using the terminology of Gabor) that are easily found analytically or numerically, propagation can be reduced to a series of 2D integrals. Unlike the coherent-mode expansion, the elementary function method is not mathematically exact: certain approximations limit its application to relatively well-behaved fields such as short-wavelength partially coherent excimer sources. An elementary function expansion has also previously been applied in both the space-time and space-frequency domains [71], [78], [79], [80]. This section is based on our publication [65].

We begin by examining the continuous case. We assume once again that we are dealing with a partially coherent field, described by the cross-spectral density $W(\mathbf{r}_1,\mathbf{r}_2)$. The intensity is $I_0 = W(\mathbf{r},\mathbf{r})$. All quantities are implicitly assumed to depend on the frequency ν of the light, e.g., $W(\mathbf{r}_1,\mathbf{r}_2,\nu)$, but the notation is dropped for the rest of this discussion. It is assumed that we know $W_0(\mathbf{r}_1,\mathbf{r}_2)$ over a plane, two-dimensional region where $\mathbf{r}_j = (x_j, y_j)$, and we wish to propagate it to any other region. We also need to make an important assumption: that the cross-spectral density is real in the plane where we start the propagation. This is a limitation, but it can be said of most beams if the first plane is chosen cleverly and e.g. the phase functions of lenses included in the propagation instead. The propagation is performed using

$$W(\mathbf{r}_1,\mathbf{r}_2) = \iint_\infty \mathrm{d}^2 r_1' \iint_\infty \mathrm{d}^2 r_2' h^*(\mathbf{r}_1,\mathbf{r}_1') h(\mathbf{r}_2,\mathbf{r}_2') W_0(\mathbf{r}_1',\mathbf{r}_2'), \tag{3.37}$$

where $h(\mathbf{r},\mathbf{r}')$ is an impulse response function and is system-specific. Introducing the transform, we assume the initial cross-spectral density may be written as

$$W_0(\mathbf{r}_1,\mathbf{r}_2) = \iint_\infty \mathrm{d}^2 r' \, a(\mathbf{r}') f(\mathbf{r}_1 - \mathbf{r}') f(\mathbf{r}_2 - \mathbf{r}') \tag{3.38}$$

where the elementary function $f(\mathbf{r})$ is real and even. This expression differs from that defined by Wald et al [64] in that we use a general coefficient $a(\mathbf{r}')$, while they force positivity by using $a^2(\mathbf{r}')$. The expression describes a transform, and provided that the functions $a(\mathbf{r}')$ and $f(\mathbf{r})$ exist, it is exact. Wald et al [64] introduce a method for finding $a(\mathbf{r}')$ and $f(\mathbf{r})$, thus proving that the transform in Equation 3.38 exists.

Taking the Fourier transform of the cross-spectral density, we get

$$\begin{aligned} \widehat{W_0}(\mathbf{r}_1,\mathbf{r}_2) &= \iint_\infty \mathrm{d}^2 r_1 \iint_\infty \mathrm{d}^2 r_2 \iint_\infty \mathrm{d}^2 r' a(\mathbf{r}') f(\mathbf{r}_1 - \mathbf{r}') f(\mathbf{r}_2 - \mathbf{r}') \\ &\quad \times exp(-2\pi i \mathbf{u}_1 \cdot \mathbf{r}_1) exp(-2\pi i \mathbf{u}_2 \cdot \mathbf{r}_2) \end{aligned} \tag{3.39}$$

where the spatial frequencies are $\mathbf{u}_j = (u_j, v_j)$. Changing the order of integration and applying the shift theorem twice yields

$$\widehat{W_0}(\mathbf{r}_1,\mathbf{r}_2) = \hat{f}(\mathbf{u}_1)\hat{f}(\mathbf{u}_2) \iint_\infty \mathrm{d}^2 r' a(\mathbf{r}') exp(-2\pi i(\mathbf{u}_1 + \mathbf{u}_2) \cdot \mathbf{r}'), \tag{3.40}$$

where $\hat{f}(\mathbf{u})$ is the two-dimensional Fourier transform of $f(\mathbf{r})$. Since we have assumed that $f(\mathbf{r})$ is real and even, so is its Fourier transform $\hat{f}(\mathbf{u})$, i.e., $\hat{f}(\mathbf{u}) = \hat{f}(-\mathbf{u})$. We can thus evaluate

$$\widehat{W_0}(\mathbf{u}, -\mathbf{u}) = C[\hat{f}(\mathbf{u})]^2 \tag{3.41}$$

provided the integral

$$C = \iint_\infty d^2 r' \, a(\mathbf{r}') \tag{3.42}$$

exists. We now know that

$$\hat{f}(\mathbf{u}) \propto \sqrt{\widehat{W}_0(\mathbf{u}, -\mathbf{u})} \tag{3.43}$$

and the shape of the function $f(\mathbf{r})$ can be found from the inverse Fourier transform. It can then be normalized.

The coefficient function $a(\mathbf{r}')$ must also be found. As mentioned earlier, the intensity may be written as

$$I_0(\mathbf{r}) = W_0(\mathbf{r}, \mathbf{r}) = \int_\infty d^2r' a(\mathbf{r}') f^2(\mathbf{r} - \mathbf{r}'). \tag{3.44}$$

Taking the Fourier transform, changing the order of integration and applying the shift theorem gives

$$\hat{I}_0(\mathbf{u}) = \hat{a}(\mathbf{u}) \hat{f^2}(\mathbf{u}), \tag{3.45}$$

where $\hat{I}_0(\mathbf{u})$ is the Fourier transform of $a(\mathbf{r})$, $\hat{a}(\mathbf{u})$ is the Fourier transform of $a(\mathbf{r})$, and $\hat{f^2}(\mathbf{u})$ is the Fourier transform of $f^2(\mathbf{r})$, the square of $f(\mathbf{r})$. Rearranging, we get

$$\hat{a}(\mathbf{u}) = \frac{\hat{I}_0(\mathbf{u})}{\hat{f^2}(\mathbf{u})}, \tag{3.46}$$

and knowing that $a(\mathbf{r}')$ and $f(\mathbf{r})$ exist, we can proclaim the transform in(3.38) as valid. Propagation is performed by inserting Equation 3.38 into Equation 3.37. In practice, instead of the exact transform in Equation 3.38, we use the corresponding discrete sum

$$W_0(\mathbf{r}_1, \mathbf{r}_2) = \sum_n \sum_m a_{mn} f(\mathbf{r}_1 - \mathbf{r}_{mn}) f(\mathbf{r}_2 - \mathbf{r}_{mn}). \tag{3.47}$$

Here $\mathbf{r}_{mn} = (x_m, y_n)$ are sampled on a grid as $x_m = m\Delta x$ and $y_n = n\Delta y$. The values of Δx and Δy are found by establishing a sampling criterion based on the partition of unity condition [81]. We take the Fourier transform of the cross-spectral density to give

$$\begin{aligned} \widehat{W}_0(\mathbf{u}_1, \mathbf{u}_2) \quad = \quad & \int\int_\infty d^2r_1 \int\int_\infty d^2r_2 \sum_m \sum_n a_{mn} f(\mathbf{r}_1 - \mathbf{r}_{mn}) \\ & \times f(\mathbf{r}_2 - \mathbf{r}_{mn}) exp(-2\pi i \mathbf{u}_1 \cdot \mathbf{r}_1) exp(-2\pi i \mathbf{u}_2 \cdot \mathbf{r}_2). \end{aligned} \tag{3.48}$$

Interchanging the order of integrations and sum, and performing the Fourier transforms using the shift theorem, yields

$$\widehat{W}_0(\mathbf{u}_1,\mathbf{u}_2) = \hat{f}(\mathbf{u}_1)\hat{f}(\mathbf{u}_2)\sum_m\sum_n a_{mn} exp(-2\pi i(\mathbf{u}_1+\mathbf{u}_2)\cdot\mathbf{r}_{mn}). \tag{3.49}$$

We once again assume $f(\mathbf{r})$ is real and even, so we can say

$$\widehat{W}_0(\mathbf{u},-\mathbf{u}) = \hat{f}^2(\mathbf{u})\sum_m\sum_n a_{mn} \tag{3.50}$$

and, provided that the sum exists, this leads us to

$$\hat{f}(\mathbf{u}) \propto \sqrt{\widehat{W}_0(\mathbf{u},-\mathbf{u})} \tag{3.51}$$

3.4.1 Analytical Examples

In the case of a fully coherent Schell-model source, the cross-spectral density may be written as

$$W_0(x_1,x_2) = \sqrt{I_0(x_1)}\sqrt{I_0(x_2)} \tag{3.52}$$

where $I_0(x)$ is the intensity. It is separable in x_1 and x_2, so its Fourier transform may be written as the product

$$\widehat{W}_0(u_1,u_2) = \mathcal{F}\left\{\sqrt{I_0(x_1)}\right\}[u_1]\cdot\mathcal{F}\left\{\sqrt{I_0(x_2)}\right\}[u_2] \tag{3.53}$$

where $\mathcal{F}$ represents the Fourier transform. Evaluating $W(u,-u)$ quickly yields

$$\hat{f}(u) \propto \mathcal{F}\left\{\sqrt{I_0(x)}\right\}[u] \tag{3.54}$$

and consequently

$$f(x) \propto \sqrt{I_0(x)}. \tag{3.55}$$

This confirms what was obvious already from Equation 3.52: the cross-spectral density is separable, with $f(x) = \sqrt{I_0(x)}$, $a_0 = 1$, and $a_m = 0$ for $m \neq 0$. Only one mode/function is required.

In the case of incoherent light, the cross-spectral density is

$$W_0(x_1,x_2) = \sqrt{I_0(x_1)}\sqrt{I_0(x_2)}\delta(x_1 - x_2), \tag{3.56}$$

and its Fourier transform

$$\widehat{W}_0(u_1,u_2) = \hat{I}_0(u_1 + u_2). \tag{3.57}$$

So $W(u,-u) = \hat{I}_0(0)$ is a constant, and $\hat{f}(u)$ must be a constant too. Then $f(x) \propto \delta(x)$, and since a delta function is used for interpolation the sampling interval Δx must be infinitely small and the continuous theory must be used. This gives no advantage in computation speed. Besides, for incoherent light the normal method of propagation reduces to solving 2D integrals, so the modal method is not needed from numerical point of view.

In the case of a Gaussian Schell-model beam (a partially coherent source), the elementary function is given by

$$W_0(x_1,x_2) = A\exp\left(\frac{x_1^2}{\sigma_\mathrm{I}^2}\right)\exp\left(\frac{x_2^2}{\sigma_\mathrm{I}^2}\right)\exp\left[\frac{(x_1 - x_2)^2}{\sigma_\mathrm{g}^2}\right] \tag{3.58}$$

where σ_I represents the width of the intensity distribution and σ_g the width of the coherence distribution. Following a series of calculations, we arrive at an expression for the elementary function

$$f(x) = \frac{1}{(\sigma_\mathrm{I}\sigma_\mathrm{g})^{1/2}}\left(\frac{\sigma_\mathrm{I}^2 + \sigma_\mathrm{g}^2}{\pi}\right)^{1/4}\exp\left[-\frac{1}{2}x^2\left(\frac{1}{\sigma_\mathrm{I}^2} + \frac{4}{\sigma_\mathrm{g}^2}\right)\right] \tag{3.59}$$

We can see that it is a Gaussian, and that its width depends on both σ_I and σ_g. For example, if $\sigma_\mathrm{g} \gg \sigma_\mathrm{I}$ the light is nearly coherent, and the elementary function will be proportional to the field distribution $\sqrt{I_0(x)}$. This implies only one such function is required to represent the field. On the other hand, if $\sigma_\mathrm{g} \ll \sigma_\mathrm{I}$ the light is nearly incoherent, and the elementary function will become very narrow and resemble a delta function. Very small Δx is then required.

Continuing the analysis to find the coefficients a_m, we get

$$a_m = \Delta x A\sigma_\mathrm{I}\sqrt{\pi}\int_{-1/2\Delta x}^{1/2\Delta x} \mathrm{d}u \exp\left(-\pi^2 u^2 \frac{4\sigma_\mathrm{I}^4}{4\sigma_\mathrm{I}^2 + \sigma_\mathrm{g}^2}\right)\exp\left(2\pi \mathrm{i} m\Delta x u\right). \tag{3.60}$$

and the sampling distance Δx is given by

$$\Delta x = \frac{\pi \sigma_{\mathrm{I}} \sigma_{\mathrm{g}}}{\sqrt{(-\ln c)(\sigma_{\mathrm{g}}^2 + 4\sigma_{\mathrm{I}}^2)}}. \tag{3.61}$$

where the value of c depends on our definition of the full width of the Gaussian distribution. This will be discussed further in Chapter 5.

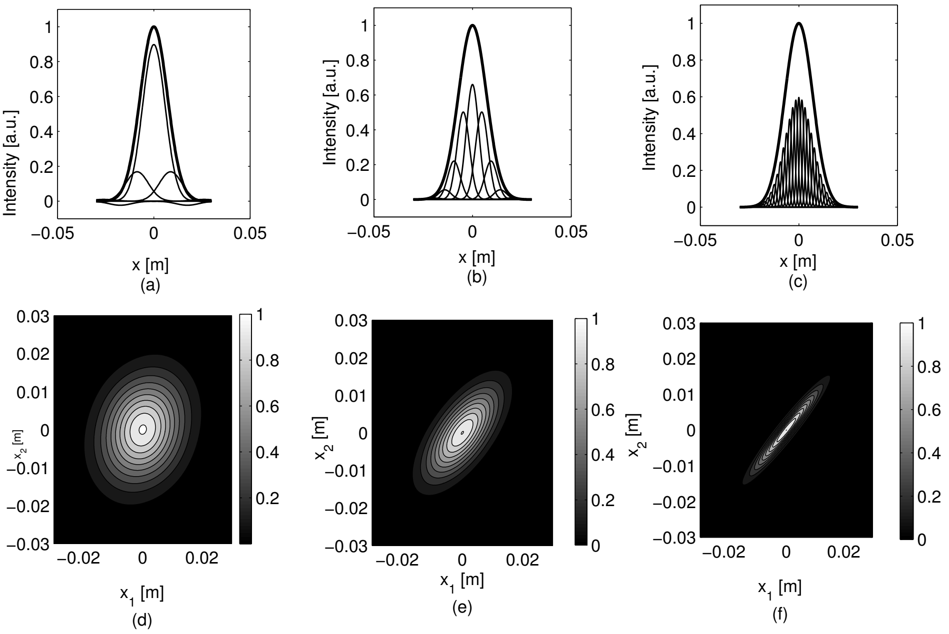

Figure 3.2: (a)-(c) Intensity and reconstructed intensity (thick line; they are almost identical in all three cases) are shown for 1-D Gaussian Schell-model beams of different degrees of coherence, along with the scaled and shifted squares of the elementary functions (thin lines). Adding those together gives the reconstructed intensity. The intensity distribution is the same for all beams, with $\sigma_I = 0.01\ m$, while the coherence varies from high to low as (a) $\sigma_g = 0.03\ m$ with 7 elementary functions required, (b) $\sigma_g = 0.01\ m$ with 13 elementary functions required, and (c) $\sigma_g = 0.003\ m$ with 39 elementary functions required. (d)-(f) Reconstruction of the cross-spectral density for the parameters in (a), (b), and (c) respectively.

Figure 3.2 illustrates the position and weighting of the elementary functions and the reconstructed intensity.

3.4.2 Numerical Analysis

To evaluate this transform numerically, we construct the cross-spectral density of the Schell-model beam using measurable quantities: the intensity profile of the beam I_0, and the coherence distribution γ.

$$W_0(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{I_0(\mathbf{r}_1)}\sqrt{I_0(\mathbf{r}_2)}\gamma(\mathbf{r}_1 - \mathbf{r}_2). \tag{3.62}$$

Both $I_0(\mathbf{r})$ and $\gamma(\mathbf{r})$ are 2D quantities, and so finding the Fourier transform of $W_0(\mathbf{r}_1, \mathbf{r}_2)$ involves a 4D integral. To avoid this calculation, we go directly to the 2D quantity $\widehat{W}_0(\mathbf{u}, -\mathbf{u})$ by

$$\widehat{W}_0(\mathbf{u}, -\mathbf{u}) = \iint_\infty d^2u' \hat{U}^2(\mathbf{u} - \mathbf{u}')\hat{\gamma}(\mathbf{u}') \tag{3.63}$$

where $\hat{U}(\mathbf{u})$ and $\hat{\gamma}(\mathbf{u})$ are the Fourier transforms of $U(\mathbf{r})$ and $\gamma(\mathbf{r})$ respectively. The detail of this calculation can be found in Reference [65].

From the various methods discussed in this chapter, we proceed with just one: the Elementary Function Method. To apply this model to a real system, we require real beam data. In the following chapter, methods to measure the spatial coherence of the source are discussed, and the laboratory design and results are presented. In Chapters 5 and 6, the Elementary Function Method is applied to an ideal Gaussian Schell-model beam, and to a real excimer laser beam.

Chapter 4

Measuring Spatial Coherence

Excimer lasers are operated in pulsed mode with a broad linewidth, meaning that both the temporal and spatial coherence are low. The linewidth of a laser is the full width at half-maximum (FWHM) of its optical spectrum. More precisely, it is the FWHM of the power spectral density of the emitted electric field in terms of frequency, wavenumber or wavelength. Excimer lasers are quite often line-narrowed for use in lithographic systems which can increase the spatial coherence of the laser quite dramatically. Many lithographic systems also use homogenizing optics to smooth the intensity profile of the beam which can have an effect on the spatial coherence length of the source (see Chapter 5). In order to develop an accurate model of a lithographic imaging system, precise information about the degree of spatial coherence of the source (and the effects of beam homogenization on that coherence) is needed. In this chapter, some established methods to measure the spatial coherence of a laser source are explored, including methods involving interferometry (Section 4.1) and the use of coded arrays (Section 4.2). Experimental results for the measurement of spatial coherence are presented and analysed in Section 4.3.

4.1 Interferometric Methods

The concept of optical coherence has long been associated with interference, presumably because interference is the simplest phenomenon that reveals correlation between light beams. Many methods to measure spatial coherence of a light source are based on the Young's double slit interferometer as discussed in an earlier section (see Figure 2.3). In 1938, Zernike determined the degree of coherence directly from the visibility of the interference fringes formed in a Young interferometer. Visibility, which is a measure of contrast of interference fringes, is a concept introduced by Michelson (1890). The visibility $V(\mathbf{r})$ at a point $P(\mathbf{r})$ in an interference pattern is defined as

$$V(\mathbf{r}) = \frac{\langle I\rangle_{max} - \langle I\rangle_{min}}{\langle I\rangle_{max} + \langle I\rangle_{min}}, \tag{4.1}$$

where $\langle I\rangle_{max}$ and $\langle I\rangle_{min}$ are the maximum and minimum values that the averaged intensity assumes in the immediate neighbourhood of P. Methods based on the measurement of fringe visibility have involved masks containing two apertures, but this has the disadvantage that the mask must be moved laterally to sample all parts of the beam. Also, if the phase and modulus of the complex degree of spatial coherence are shift variant, a single interferogram is not sufficient to fully characterise the light field. Methods using coded arrays have an advantage over two aperture systems: the data gathered from a single exposure of an array of many pinholes can be equivalent to several exposures of a double aperture system.

4.2 Methods using Coded Arrays

Nugent and Trebes [82] and Castañeda [56], [57] proposed an alternative to the double pinhole measurement: using a mask with multiple apertures spaced evenly and analysing the Fourier spectrum. This type of mask became known as a coded array. A coded array is defined to be a pattern on a periodic two-dimensional lattice which associates with each lattice point a 0 or a 1 indicating whether the lattice point is "open" or "closed". In coded aperture imaging, the open and closed lattice points become open and closed cells in an opaque mask. The spacing between apertures in the array determines the class of coded array: if the spacing between each pair of apertures is unique, the array is called *non-redundant*; if the spacing between each set of apertures is equal, then the array is called *uniformly redundant*.

Coded arrays were originally conceived for applications in X-ray imaging by Mertz & Young 1961 and by Dicke in 1968 [83]. X-ray sources are usually so weak that a single pinhole camera would have to have a very large opening in order to obtain a reasonable signal to noise ratio (SNR), which reduces the resolution of the system. If, instead, N pinholes are used to image the object, the image produced consists of N overlapping images of the object. The coded aperture technique (for a point source) can improve the SNR by roughly $\sqrt{N}$ when compared to the single pinhole camera [83].

4.2.1 Uniformly Redundant Arrays

A uniformly redundant array (URA) is a particular form of coded array [84]. For a URA, each possible vector displacement between pairs of equivalent open lattice points occurs a uniform or precisely equal number of times. The power spectrum of the array is flat up to some limit determined by the minimum aperture separation and so has a perfectly sharp autocorrelation function with perfectly flat sidelobes, provided the autocorrelation function is calculated on a discrete grid with separations equal to the minimum aperture separation [82]. The spacing between each aperture pair is not unique and corresponds to a single frequency in the Fourier analysis of the resultant interference pattern. As a result, URAs are not suited to making spatial coherence measurements.

4.2.2 Non-Redundant Arrays

A non-redundant array (NRA) is an array of points on a regular lattice such that the vector difference between any two points in the array is unique. This type of coded array is most suitable for spatial coherence measurements: all classes of aperture pairs of the mask are composed by only one pair and so contribute only once to the resulting interference pattern. Each frequency is represented only once. There are some considerations in the design of the array. If the radius of the apertures is much smaller than the dimensions of the coherence area, then the variations of the light field within each aperture are not significant to make an accurate measurement [85].

Mejía and Gonzáles [85] have shown experimental results of a method involving an NRA. Their experimental set-up includes an array with five apertures. Each spacing is unique and an integer multiple of the smallest spacing. The mask is illuminated by a coherent laser source (633 *nm*) decohered using a piece of rotating ground glass. When the array of pinholes is illuminated by the source, a pattern is produced which is the result of the interference of each aperture pair adding together.

Fourier analysis of the far-field interferogram follows. The degree of spatial coherence can be deduced from the visibility of the interference fringes multiplied by a factor dependent on the intensity at each aperture as shown by Wolf [86].

4.2.3 Experimental Design

In our design, we considered the advantages of the pinhole array over a two-pinhole arrangement. When using a mask with multiple apertures, one measurement is sufficient to deduce the spatial coherence of the source. The sizes of the pinholes were chosen to allow the mask to be used with sources of $193nm$ and $248nm$ wavelengths. The design of the pinhole plate is shown in Figure 4.1. The layout of the pinholes is suitable for measurements of spatial coherence in both the horizontal and vertical direction, or the long and short axes of the beam. The diameter of the pinholes differs in the arrays: the pinholes in array 'A' are 5 μm in diameter, the pinholes in the arrays labelled 'B' are 10 μm in diameter, and in arrays labelled 'C' the pinhole diameter is 15 μm. Figure 4.2 is an expanded diagram of the layout of the pinholes in each array. The spacings chosen form an array with one degree of redundancy, i.e. one aperture pair spacing occurs more than once (see Table 4.1). The chosen design will allow coherence measurements at several spacings within the expected coherence length of the sources, but the trade-off is one separation occurring twice. This can be avoided by increasing the pinhole separations, but this results in fewer pinhole pairs within the coherence length of the source. To optimise the measurements, his compromise was accepted, and the data corresponding to the redundant spacing was ignored during data analysis. Golay presents a discussion on the optimization of nonredundant arrays in [87].

In Table 4.1, it is shown that the interference between pinholes 2 and 4 and pinholes 4 and 5 arises from a separation of $0.40mm$. This will affect the distribution of intensity among the interference fringes. The number of pinholes (five) will give nine classes of aperture pair . Under spatially partially coherent illumination, the effective contributing classes will be chosen by the modulus of the complex degree of spatial coherence, in such a way that its magnitude specifies the weight of the contribution and its support the number of contributing classes. From this, the modulus and complex degree of coherence can be determined.

The expected results can be derived using the elementary equations for the theory of partial coherence found in Born & Wolf [37]. Born & Wolf show that when two quasi-monochromatic light beams interfere, the intensity at point Q in the interference

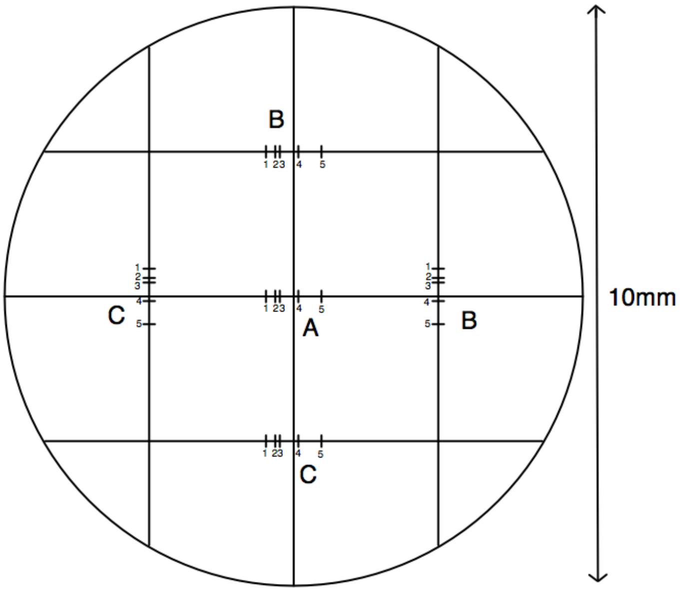

Figure 4.1: The finished design of mask with arrays of five pinholes. The pinhole diameters are $A = 5\ \mu m$, $B = 10\ \mu m$, and $C = 15\ \mu m$.

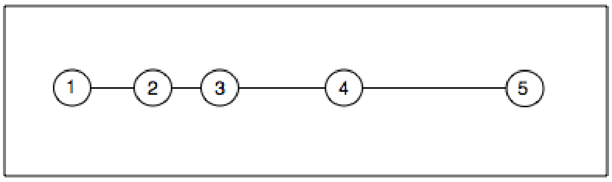

Figure 4.2: Layout of five-pinhole array. Spacings: $1-2 = 0.16\ mm$, $2-3 = 0.08\ mm$, $3-4 = 0.32\ mm$, and $4-5 = 0.40\ mm$.

Pair	Class	Separation (mm)
2,3	1	0.08
1,2	2	0.16
1,3	3	0.24
3,4	4	0.32
2,4	5	0.40
4,5	5	0.40
1,4	7	0.56
3,5	9	0.72
2,5	10	0.80
1,5	12	0.96

Table 4.1: Classes of aperture pairs yielded by the mask of Figure 4.2.

pattern is given by

$$I^{(2)}(Q) = I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|j_{12}|cos[\beta_{12} - \delta]. \tag{4.2}$$

This equation is valid as long as the path difference $|s_2 - s_1| = c|\tau|$, introduced between the interfering beams, is small compared to the coherence length, i.e.

$$|\Delta S| = |s_2 - s_1| = \frac{\bar{\lambda}}{2\pi}\delta << \frac{\bar{\lambda}^2}{\Delta\lambda}, \tag{4.3}$$

where $c/\Delta\nu = \lambda^2/\Delta\lambda$. Within this range of validity, the correlation between the vibrations at any two points P_1 and P_2 in the wave field is characterized by the mutual intensity, J_{12}, rather than by $\Gamma_{12}(\tau)$, i.e. by a quantity which depends on the positions of the two points, but not on the time difference τ. It follows that

$$|\gamma_{12}(\tau)| \approx |j_{12}|, \tag{4.4}$$

so that $|j_{12}|$ represents the degree of coherence of the vibrations at P_1 and P_2 (and $0 \leq |j_{12}| \leq 1$). The phase β_{12} of $|j_{12}|$ represents their effective phase difference. j_{12}, just like $\gamma_{12}(\tau)$ of which it is a special case, is usually called the *complex degree of coherence* and J_{12} is called the *mutual intensity*.

Equation 4.2 can be extended for interference of multiple sources, as in the case of a pinhole array. In the case of three pinholes, the intensity at point Q in the interference pattern is given by

$$\begin{aligned} I^{(3)}(Q) &= I^{(1)}(Q) + I^{(2)}(Q) + I^{(3)}(Q) \\ &\quad + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|j_{12}|cos[\beta_{12} - \delta] \\ &\quad + 2\sqrt{I^{(2)}(Q)}\sqrt{I^{(3)}(Q)}|j_{23}|cos[\beta_{23} - \delta] \\ &\quad + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(3)}(Q)}|j_{13}|cos[\beta_{13} - \delta] \end{aligned} \tag{4.5}$$

If we assume uniform intensity across the original laser source, then the intensity of the light leaving each pinhole is equal. We substitute $I'(Q)$ for the intensity from each pinhole i.e.

$$I^{(1)}(Q) = I^{(2)}(Q) = I^{(3)}(Q) = I'(Q) \tag{4.6}$$

Substituting $I'(Q)$ into the previous equation gives

$$\begin{aligned} I^{(3)}(Q) &= 3I'(Q) + 2I'(Q)|j_{12}|cos[\beta_{12} - \delta] \\ &\quad + 2I'(Q)|j_{23}|cos[\beta_{23} - \delta] \\ &\quad + 2I'(Q)|j_{13}|cos[\beta_{13} - \delta] \end{aligned} \tag{4.7}$$

Performing a Fourier transform of the intensity expression gives

$$\begin{aligned} \tilde{I}^{(3)}(u) &= 3I'(Q)\left[\delta(u) + \frac{|j_{12}|}{3}(\delta(u-u_1) + \delta(u+u_1))\right. \\ &\quad + \frac{|j_{23}|}{3}(\delta(u-u_2) + \delta(u+u_2)) \\ &\quad \left. + \frac{|j_{13}|}{3}(\delta(u-u_3) + \delta(u+u_3))\right] \end{aligned} \tag{4.8}$$

This shows that a Fourier transform of the intensity in the image plane of the interference pattern will produce a delta function with magnitude proportional to the sum of the intensities of the interfering beams. On either side of this delta function, there are delta functions of a lower magnitude, corresponding to a different frequency or interaction between a pair of apertures. For an interference problem with three apertures, there are three resulting delta functions located on either side of the central maximum. The spacing of the functions is related to the distance between interfering

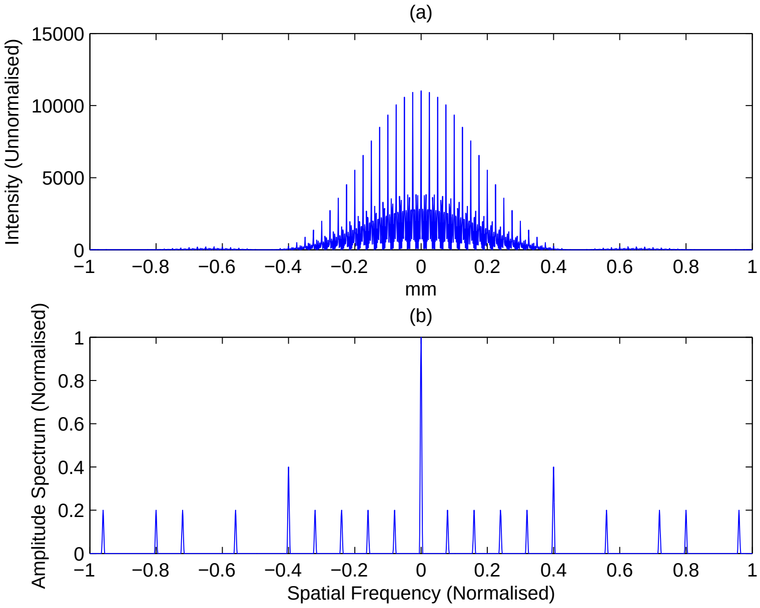

Figure 4.3: (a) Simulated intensity cross-section of the interference pattern produced by five pinholes illuminated with spatially coherent light, (b) The normalised Fourier magnitude spectrum of (a). Each delta function has a magnitude equal to one fifth of the normalised central maxiumum, except the class 5 which is doubled as it occurs twice in the pinhole mask. Data for class 5 spacing is omitted.

apertures in the array, their height depends on the number of apertures in the array. In the case of three pinholes illuminated with fully spatially coherent light, the height of the sidebands is one third of the normalised central delta function. For four pinholes, the sideband height is one quarter that of the central delta function; for five pinholes, the magnitude is one fifth. In Figure 4.3(a), the intensity cross-section for a simulation of the pinhole array in Figure 4.2 illuminated with perfectly spatially coherent light is shown. Figure 4.3(b) shows the corresponding normalised Fourier spectrum in this case. As expected, for an array of 5 pinholes, the magnitude of the delta functions on either side of the central maximum is one fifth the magnitude of the central peak.

The magnitude of the sidebands decreases with decreasing spatial coherence. According to Equation 4.8, taking the ratio of the magnitude of each sideband with respect to the central (unnormalised) delta function and multiplying by a factor related to the number of pinholes in the array, the visibility of the interference fringes can be calculated. A plot of pinhole separation versus fringe visibility will give a value for the complex degree of coherence, j_{12}, which is analogous to γ_{12} defined previously.

4.3 Experimental Results

4.3.1 Measurements at 193nm

Spatial coherence measurements were carried out by illuminating the 10 μm and 15 μm pinhole arrays with an ArF excimer laser operating at 193 nm. This is an ATLex 300 laser source with a high voltage electrical discharge ($12 - 25$ kV) capable of generating 7 ns duration pulses at 193nm. The laser has a stable resonator configuration with two plane parallel mirrors. The output coupler is made of an MgF_2 substrate and is coated with $R = 25$ % reflection coating. The rear mirror is also made of MgF_2, but is covered with a high reflective coating ($R > 98$ %).

To avoid saturating the detector, the beam is attenuated. The attenuation used was a programmable Optec AT4020 Energy Controller. It relies on the principle of the shift in cut-off wavelength which occurs when a multi-layer dielectric high-pass edge filter is tilted with respect to the beam. The attenuating element is a fused silica plate for use at wavelengths down to 193nm and below. Transmission reduces progressively with tilt angle in a non-linear fashion to a low value of typically < 5 % at 45 ° tilt. The part of the beam not transmitted is reflected at twice the title angle (90 ° for minimum transmission). In the AT4020 this reflected component is scattered/absorbed in a light trap/heat sink surrounding the optics. One complete turn of the operating dial gives a 45 ° plate tilt.

The attenuated beam illuminates the mounted pinhole arrays. The arrays are mounted with three degrees of freedom, and connected to the detector using a bag bellows to avoid contamination from background light and minimise reflections. The detector used was a Hamamatsu C8484-16C UV Digital Camera. Measurements were carried out using 10μm, and 15μm pinhole arrays in both horizontal and vertical orientation to correspond to the long and short axes of the beam respectively.

Figures 4.4 and 4.6 show the raw images from the 193nm source after passing through 10 μm and 15 μm pinhole arrays respectively (arrays B and C in Figure 4.1). In both cases, we can almost observe the classic Airy disk diffraction pattern created by the laser light passing through the circular apertures. The Airy disk rings are clearer in Figure 4.5, which is a greyscale version of Figure 4.4. The size and form of the Airy disk is given by

$$I(\theta) = I_0 \left(\frac{2J_1(ka\,sin\theta)}{ka\,sin\theta} \right)^2 \tag{4.9}$$

where a is the radius of the circular aperture (in this case 5 μm and 7.5 μm), k is

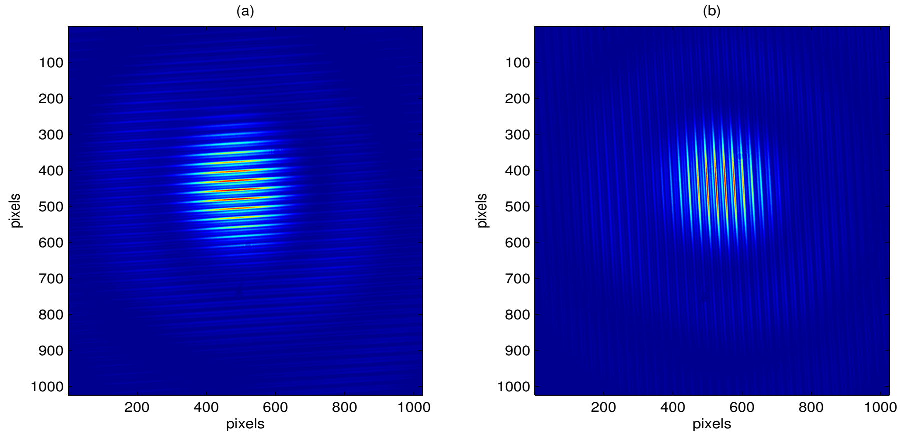

Figure 4.4: Raw intensity images using 10 μm pinholes on the 193 nm ArF laser source. (a) Interference pattern generated by pinholes orientated along the long axis, and (b) along the short axis of the beam.

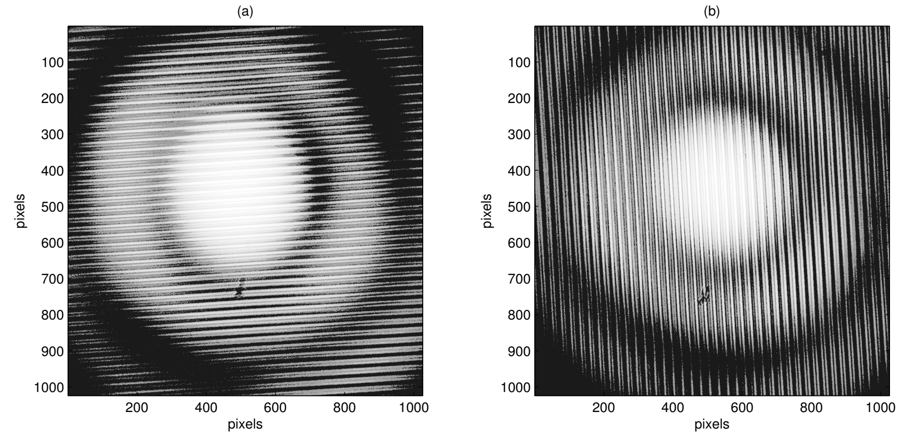

Figure 4.5: A greyscale version of the raw intensity images using 10 μm pinholes on the 193 nm ArF laser source. (a) Interference pattern generated by pinholes orientated along the long axis, and (b) along the short axis of the beam. The classic Airy pattern becomes more visible with a change of contrast.

given by $2\pi/\lambda$, and J_1 is a Bessel function of the first order. The size of the aperture determines the size of the central spot in the diffraction pattern; the smaller the aperture, the larger the spot size. The central spot size observed in the diffraction pattern produced by the 15 μm is indeed smaller than the pattern produced by the 10 μm pinholes.

Within the Airy disk envelope we observe the interference pattern created by the multiple pinholes illuminated by spatially partially coherent light. A larger number of interference lines are observed along the short axis of the beam in data taken with both the 10 μm and 15 μm pinholes. In Figures 4.8(a) and 4.10(a), we observe deep modulation in the intensity profile for measurements taken along the short axis of the beam, indicating a higher degree of coherence. This modulation is less pronounced in the data gathered along the long axis of the beam (Ref: Figures 4.7(a) and 4.9(a)), indicating a lower degree of coherence in this plane. This is especially obvious in the data taken using the 15 μm pinholes (see Figure 4.9).

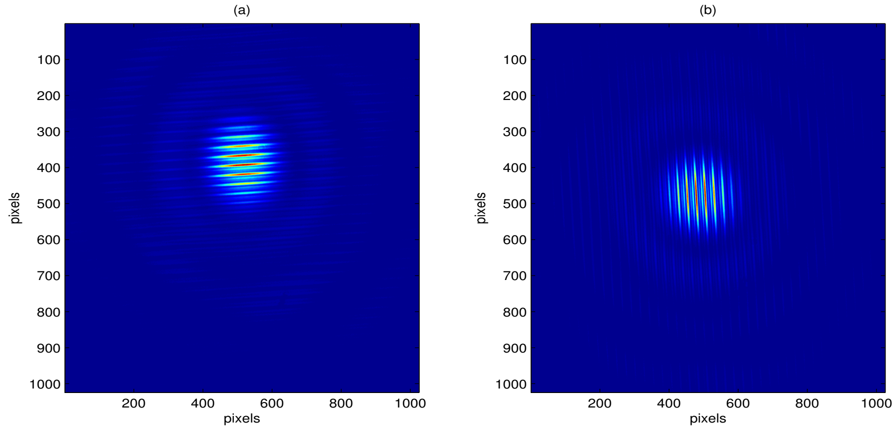

Figure 4.6: Raw intensity images using 15 μm pinholes on the 193nm ArF laser source. (a)Interference pattern generated by pinholes orientated along the long axis, and (b) along the short axis of the beam.

To process this interference data, a line sample was selected perpendicular to the interference lines, through the centre of the image, to include the pixels of maximum intensity. This sample was extended to include 10 pixels either side (total 21 lines of intensity values) and integrated to calculate an average intensity value. Finally, a Fourier transform was performed on the resulting $1 \times n$ array of averaged intensity values, where n is the number of pixels in the original line sample. The degree of

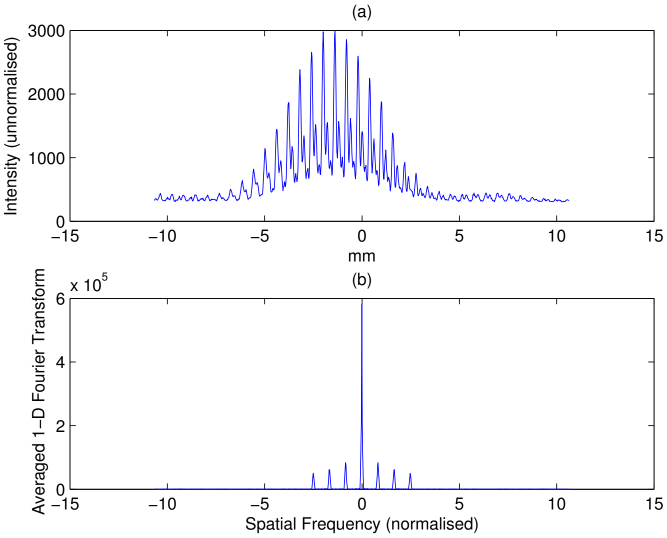

Figure 4.7: (a) Intensity cross-section (unnormalised) and (b) corresponding averaged 1-D Fourier transform of a sample taken along the long axis of the beam using 10 μm pinholes illuminated at 193 nm (see Figure 4.4 (a)).

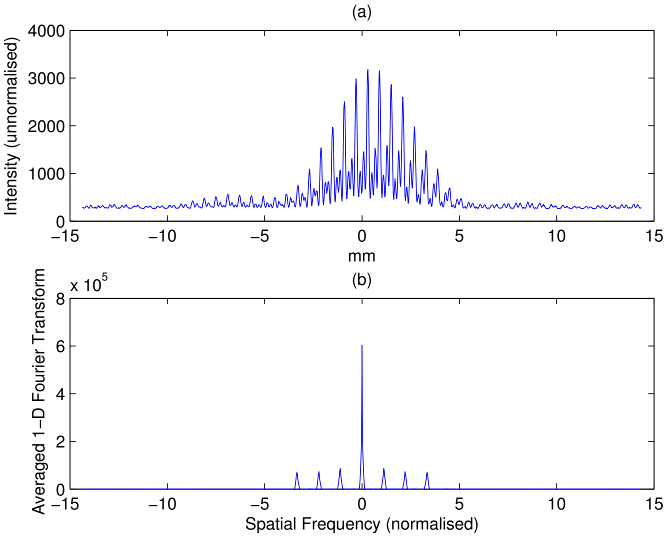

Figure 4.8: (a) Intensity cross-section (unnormalised) and (b) corresponding averaged 1-D Fourier transform of a sample taken along the short axis of the beam using 10 μm pinholes illuminated at 193 nm (see Figure 4.4 (b)).

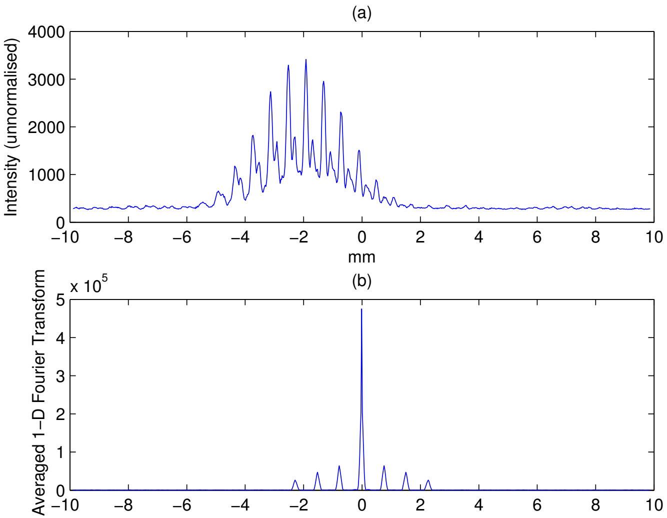

Figure 4.9: (a) Intensity cross-section (unnormalised) and (b) corresponding averaged 1-D Fourier transform of a sample taken along the long axis of the beam using 15 *μm* pinholes illuminated at 193 *nm* (see Figure 4.6 (a)).

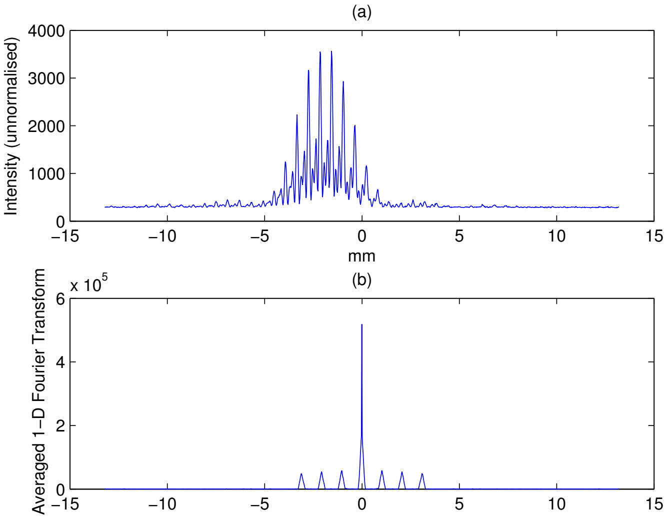

Figure 4.10: (a) Intensity cross-section (unnormalised) and (b) corresponding averaged 1-D Fourier transform of a sample taken along the short axis of the beam using 15 *μm* pinholes illuminated at 193 *nm* (see Figure 4.6 (b)).

spatial coherence can be deduced from the Fourier spectrum as follows. Taking the unnormalised data in the Fourier magnitude spectrum, the ratio of the magnitude of each delta function with respect to the central delta function is calculated. This ratio is then multiplied by a factor relating to the number of pinholes in the array, which is 5 in this case. The result is a measure of the *visibility*, which we can equate to j_{12}, the complex degree of coherence.

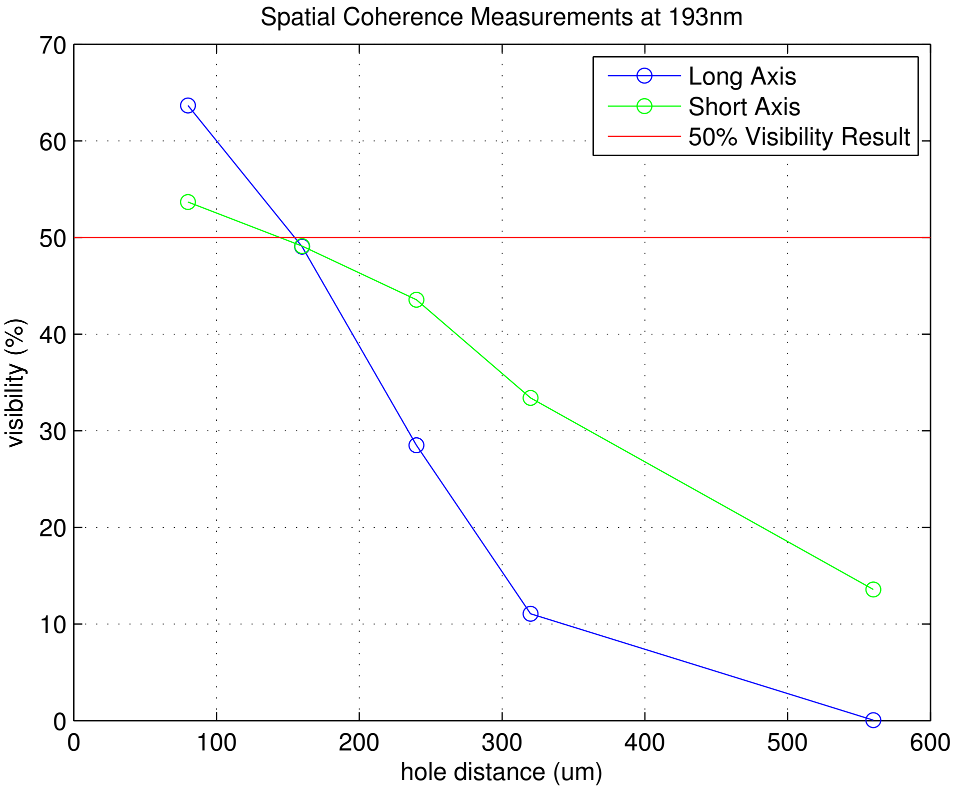

Figure 4.11: Plot of average visibility(%) as a function of pinhole separation (μm) for the long and short axis of a 193 nm source. The pinholes used in the measurements were 10 μm and 15 μm in diameter.

In Figure 4.11, we plot visibility as a function of hole distance. The shortest hole distance corresponds to the smallest separation in the pinhole array. Each subsequent hole distance on the plot relates directly to a pinhole separation, or the interference between a pinhole pair. According to Zernike [45] *"the degree of coherence of two light-vibrations shall be equal to the visibility of the interference fringes that may be obtained from them under the best circumstances, that is, when both intensities are made equal and only* *small path-differences introduced"*. We can assume that the experimental conditions satisfy these conditions, and thus the maximum value of the visibility in the long and short axis of the beam for the 193 *nm* source (and also the 248 *nm* source) is equal to the degree of coherence γ. For this 193 *nm* source, the coherence length in the long axis is slightly greater than along the short axis.

4.3.2 Measurements at 248nm

Further spatial coherence measurements were taken using 248 *nm* KrF Excimer source. This is a pulsed source running at 5 *Hz* with an output of 4.2 *mJ*. The raw laser beam had a rectangular profile 8 *mm* by 3*mm* approx arriving at the pinhole plate. The setup used in these measurements was the same as used for the 193nm source, with the exception that a different attenuation method was used when necessary. This laser is in use in a lithographic system at HP DIMO (Leixlip), with a beamsplitter 'dumping' some of the light before entering the lithographic optics. Measurements were taken using this unused arm of the system.

The measurement apparatus is similar to that used for the measurements at 193 *nm*. The laser light is incident on the pinhole array which is attached to the detector by a flexible bag bellow. When necessary (to avoid saturation of the detector), the source was attenuated using a Neutral Density Filter (ODE = 0.8), giving a transmission of $15-16$ % at 248 *nm*.

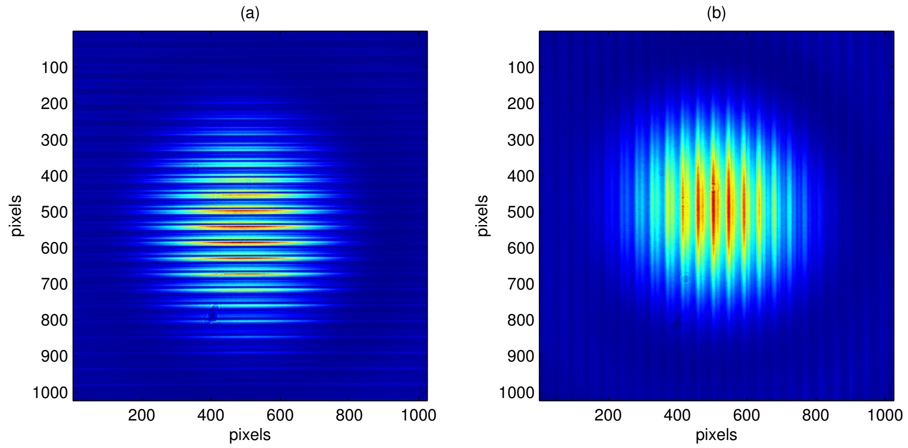

Figure 4.12: Raw intensity images using 10 *µm* pinholes on the 248 *nm* KrF laser source. (a)Interference pattern generated by pinholes orientated along the short axis, and (b) along the long axis of the beam.

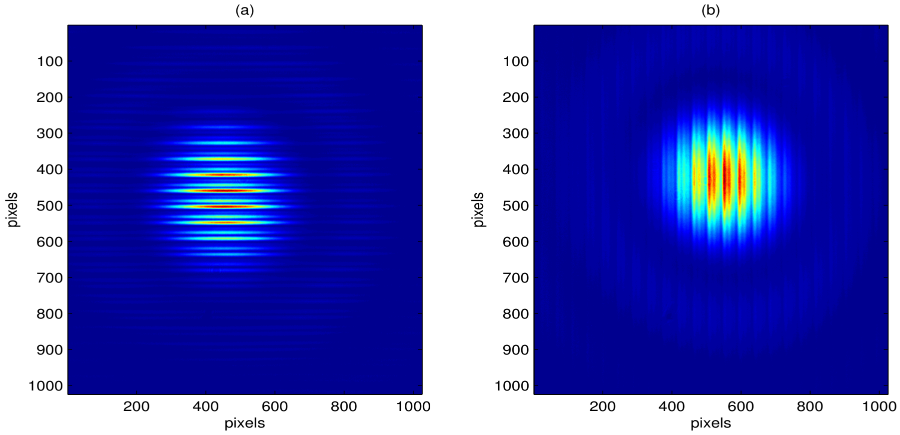

Figure 4.13: Raw intensity images using 15 μm pinholes on the 248 nm KrF laser source. (a) Interference pattern generated by pinholes orientated along the short axis, and (b) along the long axis of the beam.

As before, the pinhole plate was orientated such that both vertical and horizontal measurements were taken, corresponding to the long and short axis of the beam profile respectively. Figure 4.12 shows the output image from the camera for the case of the 10 μm pinholes. Figure 4.12(a) shows the interference resulting from the spatial correlations along the short axis of the beam. Figure 4.12(b) represents the spatial correlations along the long axis of the beam. Figure 4.13 shows the interference patterns observed on illuminating the 15 μm pinholes at 248 nm; Figure 4.13(a) is the interference observed along the short axis of the beam, and Figure 4.13(b) is the result along the long axis of the beam.

As with the 193 nm data, a line sample was selected perpendicular to the interference lines, through the centre of the image, to include the pixels of maximum intensity. This sample was extended to include 10 pixels either side and integrated to calculate an average intensity value. A Fourier transform was then performed on the the resulting $1 \times n$ array of averaged intensity values, where n is the number of pixels in the original line sample. Figure 4.14 shows the normalised intensity cross-section and the unnormalised Fourier transform of the raw image data as seen in 4.12(a). Similarly, Figure 4.15 corresponds to the image seen in Figure 4.12(b).

The data from the 15 μm pinholes is given in Figures 4.16 and 4.17. Observing the intensity cross-section (Figures 4.14(a), 4.15(a), 4.16(a), and 4.17(a)), we can deduce that the spatial coherence of this source is slightly greater along the short axis of the

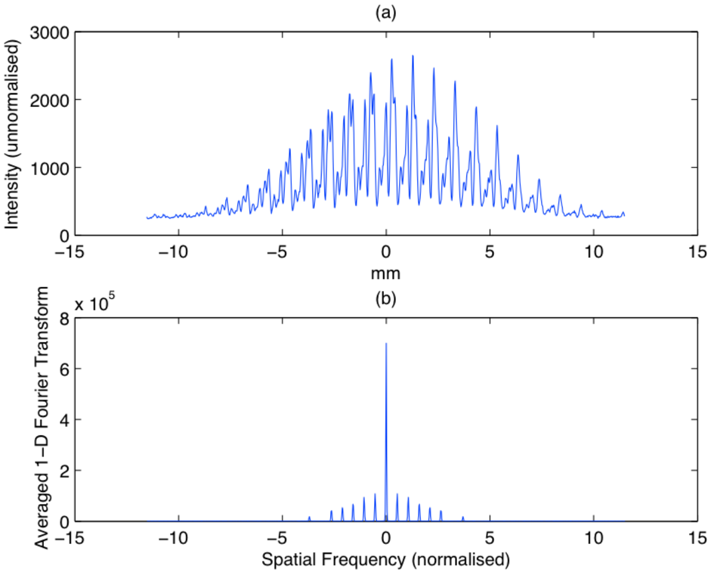

Figure 4.14: (a) Intensity cross-section (normalised) and (b) corresponding 1-D Fourier transform of an integrated sample (unnormalised) taken from the correlation data along the short axis of the beam (see Figure 4.12(a)).

beam than along the long axis. To investigate this further, we plot the visibility of the interference fringes, in both cases, as a function of pinhole separation.

Figure 4.18 shows the visibility of the interference fringes as a function of pinhole separation for both the long and short axes for the 248 nm source. An estimate of the spatial coherence of the source in both directions can be taken at a reference level of 50 % visibility, which in this case gives a coherence length of 285 μm in the short axis. If we compare these results with those obtained by Coherent Inc following the most recent off-site repair, they are consistent. The source has a greater coherence length in the short axis of the beam than in the long axis. Coherent give a value of 275 μm and our results give an initial estimate of $\approx$ 270 μm for the coherence length in the short axis of the beam. In the long axis, Coherent Inc report a coherence length of 75 μm. Our apparatus did not register a measurement for pinhole separation of less than 80 μm which falls below the 50 % visibilty reference level that we used to estimate the spatial coherence length. However, this result is still suitable for inclusion in our beam model as the entire coherence distribution is taken into account.

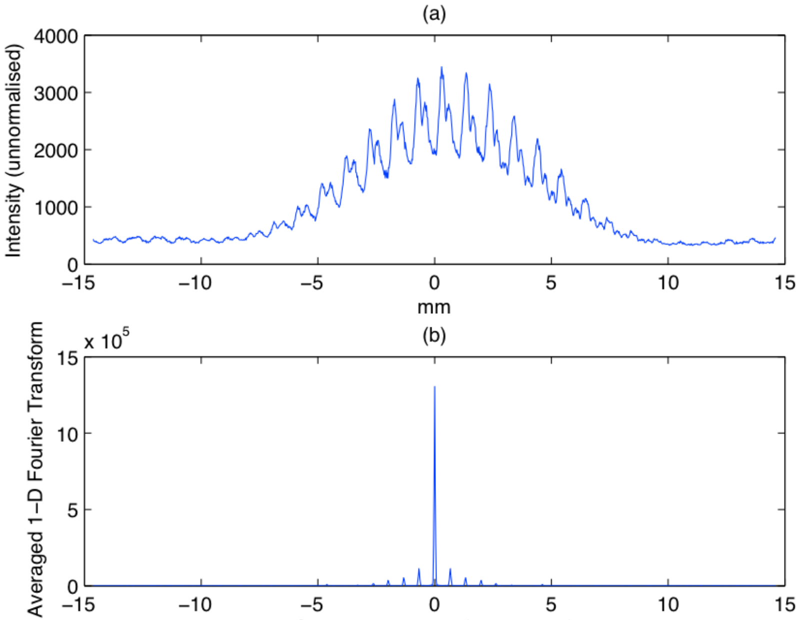

Figure 4.15: (a) Intensity cross-section (normalised) and (b) corresponding 1-D Fourier transform of an integrated sample (unnormalised) taken from the correlation data along the long axis of the beam (see Figure 4.12(b)).

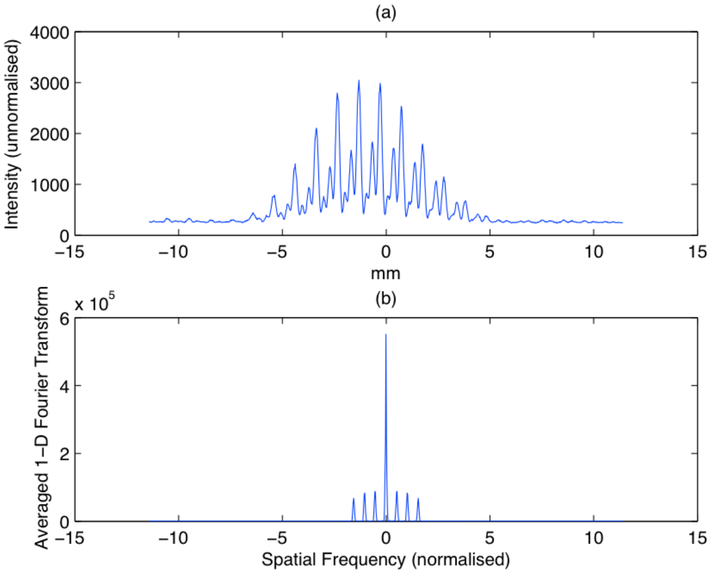

Figure 4.16: (a) Intensity cross-section (normalised) and (b) corresponding 1-D Fourier transform of an integrated sample (unnormalised) taken from the correlation data along the short axis of the beam (see Figure 4.13(a)).

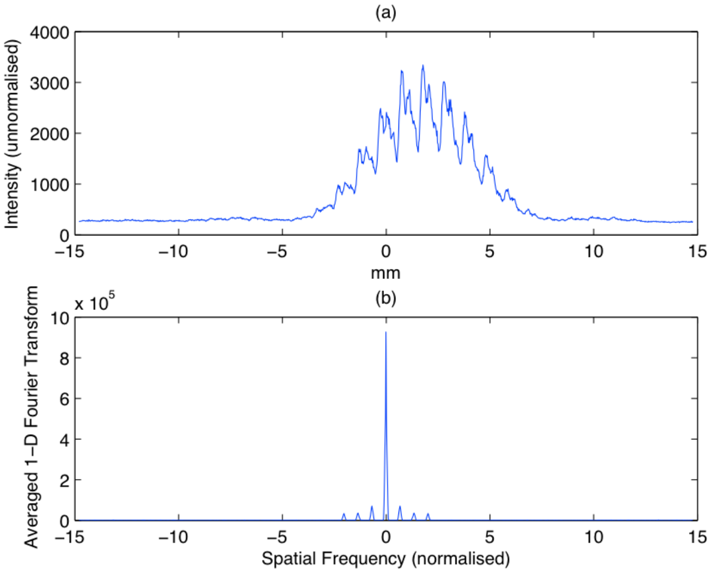

Figure 4.17: (a) Intensity cross-section (normalised) and (b) corresponding 1-D Fourier transform of an integrated sample (unnormalised) taken from the correlation data along the long axis of the beam (see Figure 4.13(b)).

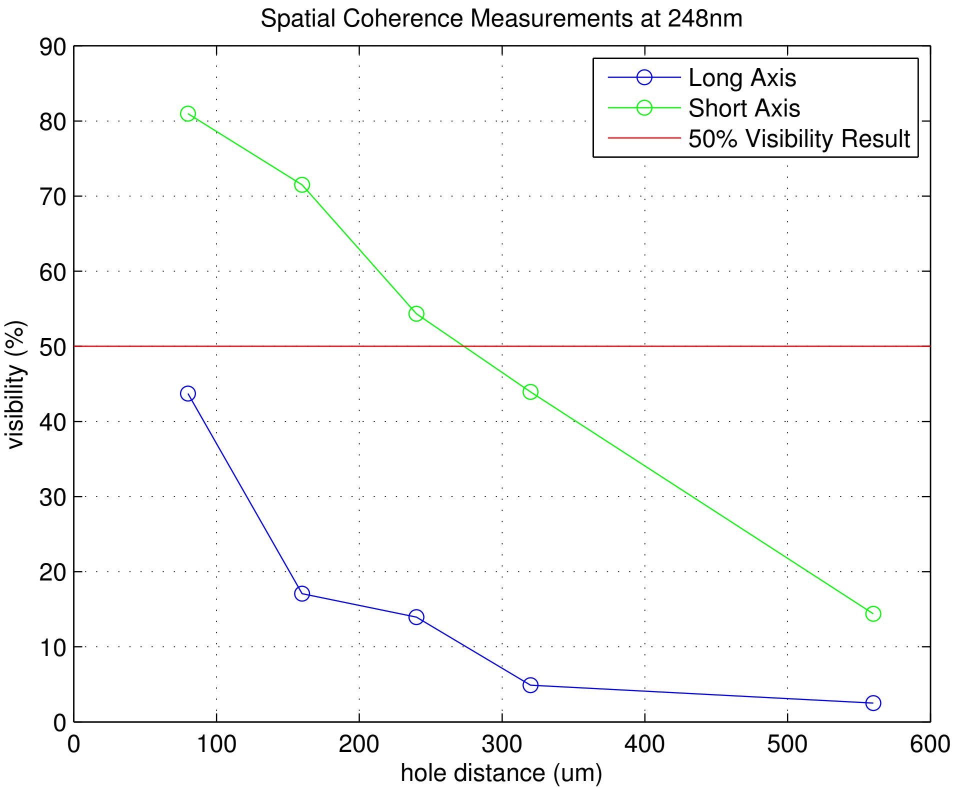

Figure 4.18: Plot of average visibility (%) as a function of pinhole separation (μm) for the long and short axis of a 248 nm source. The pinholes used in the measurements were 10 μm and 15 μm in diameter. Overall, the trend seen here is similar to the results for the 193 nm beam (see Figure 4.11), with a higher degree of coherence present along the short axis of the beam.

4.3.3 The Coherence Distribution

A Gaussian fit was applied to the coherence distribution data for both the 193 *nm* and 248 *nm* sources. In both cases, the fit was applied to the experimental data using the curve fitting toolbox in Matlab. The Gaussian curve was not constrained to pass through the maximum visibility point. The reason for this is that the maximum visibility was not measured experimentally. In theory, a maximum visibility measurement of 100 % is achievable for two overlapping pinholes illuminated with partially coherent light. If the pinholes overlap completely, the incident light passes through the overlying pinholes and no interference occurs. The visibility in this case is, in theory, 100 %. However, we can assume that if we were to have completed this measurement in the laboratory, it would be subject to the same error as the other measurements taken. Therefore, it is inaccurate to constrain the Gaussian fit to pass through this point.

Coherence Distribution at 193nm

A Gaussian distribution was applied to the coherence distribution data gathered from the 193 *nm* source. In Figure 4.19 (a), the original coherence distribution data for the short axis of the 193 *nm* source is shown as the starred data points. In Figure 4.19 (b), the starred data points refer to the original coherence distribution data for the long axis of the 193 *nm* source. In both plot (a) and (b), the Gaussian fit is shown as a solid red line.

The Root Mean Square (RMS) error for the Gaussian fit applied to the data taken for the short axis of the 193 *nm* beam (Figure 4.19 (a)) is 12.24 %. For the fit applied to the data corresponding to the long axis of the beam (Figure 4.19 (b)), the RMS error is 7.20 %.

Coherence Distribution at 248nm

Similarly, a Gaussian fit was applied to the coherence distribution data gathered from the 248 *nm* source. Figure 4.20 (a) shows the measured coherence distribution data for the short axis of the source (starred data points) and the best Gaussian fit is given as a solid red line. The coherence distribution data for the long axis of the source beam (starred data points) and corresponding Gaussian fit (solid red line) is plotted in Figure 4.20 (b).

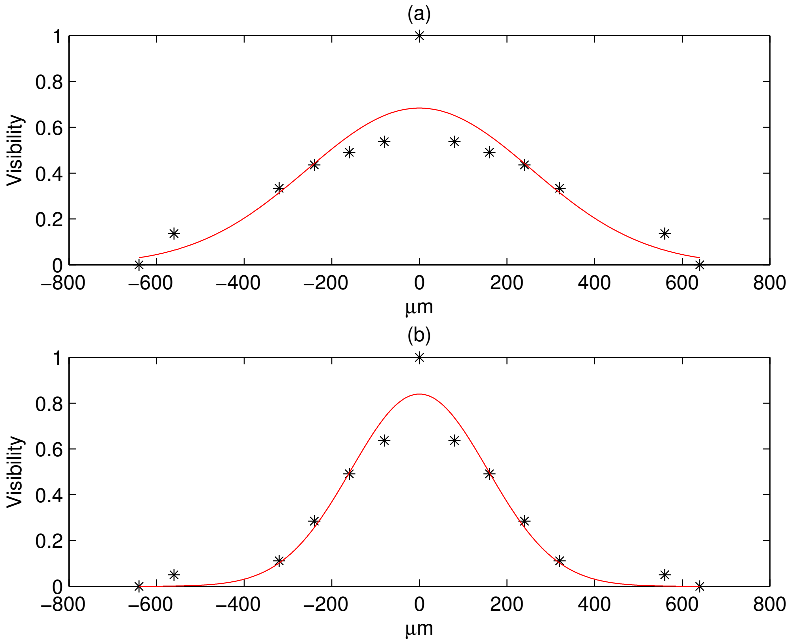

Figure 4.19: Best Gaussian fit of coherence distribution data across the (a) short and (b) long axes of the 193 *nm* source.

The Root Mean Square (RMS) error for the Gaussian fit applied to the data taken for the short axis of the 248 *nm* beam (Figure 4.20 (a)) is 5.47 %. For the fit applied to the data corresponding to the long axis of the beam (Figure 4.20 (b)), the RMS error is 8.53 %.

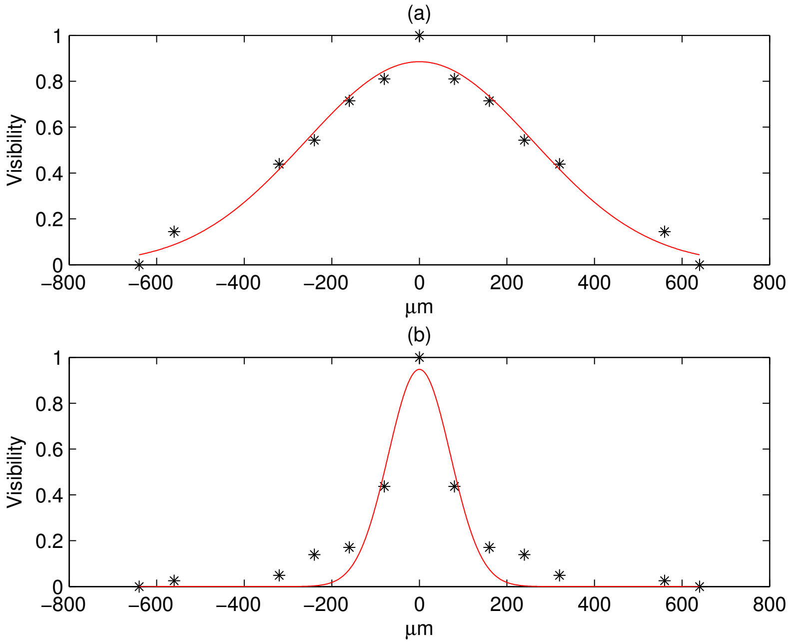

Figure 4.20: Best Gaussian fit of coherence distribution data across the (a) short and (b) long axes of the 248 *nm* source.

4.3.4 Error Discussion

The error in the measurements and subsequent data processing is difficult to quantify. Experiments were carried out over two visits to DIMO in May and June 2010. It is necessary to take the measurements manually as the optimum position must be found for each pinhole size and orientation. The measurements are carried out on the raw laser beam, which has a different intensity distribution in the horizontal and vertical axes; the beam is not uniform in intensity. Thus for each series of measurements, the position of the pinhole array was adjusted manually to be as close to the centre of the beam as possible, as this is the area of the beam that is used in the subsequent imaging system. If it were possible to automate the data collection process, several sets of measurements could be taken to reduce the experimental error. However, this is not feasible here.

Some error will also occur in the processing of the data. A line of pixels is user-selected, and their values are averaged along with 10 adjacent lines of pixel values on either side. There will be some error present in the line selection. It is desirable to select a line of pixels that is perpendicular to the interference lines observed in the diffraction envelope, and an angular deviation from the ideal perpendicular line may cause error in the subsequent data analysis.

However, taking these sources of error into account, the coherence results presented here are in line with what we expect from partially spatially coherent excimer sources. In the next chapter, we will apply this coherence data to the numerical model of an excimer beam using the Elementary Function Method.

Chapter 5

Laser Beam Modelling

In the previous chapter, spatial coherence was measured for two different UV excimer sources. These spatial coherence measurements, along with those of the intensity profile of the beam, are the key inputs to a real beam model. In this chapter, intensity profile measurements of the 193 *nm* and 248 *nm* sources are presented (see Section 5.1). Next, in Section 5.2, we apply the Elementary Function method to a 1-D Gaussian Schell-model beam, noting the shape of the cross-spectral density, elementary function and weighting function. We also discuss the sampling criterion. In Section 5.3, the spatial coherence and beam profile measurements are introduced to the model, and the elementary function for the real 248 *nm* beam is generated. Finally, in Section 5.4, the method is applied to a simple edge-imaging problem, and compared with the theoretical results from spatially coherent and incoherent sources.

5.1 Beam Intensity Profile Measurements

Measurements of intensity in the beam profile were taken for both the 193 *nm* and 248 *nm* sources. In both cases, a number of frames were averaged, and a cross-section taken through the center of the averaged intensity plot to give the average 1-D intensity profile for the long and short axes of the beam. Some further processing was necessary for the data from the Braggstar 248 *nm* source, including the use of a spatial filter to smooth out the profile.

Measurements at 193nm

Some initial beam profile measurements were taken of the 193 *nm* source using an Ophir Beamstar FX laser beam profiler. Figure 5.1 shows the average of 8 beam profile images. Each image was made up of 340 × 240 pixels, and the pixel size was 0.0074 × 0.0074 *mm* square (full frame dimension 2.368 × 1.776 *mm*).

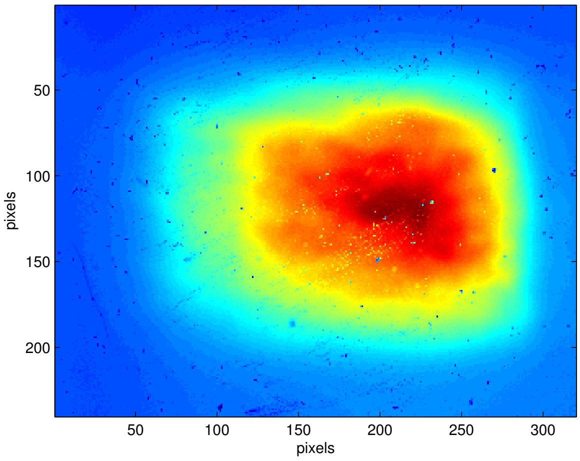

Figure 5.1: Averaged intensity profile image for the 193 *nm* ATLex source beam.

Figure 5.2 is the normalised intensity profile for (a) the long axis, and (b) the short axis of the 193 *nm* beam. The intensity profile for the short axis of the beam bears some resemblance to a broad Gaussian distribution. However, the long axis deviates from the expected top-hat intensity profile. The poor uniformity in the intensity profile means that this raw laser beam is unsuitable for lithographic applications unless some reshaping and homogenizing optics are applied to the beam. This laser is mostly used for laser ablation processes, where the surface area of the beam in use is small and the most suitable portion of the beam can be chosen. In reality, one can assume that the results in this case are not optimum and a more uniform profile is desirable.

Measurements at 248nm

Beam profiling measurements were carried out on the 248 *nm* Braggstar source using a Newport LBP-1-USB Laser Beam Profiler. The data was saved as two .avi (Audio Video Interleave) files, and the individual frames were extracted using Matlab. The source operates in pulsed mode, resulting in some 'blank' frames recorded between

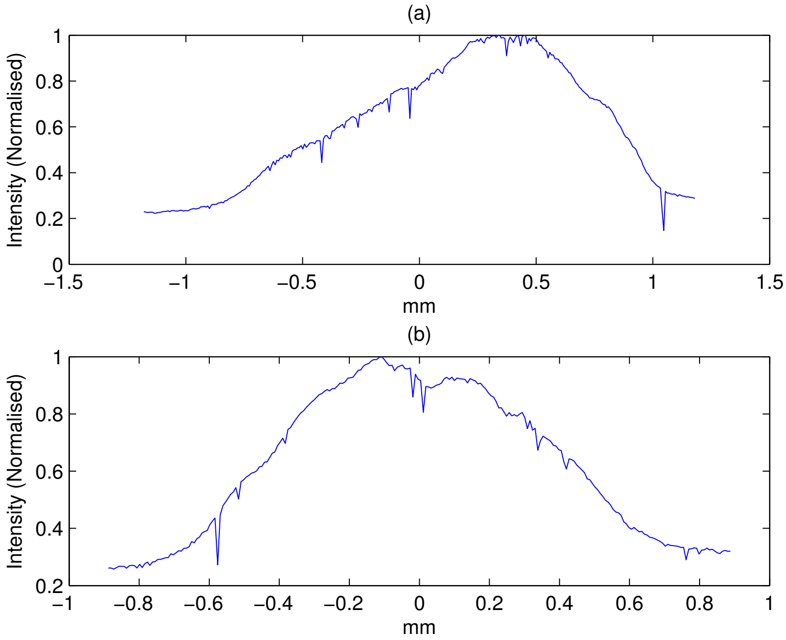

Figure 5.2: Intensity cross-section (normalised) of the (a) long and (b) short axis of 193 *nm* ATLex source. Several discontinuities are visible across the intensity profile, most likely a result of dust on the sensor.

pulses. A total of 187 frames were averaged (47 frames from the first video clip, and 140 frames from the second clip). The averaged imaged is shown in Figure 5.3.

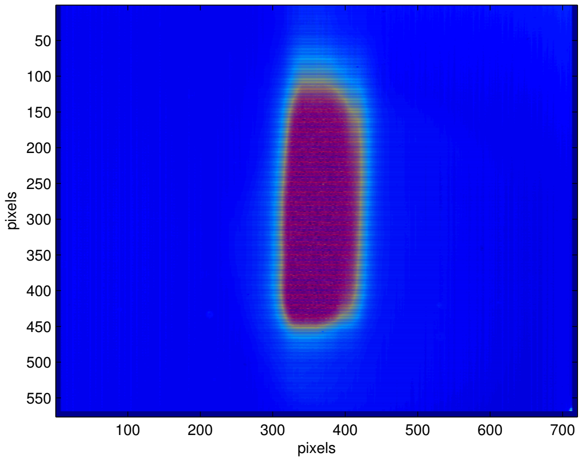

Figure 5.3: Average of 187 beam profile images of the 248 *nm* Braggstar source. Artifacts in the form of horizontal lines are visible in the image as a result of interlacing.

Each frame grabbed from the .avi file was 720 × 576 pixels. The short axis of the beam was oriented along the 720 pixel side, and with each pixel corresponding to 9.0595 μm, this represented 6.52 *mm*. The profile of the long axis of the beam comprised 576 pixels, each measuring 8.3 μm, corresponding to 4.78 *mm*.

An artifact is visible in Figure 5.3 in the form of horizontal lines in the image. This artifact seems to be the result of 'interlacing'. Normal videos record two frames at 50 *Hz* and they are then interlaced to produce the final 25 *Hz* image. The .avi files saved by the beam profiling software appears to be missing one of the images. Figure 5.4 shows the cross-section of the long axis of the averaged raw beam profile image from Figure 5.3. The missing data creates a series of low values across the profile representing the background level/dark noise in the ccd. To overcome this issue, the missing data was filled in by interpolation and extrapolation of the missing data points using the nearest neighbour method in the curve fitting toolbox in Matlab.

The background data was removed by defining four 21 × 21 arrays of intensity values taken from four different areas around the beam in the raw images. The average background intensity was subtracted from the intensity cross-section profiles for the long and short axes. Figure 5.5 shows the normalised intensity cross-section (background removed) for (a) the short axis and (b) the long axis of the 248 *nm* Braggstar

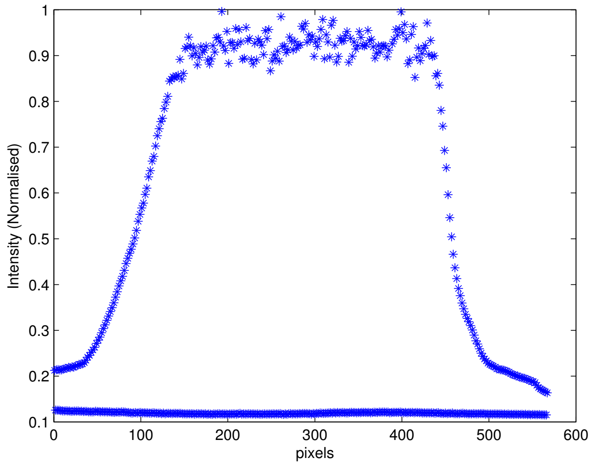

Figure 5.4: Intensity cross-section (normalised) of the long (y) axis of the raw 248 *nm* Braggstar beam

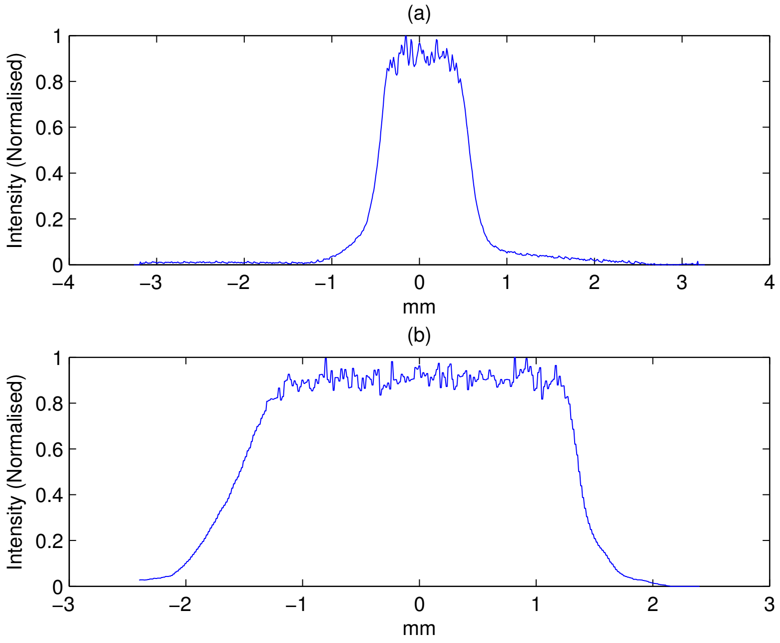

Figure 5.5: Intensity cross-section (normalised) of the (a) short and (b) long axis of 248 *nm* Braggstar source including interpolated data.

beam. In both cases, the intensity profile is a similar shape to that expected from an excimer laser beam: the intensity profile along the short axis of the beam is a Gaussian distribution, and the expected top-hat distribution of intensity is seen along the long axis of the beam.

To use this beam profile data in the system model, the noisy intensity curves must be smoothed. The intensity data for the long and short axes of the beam were smoothed by applying a Savitzky-Golay filter. Savitzky-Golay smoothing filters (also called digital smoothing polynomial filters or least-squares smoothing filters) are typically used to smooth out a noisy signal whose frequency span (without noise) is large. The Savitzky-Golay filter can be thought of as a generalized moving average which preserves features of the distribution such as relative maxima, minima and width, which are often distorted by other averaging techniques. The filter coefficients are derived by performing an unweighted linear least-squares fit using a polynomial of a given degree. The smoothing is computed using the algorithm described by Savitzky and Golay [88]. The algorithm computes a local polynomial regression on the input data by solving the following equation:

$$y(t) = a_0 + a_1 t + a_2 t^2 + \dots \tag{5.1}$$

up to order R. That is

$$y(t) = \sum_{r=0}^{R} a_r t^r \tag{5.2}$$

In general, higher degree polynomials can more accurately capture the heights and widths of narrow peaks, but can do poorly at smoothing wider peaks. In this data set, we are dealing with one wide "peak" in the long and short axis, so a second order polynomial was sufficient to smooth the curves. Figure 5.6 shows a comparison between the raw beam profile data (solid red line) and the filtered data (solid black line) for (a) the short axis and (b) the long axis of the 248 *nm* source. In both cases the raw data has been smoothed to give a curve with less noise which is more easily modelled.

The Elementary Function method involves a series of Fourier transforms which require that the intensity profile of the beam is centered correctly. It is also necessary that the intensity profile is symmetrical: an assumption of the Elementary Function theory is that the cross spectral density, and thus the elementary function, is real and even. In order to symmetrize the beam, the centroid of the smoothed data was found using the following formula:

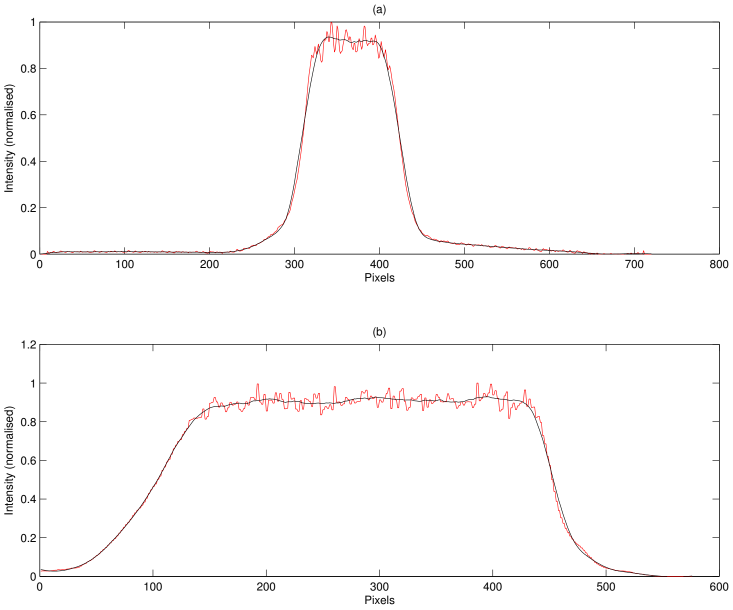

Figure 5.6: Comparison of raw beam profile data (red) and data smoothed using a Savitzky-Golay filter (black) for (a) short and (b) long axis of 248 *nm* Braggstar source. In both cases, the extrapolated data was used (background removed).

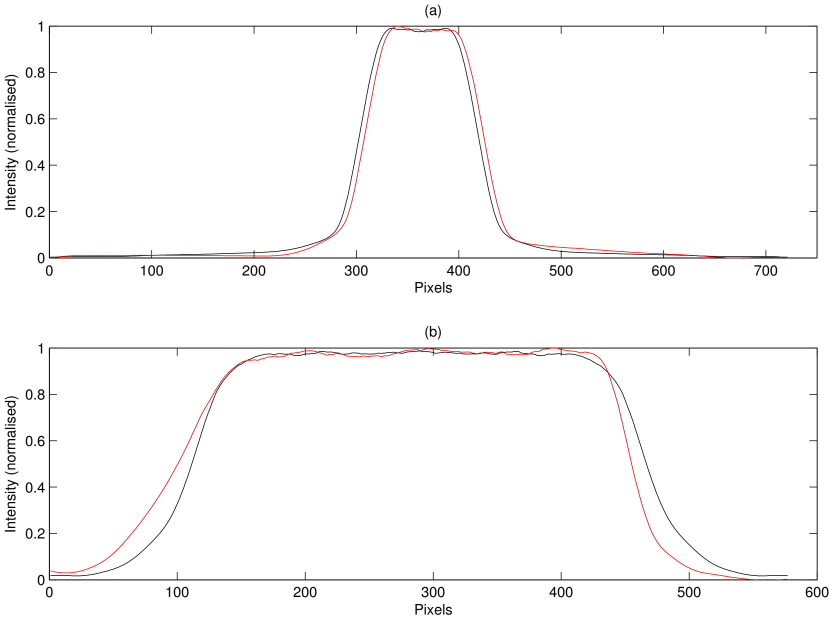

Figure 5.7: Comparison of centered, symmetrized beam profiles and smoothed beam profile data for (a) short and (b) long axis of 248 *nm* Braggstar source. The original smoothed data is represented by the solid red line; the symmetrized beam is given by the solid black line.

$$C = \frac{\sum_n c_n x_n}{\sum_n c_n}, \tag{5.3}$$

where C is the centroid of a set of n elements. Each element is represented by c_n, and its position is denoted as x_n. Once the location of the centroid was found, the beam intensity profile was symmetrized and centered about that point. A comparison of the original smoothed profile data and the symmetrized, centred data is presented in Figure 5.7.

5.2 Elementary Function Method: Gaussian Schell-Model Beam

The Elementary Function model was tested in the first instance for a Gaussian Schell-model beam. We follow the method referred to in Section 3.4.2, where we avoid 4D calculations by going directly to the 2D cross-spectral density, $\widehat{W}_0(\mathbf{u}, -\mathbf{u})$. For clarity, we recall Equation 3.63:

$$\widehat{W}_0(\mathbf{u}, -\mathbf{u}) = \iint_\infty d^2u' \hat{U}^2(\mathbf{u} - \mathbf{u}')\hat{\gamma}(\mathbf{u}') \tag{5.4}$$

where $\hat{U}(\mathbf{u})$ and $\hat{\gamma}(\mathbf{u})$ are the Fourier transforms of $U(\mathbf{r})$ (the intensity distribution) and $\gamma(\mathbf{r})$ (the coherence distribution) respectively. For this test case, both the intensity and the coherence have a Gaussian distribution. To find the cross-spectral density, the Fourier transform of the intensity distribution is calculated, and then squared to give $\hat{U}^2$. A Fourier transform is applied to the coherence distribution to give $\hat{\gamma}$. These two quantities are convolved to give $\widehat{W}_0(\mathbf{u}, -\mathbf{u})$.

Referring to Equation 3.51, the Elementary Function can be found from $\widehat{W}_0(\mathbf{u}, -\mathbf{u})$ using the following relation

$$\hat{f}(\mathbf{u}) \propto \sqrt{\widehat{W}_0(\mathbf{u}, -\mathbf{u})}. \tag{5.5}$$

The intensity distribution, the cross-spectral density, and the generated elementary function are shown in Figure 5.8.

Next, using the sampling criterion outlined in [65], we calculate the number of functions required to represent the source. The elementary functions are, in general,

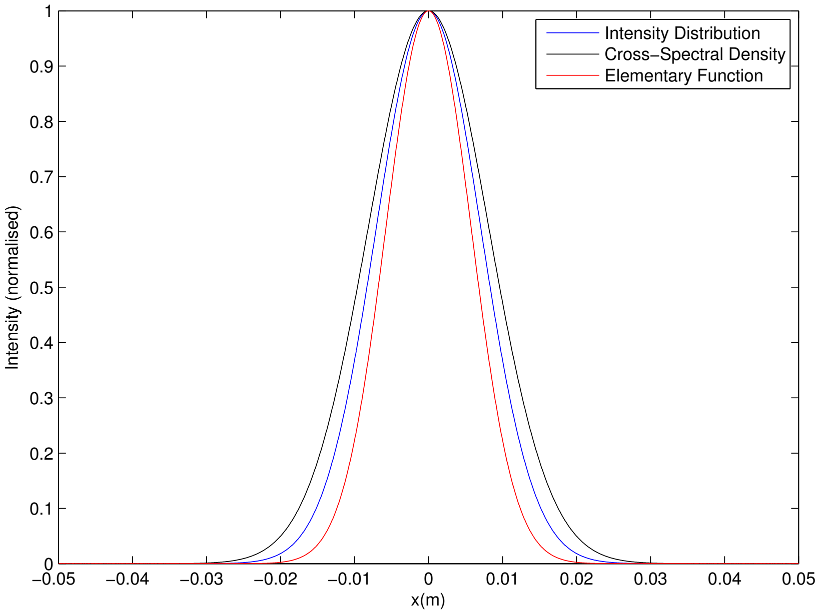

Figure 5.8: Comparison of the intensity distribution, cross-spectral density, and elementary function for a Gaussian Schell-model beam.

not orthogonal, and thus finding the expansion coefficients becomes more difficult than for orthogonal sets of functions. Unser [81] presented a sampling and interpolation theory for nonorthogonal basis sets, which we use to find a sampling criterion and to retrieve the values of the coefficients. The traditional Shannon-Whittaker sampling criterion cannot be applied, since it implies that sinc functions of appropriate width will be used as interpolation functions. In our case the interpolating function (i.e. the elementary function) will not be a sinc function, and its width will be different from that assumed in traditional sampling and interpolation.

We can write the intensity in terms of the cross spectral density

$$I_0(\mathbf{r}) = W_0(\mathbf{r},\mathbf{r}) = \sum_m \sum_n a_{mn} f^2(\mathbf{r} - \mathbf{r}_{mn}), \tag{5.6}$$

and, from this, define our basis functions

$$\varphi_{mn}(\tilde{\mathbf{r}}) = D f^2(\Delta x(\tilde{x} - m), \Delta y(\tilde{y} - n)). \tag{5.7}$$

The new coordinates $\tilde{\mathbf{r}} = (\tilde{x}, \tilde{y}) = x/\Delta x, y/\Delta y)$ ensure that the function is sampled at integer values, and the constant D ensures that $\varphi(0) = 1$. These changes allow us to use the three conditions for expansion presented by Unser. First, the sequence of coefficients must be square-integrable, i.e.

$$\|a\|^2 = \sum_m \sum_n |a_{mn}|^2 \tag{5.8}$$

must exist. From Equation 5.6, we are aware of the dependence of a_{mn} on the intensity, $I_0(\mathbf{r})$ and the square of the elementary function $f^2(\mathbf{r})$. If we consider just one dimension, and rewrite Equation 5.6 referring to the sampling distance, Δx, we get

$$\sum_m |a_m|^2 = \frac{1}{\Delta x} \int_{-1/2\Delta x}^{1/2\Delta x} du \left| \frac{\hat{I}_0(u)}{\widehat{f^2}(u)} \right|^2. \tag{5.9}$$

This expression exists if $\hat{I}_0(u)/\hat{f^2}(u)$ is square integrable on $[-1/2\Delta x, 1/2\Delta x]$, which is fulfilled as long as the function has no singularities, i.e., as long as $\hat{f^2}(u)$ has no zeros in the interval. This has already been assumed. Thus Equation 5.8 is fulfilled for any function of interest.

Second, the set of basis functions needs to form a Riesz basis. That means that the condition

$$A \cdot \|a\|^2 \leq \left\| \sum_m \sum_n a_{mn} \varphi_{mn}(\mathbf{r}) \right\|^2 \leq B \cdot ||a||^2, \tag{5.10}$$

where $A > 0$ and $B < \infty$, must be fulfilled. Note that if $A = B = 1$, then the basis functions are orthonormal. The first inequality implies that the functions are linearly independent, while the second inequality holds for any physical elementary functions [65]. The third condition is the most demanding: the partition of unity condition:

$$\sum_n \sum_m \varphi(\tilde{x} + m, \tilde{y} + n) = 1, \tag{5.11}$$

for all real $(\tilde{x}, \tilde{y})$. In the Fourier domain it translates as

$$\varphi(m,n) = \delta_m \delta_n, \tag{5.12}$$

where the difference from Equation 5.6 originates from different definitions of the Fourier transform and

$$\delta_m = \begin{cases} 1 : m = 0 \\ 0 : m \neq 0 \end{cases}, \tag{5.13}$$

where m is an integer. This condition affects how closely a function can be reproduced, by making the sampling step sufficiently small. Our elementary functions do not fulfill this criterion exactly, so we are forced to accept an approximate version of the partition of unity condition. We get a good idea of the sampling required if we use the condition that

$$\hat{\varphi}(\mathbf{0}) = 1, \tag{5.14}$$

while

$$\hat{\varphi}(1,1) < c, \tag{5.15}$$

where $c \leq 0$ is a small value considered negligible in the context and we have assumed that the function is positive with a single maximum at $(\tilde{x}, \tilde{y}) = (0,0)$. Changing notation from $\varphi(\tilde{\mathbf{r}})$ to $f^2(\mathbf{r})$ changes the conditions into

$$\hat{f^2} = 1/D \tag{5.16}$$

and

$$\hat{f^2}\left(\frac{1}{\Delta x}, \frac{1}{\Delta y}\right) < c/D \tag{5.17}$$

Equation 5.17 is used to determine the sampling intervals in Δx and Δy. The sampling distance, Δx, is given as

$$\Delta x = \frac{\pi \sigma_I \sigma_g}{\sqrt{(-\ln c)(\sigma_g^2 + 4\sigma_I^2)}}, \tag{5.18}$$

where σ_I represents the FWHM of the intensity distribution, and σ_g is the width of the coherence distribution. The value given to c depends on what we consider to be the full width of the Gaussian distribution. Common practice when handling a Gaussian distribution $exp(-x^2/\sigma^2)$ is often to consider values over a width of 6σ, i.e., for $|x| \leq 3\sigma$, while values for larger $|x|$ are considered to be zero. This corresponds to using $c = exp(-9) \approx 1.23 \times 10^{-4}$. The number of elementary functions to be propagated is calculated by dividing the value for Δx into the full width of the intensity distribution (i.e., $6\sigma_I$). For this example case, we choose the values for σ_I and σ_g as 0.01 m and 0.02 m respectively. These values produce a Δx of 0.0074 m which corresponds to 9 elementary functions. Starting with an elementary function positioned at the centre of the source, we position 4 functions either side, each centered a distance of Δx away from the previous function. The weighting coefficients are calculated as a distribution, which we sample in the same way. The weighting distribution is calculated as follows

$$\sum_m a_m = \Delta x \frac{\hat{U}_0(0)}{\hat{f}^2(0)} \tag{5.19}$$

Each shifted elementary function is squared and then weighted to give the reconstructed intensity distribution $U_r(\mathbf{r})$:

$$U_r(\mathbf{r}) = f^2(\mathbf{r}) \sum_m a_m \tag{5.20}$$

A plot of the 9 shifted, squared and weighted elementary functions required for this Gaussian Schell-model beam example is shown in Figure 5.9. The sum of the weighted functions is compared to the original intensity distribution in Figure 5.10; the two are almost identical.

In Figure 5.11, the numerical results of the Elementary Function Method are compared to the analytical results previously presented in Figure 3.2. Figure 5.11 (a), (b) and (c) are calculated numerically using the same coherence width and intensity width as in the analytic calculations shown in (d), (e), and (f) respectively. In each case, the numerical result bears a close resemblance to the equivalent analytic result below it. In (d), the calculated weighting function has some negative elements which drives the two second-to-largest shifted functions negative. In the numeri-

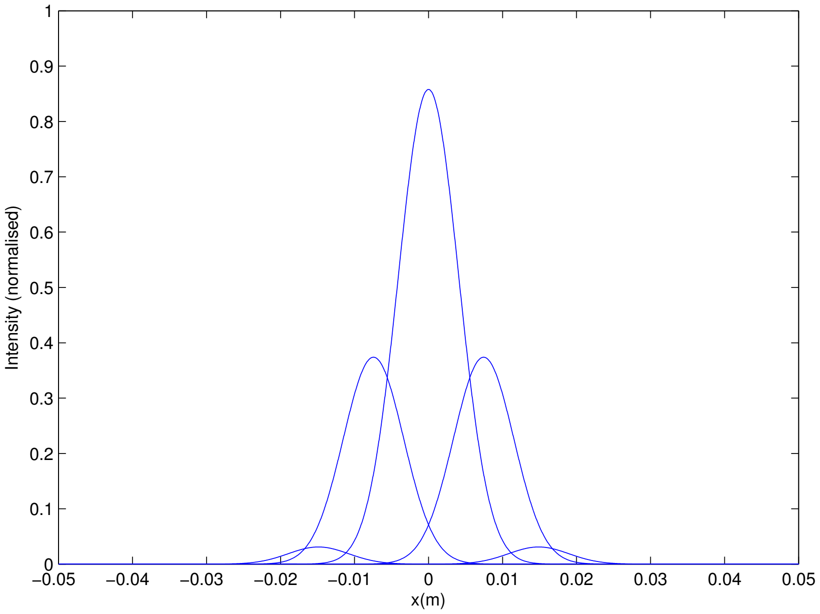

Figure 5.9: Shifted and weighted elementary functions for a Gaussian Schell-model source.

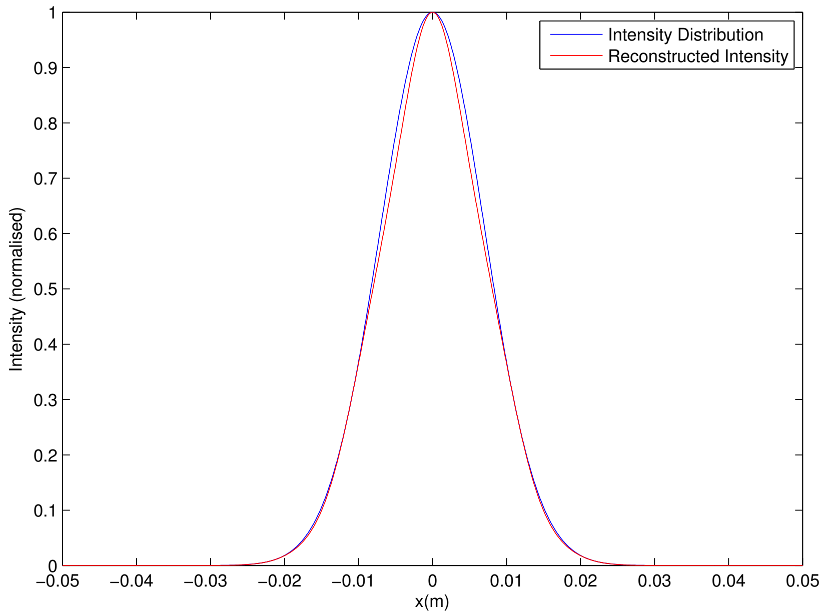

Figure 5.10: Comparison of original intensity distribution and the reconstructed intensity distribution for a Gaussian Schell-model source. The two plots are almost identical.

cal case, the weighting coefficients are positive, but the reconstructed intensity still closely matches the original intensity profile.

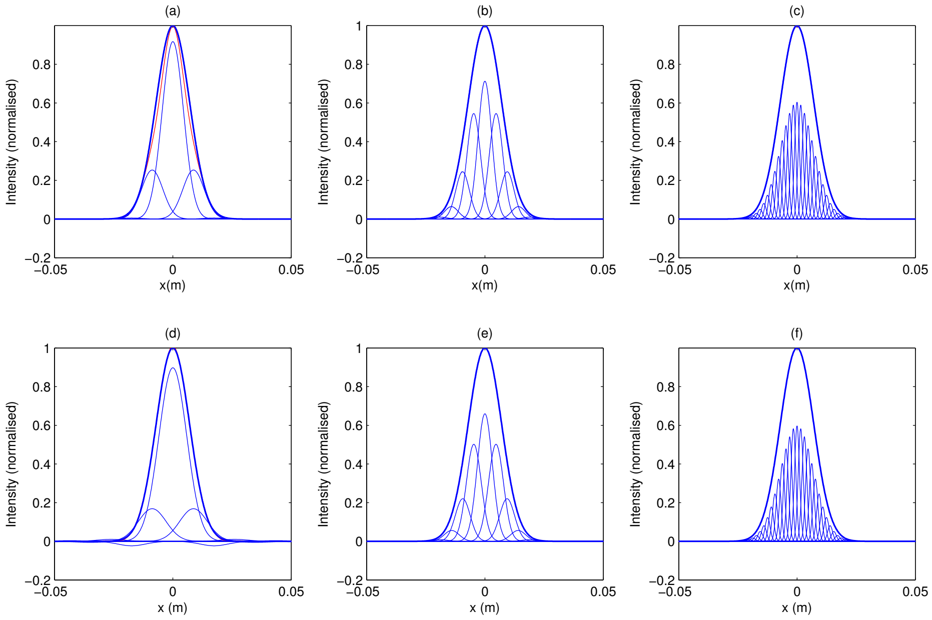

Figure 5.11: Comparison of analytical and numerical results for the Elementary Function Method applied to a Gaussian Schell-model beam. (a)-(c) Intensity (thick blue line) and reconstructed intensity (red line) are shown for 1-D Gaussian Schell-model beams of different degrees of coherence, along with the scaled and shifted squares of the elementary functions (thin lines), as calculated analytically (as seen in Figure 3.2). (d) - (f) Intensity and reconstructed intensity (thick line) are shown for 1-D Gaussian Schell-model beams of different degrees of coherence, along with the scaled and shifted squares of the elementary functions (thin blue lines), as calculated numerically. The intensity distribution is the same for all beams, with $\sigma_I = 0.01\ m$, while the coherence varies from high to low as (a), (d) $\sigma_g = 0.03\ m$ with 7 elementary functions required, (b), (e) $\sigma_g = 0.01\ m$ with 13 elementary functions required, and (c), (f) $\sigma_g = 0.003\ m$ with 39 elementary functions required.

As we discussed earlier in this section, we have considered the full width of the Gaussian intensity distribution to extend over a width of $6\sigma_I$. In practice, once the weighting distribution is applied to the squared and shifted elementary functions, the amplitude of many of the functions at the edge of the distribution drops to almost zero; in the reconstructed intensity distribution, the main contribution comes from the elementary functions placed close to the centre. To test the effect of decreasing the defined width of the intensity distribution to $4\sigma_I$, we consider the three cases in

Figure 5.11. The results are shown in Figure 5.12. In each case, the intensity distribution is the same, with $\sigma_I = 0.01\ m$, while the coherence varies as (a) $\sigma_g = 0.03\ m$, (b) $\sigma_g = 0.01\ m$, and (c) $\sigma_g = 0.003\ m$. The reconstructed intensity for a Gaussian width of $6\sigma_I$ is given by the blue line, the red line represents the $4\sigma_I$ case.

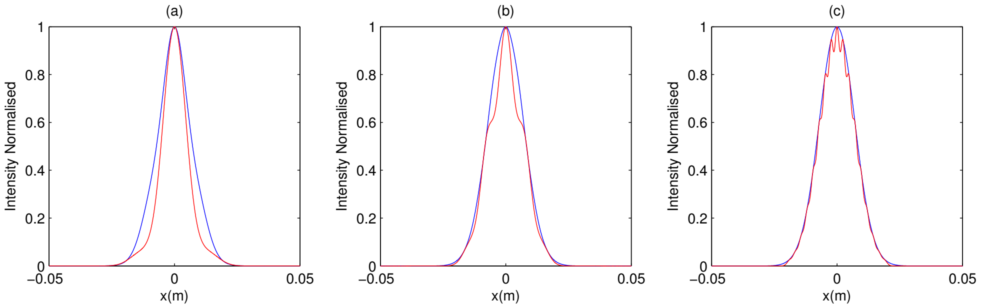

Figure 5.12: Comparison of the reconstructed intensity using the Elementary Function model for different Gaussian distribution widths. In each case, the intensity distribution is the same, with $\sigma_I = 0.01\ m$, while the coherence varies from high to low as (a) $\sigma_g = 0.03\ m$, (b) $\sigma_g = 0.01\ m$, and (c) $\sigma_g = 0.003\ m$. The reconstructed intensity for a Gaussian width of $6\sigma_I$ is given by the blue line, the red line represents the $4\sigma_I$ case.

In all three examples, when the defined Gaussian width was reduced to $4\sigma_I$, the reconstructed intensity deviated from that produced in the $6\sigma_I$ case. From this we can assume that using the $6\sigma_I$ definition of the Gaussian distribution width is necessary to ensure the shifted and weighted elementary functions sufficiently reproduce the intensity profile.

5.3 Elementary Function Method: 248 nm Excimer Laser Beam

The next step was to apply the Elementary Function model to a real beam; in this case we use the measured intensity profiles and coherence distributions of the long and short axes of the 248 nm Braggstar source. As in Section 5.2, we go directly to the 2-D cross-spectral density, $\widehat{W}_0(\mathbf{u}, -\mathbf{u})$, and following the same series of steps, we arrive at the elementary function, as shown in Figure 5.13.

Initially, we assumed that the sampling criterion outlined in the previous section for the Gaussian Schell-model beam is sufficient for this more "Super-Gaussian" intensity distribution. However, this resulted in under-sampling. In Figure 5.14 we see

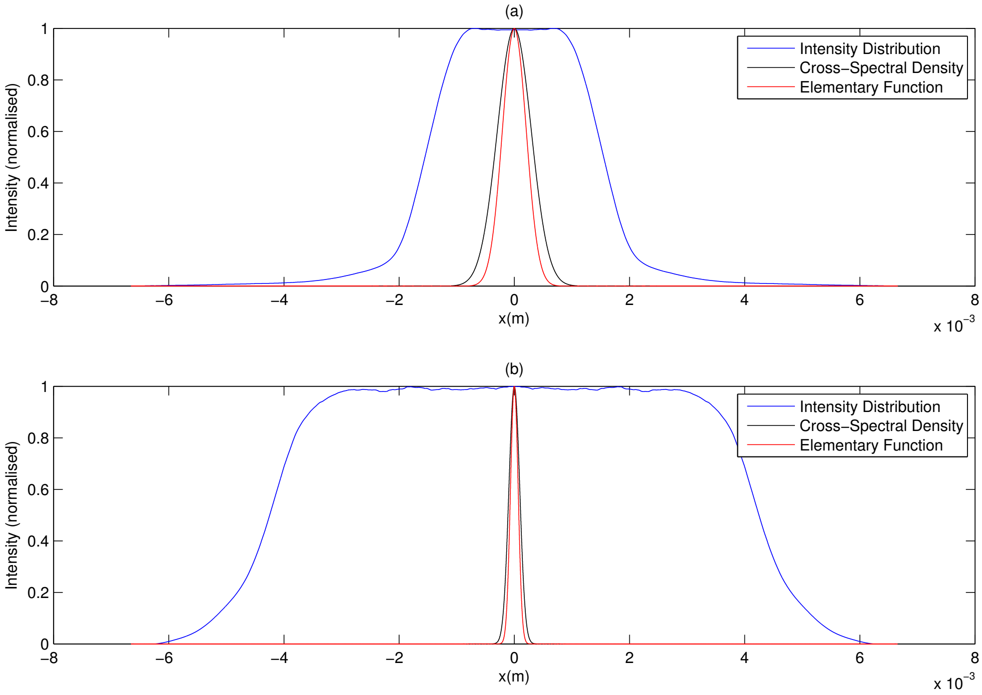

Figure 5.13: Comparison of original intensity distribution, cross-spectral density, and elementary function for the (a) short and (b) long axis of the 248 *nm* Braggstar source.

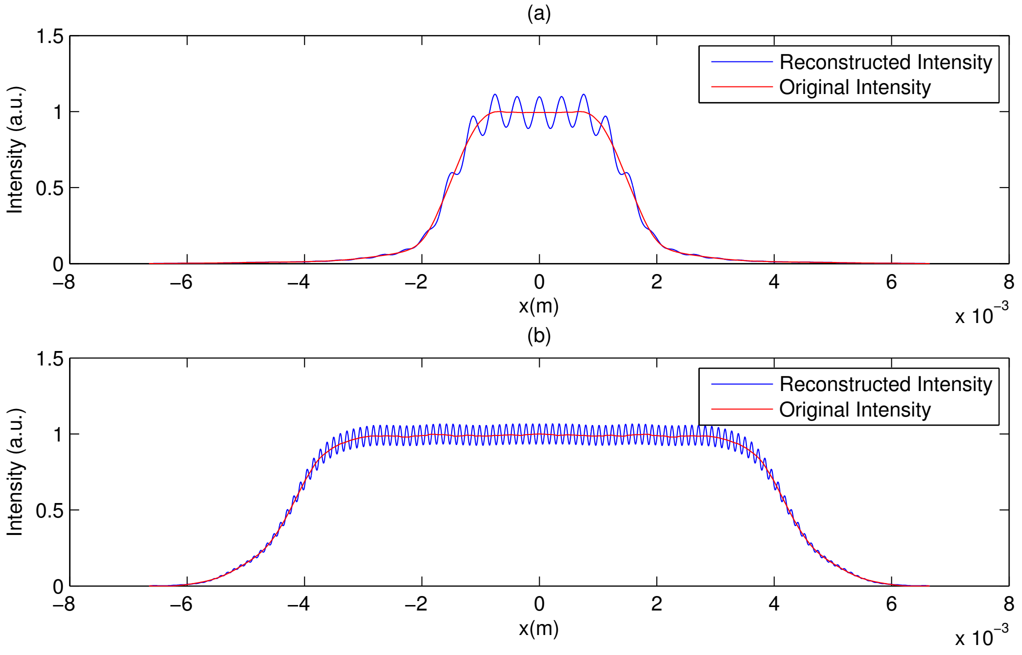

Figure 5.14: Comparison of original intensity distribution and the reconstructed intensity distribution for the (a) short and (b) long axis of the 248 *nm* Braggstar source using the sampling criterion for the Gaussian Schell-Model beam. This undersampling produces a noisy reconstructed intensity in both axes as a result of insufficient overlap of the elementary functions.

this under-sampling manifests itself as a noisy intensity distribution for the reconstructed beam. The width of the elementary function is small, and the sampling predicted by the Gaussian Schell-model case results in an insufficient overlap between adjacent functions. The top of each shifted and weighted elementary function is visible in the reconstructed intensity plots in Figure 5.14. Thus a larger overlap than predicted by the Gaussian Schell-model is needed to accurately sample the beam. We return to the initial sampling criterion outlined in [65] and, using Equation 5.17, the sampling distance Δx can be found. The weighting distribution, $a(\mathbf{r})$, is calculated by evaluating the following relation and taking the inverse Fourier transform:

$$\hat{a}(\mathbf{u}) = \frac{\widehat{U_0}(\mathbf{u})}{\widehat{f^2}(\mathbf{u})}, \tag{5.21}$$

where $\widehat{U_0}(\mathbf{u})$ is the Fourier transform of the intensity distribution, $\widehat{f^2}(\mathbf{u})$ is the Fourier transform of the square of the elementary function, and $\hat{a}(\mathbf{u})$ is the Fourier transform of the weighting distribution. Due to the deviation of the original intensity profiles from a smooth Gaussian envelope, the calculated weighting is a noisier distribution than before. The Fourier transform of the weighting distribution for the long and short axes is given in Figure 5.15. For both the long and short axes of the beam, high frequency noise is present. Carrying out an inverse Fourier transform produces the weighting distribution, which is then sampled and applied to the shifted and squared elementary functions.

The high frequency noise present in the weighting distribution, as seen in the plots of the Fourier transform of the distribution (Figure 5.15), carries through when applied to the shifted and squared elementary functions and, as a result, the reconstructed intensity is noisy. It is necessary to reduce the noise in the weighting calculation to produce a smoother reconstructed intensity.

Dealing with Noise

When faced with the issue of reducing or removing noise in an imaging system, we have a few options. An *inverse filter* can be used as a way to work back to the original system input, but the results are not always reliable, as we will discuss below. Another option available is the *Wiener filter*. This is the mean square error linear filter for images degraded by additive noise and blurring. First proposed by Wiener in 1949, its purpose is to reduce the amount of noise present in a signal by comparison with

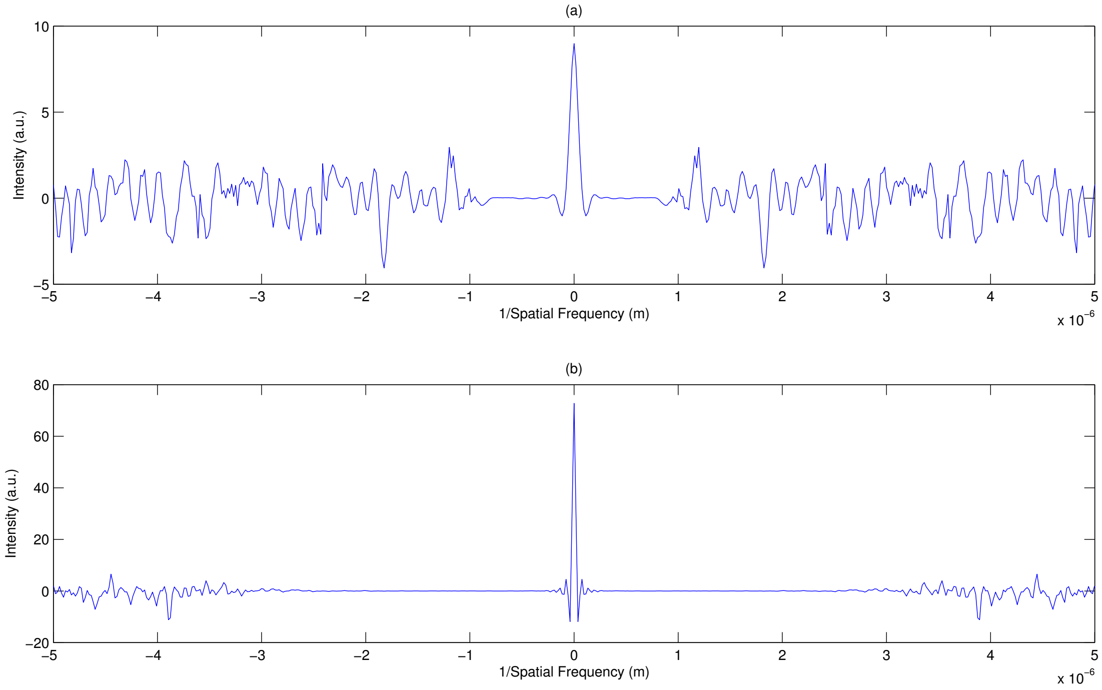

Figure 5.15: Plot of the Fourier transform of the weighting distribution calculated for the (a) short and (b) long axis of the 248 *nm* Braggstar source. Some high frequency noise is evident.

an estimation of the desired noiseless signal. Calculation of the Wiener filter requires the assumption that the signal and noise processes are second-order stationary. We discuss these filters below, with notation taken from [89].

Consider a simple imaging system:

$$g(x,y) = \int\int f(x',y')h(x-x',y-y')dx'dy' + n(x,y) \tag{5.22}$$

where $g(x,y)$ is the image (or output), $h(x,y)$ is the impulse response (PSF) of the system, and $f(x',y')$ is the quantity we wish to restore. For a real system, we must also include an additive noise term, $n(x,y)$. If we Fourier transform both sides of this equation, we get

$$G(k_x,k_y) = F(k_x,k_y)H(k_x,k_y) + N(k_x,k_y) \tag{5.23}$$

If we have an ideal situation where the additive noise term $n(x)$ is negligible, then

$$F(k_x,k_y) = \frac{G(k_x,k_y)}{H(k_x,k_y)} = Y(k_x,k_y)G(k_x,k_y) \tag{5.24}$$

and thus

$$f(x,y) = F^{-1}\{Y(k_x,k_y)G(k_x,k_y)\} \tag{5.25}$$

whereF^{-1} denotes the inverse Fourier transform. The frequency domain filters

$$Y(k_x,k_y) = \frac{1}{H(k_x,k_y)}, \tag{5.26}$$

where $H(k_x,k_y)$ is the system optical transfer function (OTF), is called the *inverse filter*. In practice, the straight inverse filter rarely works satisfactorily, because if the OTF of the system drops to zero, the magnitude of the inverse filter, $Y(k_x,k_y)$, will approach infinity. Furthermore, if we apply the inverse filter in the case where noise is present, we get

$$\begin{aligned} \hat{F}(k_x,k_y) = Y(k_x,k_y)G(k_x,k_y) &= \frac{G(k_x,k_y)}{H(k_x,k_y)} + \frac{N(k_x,k_y)}{H(k_x,k_y)} \\ &= F(k_x,k_y) + \frac{N(k_x,k_y)}{H(k_x,k_y)} \end{aligned} \tag{5.27}$$

where the ‘hat’ notation denotes an estimated quantity. The recovered frequency spectrum has an additional term: the noise spectrum $N(k_x,k_y)$ divided by the OTF $H(k_x,k_y)$. If this term is minimised, then the estimated spectrum will approach the true input spectrum. However, in practice, the noise spectrum is unknown to us, and in general is likely to have significant high frequency content.

An alternative approach in the presence of noise is the Wiener filter. We can define the Wiener filter, $G(u,v)$, as follows:

$$Y(k_x,k_y) = \frac{H^*(k_x,k_y)}{|H(k_x,k_y)|^2 + \frac{W_N(k_x,k_y)}{W_F(k_x,k_y)}} \tag{5.28}$$

where H^* is the complex conjugate of the OTF, $W_F(k_x,k_y)$ is the input power spectrum, and $W_N(k_x,k_y)$ is the noise power spectrum, i.e.

$$W_F(k_x,k_y) = \langle |F(k_x,k_y)|^2 \rangle \quad and \quad W_N(k_x,k_y) = \langle |N(k_x,k_y)|^2 \rangle \tag{5.29}$$

We can consider the term $W_N(k_x,k_y)\,/\,W_F(k_x,k_y)$ as the reciprocal of the signal-to-noise ratio. When the signal is very strong relative to the noise, $W_N(k_x,k_y)\,/\,W_F(k_x,k_y) \approx 0$ and the Wiener filter becomes $H^{-1}(k_x,k_y)$, the inverse filter for the PSF. Conversely, when the signal is very weak, noise dominates, and $Y(k_x,k_y) \to 0$.

A drawback of the Wiener filter is that it is unable to reconstruct frequency components which have been degraded by noise or components for which $H(k_x,k_y) = 0$. In some cases, an averaging of noise through interpolation is more effective than applying a filter. For the purposes of this study, we simply set the high frequency noise to zero. While this may not be the most accurate method of reducing the noise in the weighting calculation, an investigation into a more rigorous approach is beyond the scope of this project.

In Figure 5.16, the Fourier transform of the weighting distribution for the 248 nm Braggstar source is shown, with the high frequency noise components removed. An inverse Fourier transform is carried out on the noise-reduced quantities to give the weighting distribution, which is then sampled and applied to the squared and shifted elementary functions. The sum of the weighted elementary functions is compared to the original intensity distribution and presented in Figure 5.17. Reducing the noise in the weighting distribution results in a smoother reconstructed intensity profile, and for both the (a) short and (b) long axes, the reconstructed intensity plot corresponds well with the original intensity profile.

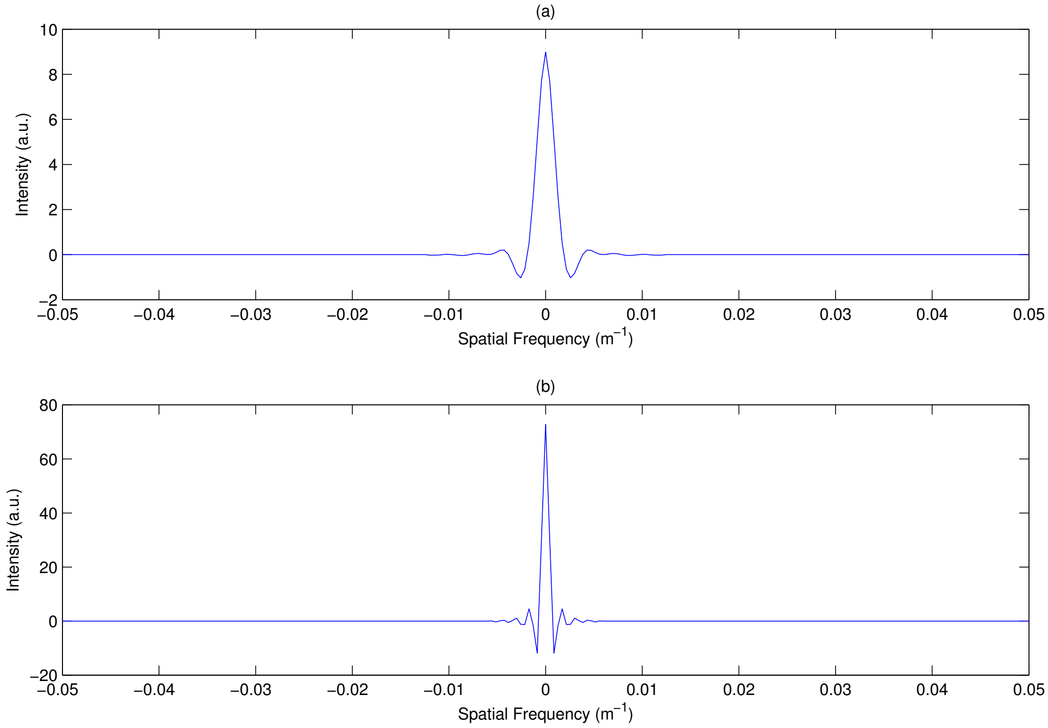

Figure 5.16: Plot of the noise-reduced Fourier transform of the weighting distribution calculated for the (a) short and (b) long axis of the 248 *nm* Braggstar source.

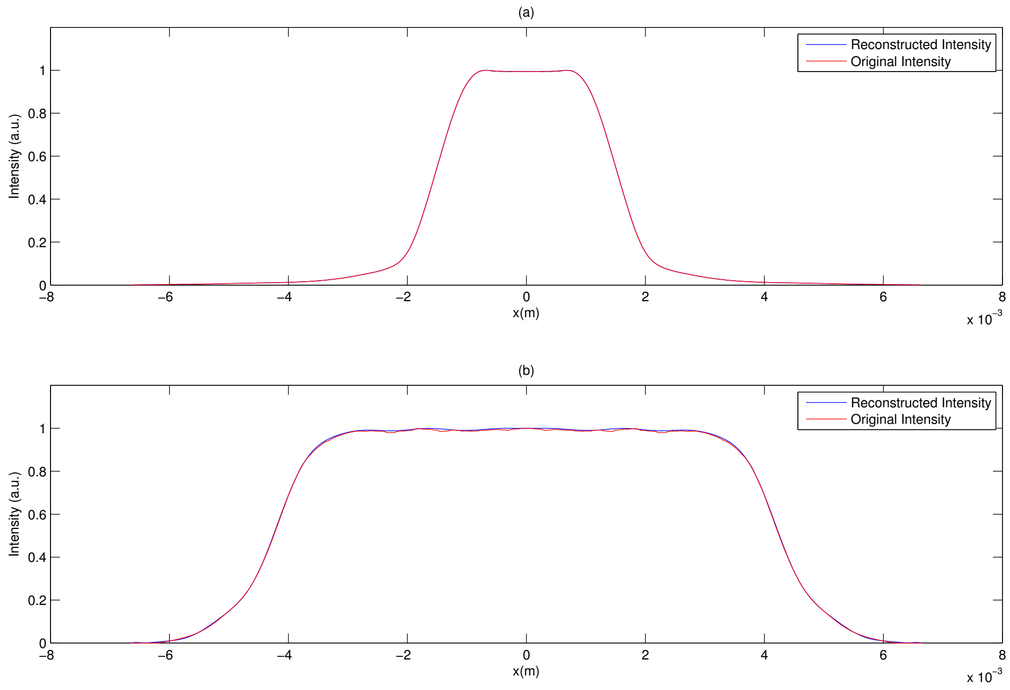

Figure 5.17: Comparison of the original intensity distribution and the noise-reduced reconstructed intensity distribution calculated for the (a) short and (b) long axis of the 248 *nm* Braggstar source.

5.4 Imaging an Opaque Edge

In 1965, Watrasiewicz published his work on the *"Theoretical calculations of images of straight edges in partially coherent illumination"* [90]. Based on Hopkins' approach to partial coherence [91], the intensity in the image plane is calculated due to one point of the effective source, followed by integration over the whole area of the effective source to give the intensity distribution due to the whole source. Watrasiewicz performed calculations for various degrees of coherence, where the coherence is defined relative to the numerical apertures of the system, i.e.

$$S = \frac{NA_c}{NA_o} \tag{5.30}$$

where NA_c and NA_o represent the numerical aperture of the condenser and the objective respectively, and S is equivalent to the degree of coherence. For illumination with coherent light, $S \to 0$, and in the presence of incoherent illumination, $S \to \infty$. For calculations involving values of S between these two extremes, the illumination is considered partially coherent. Watrasiewicz compared the intensity in the image of a straight opaque edge for various values of S. For $S \to 0$, the intensity at the edge in the image plane was 25 % of the original edge intensity. As S approaches the incoherent limit, the image intensity at the edge is 50 % of the original edge intensity. For partially coherent illumination, the edge intensity in the image lies between these values: for example, when $S = 1.0$, the intensity at the edge drops to 33 %. Figure 5.18 shows the intensity at the edge in the image of an opaque edge illuminated by coherent and incoherent plane waves. The illumination is uniform across the edge and, as expected, the intensity drops to approximately 50 % and 25 % for incoherent and coherent light respectively.

For initial partial spatially coherent calculations we assume the source to be a Gaussian Schell-model source. The partially coherent image of an opaque edge was calculated and compared with the image of an opaque edge for incoherent and coherent illumination with a Gaussian intensity distribution. The results are shown in Figure 5.19. In all cases, the width of the Gaussian intensity distribution is the same ($\sigma_I = 1000\ \mu m$) and the coherence width (σ_g) was varied. In Figure 5.19 (a) $\sigma_g = 100\ \mu m$, which requires 115 elementary functions to be propagated, (b) $\sigma_g = 10\ \mu m$, requiring 1145 elementary functions, and in (c) $\sigma_g = 1\ \mu m$, with 11459 elementary functions to be propagated. The numerical aperture of the imaging system is 0.55 and the wavelength is 248 nm. In Figure 5.19 (a), (b) and (c) the intensity at the edge of the

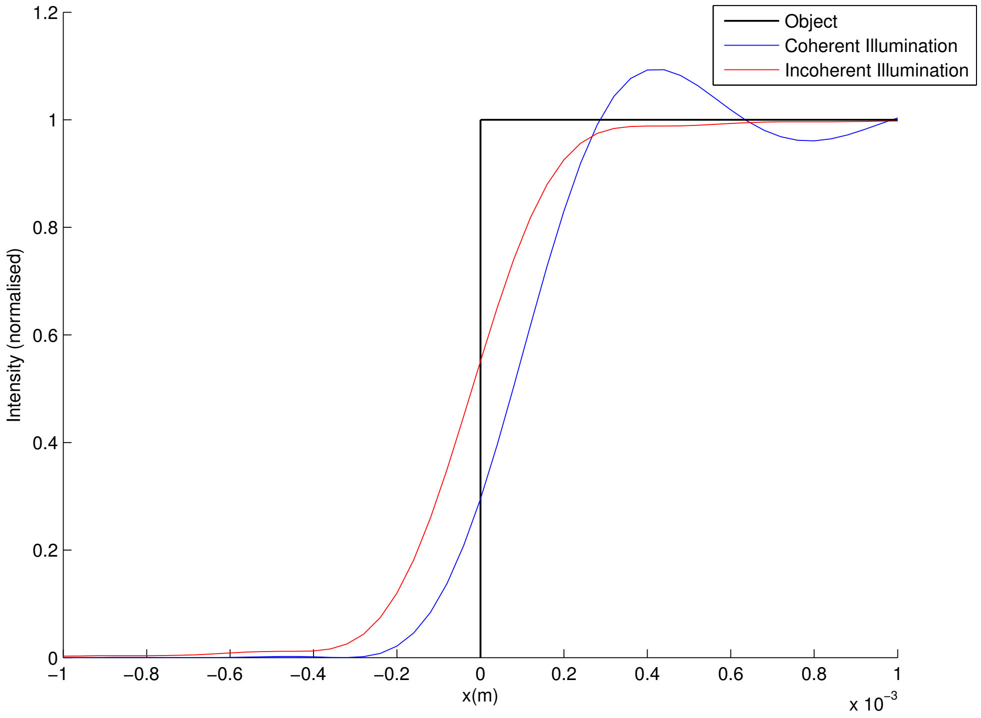

Figure 5.18: Intensity distribution in the image of a straight opaque edge illuminated by coherent and incoherent light.

image corresponds well with the purely coherent case despite the decreasing coherence. In fact, the green line representing the partially coherent data lies directly on top of the blue coherent data line in these first 3 plots. The reason for this is that in each of these cases the defined width of the coherence distribution is still large compared to the psf of the imaging system. As such, the intensity at the edge in the image of the opaque edge appears consistent with coherent imaging. In Figure 5.19 (d), σ_g is sufficiently small that the imaging appears less coherent, all three data lines are visible on the plot, and the intensity at the edge lies between the coherent and incoherent extremes.

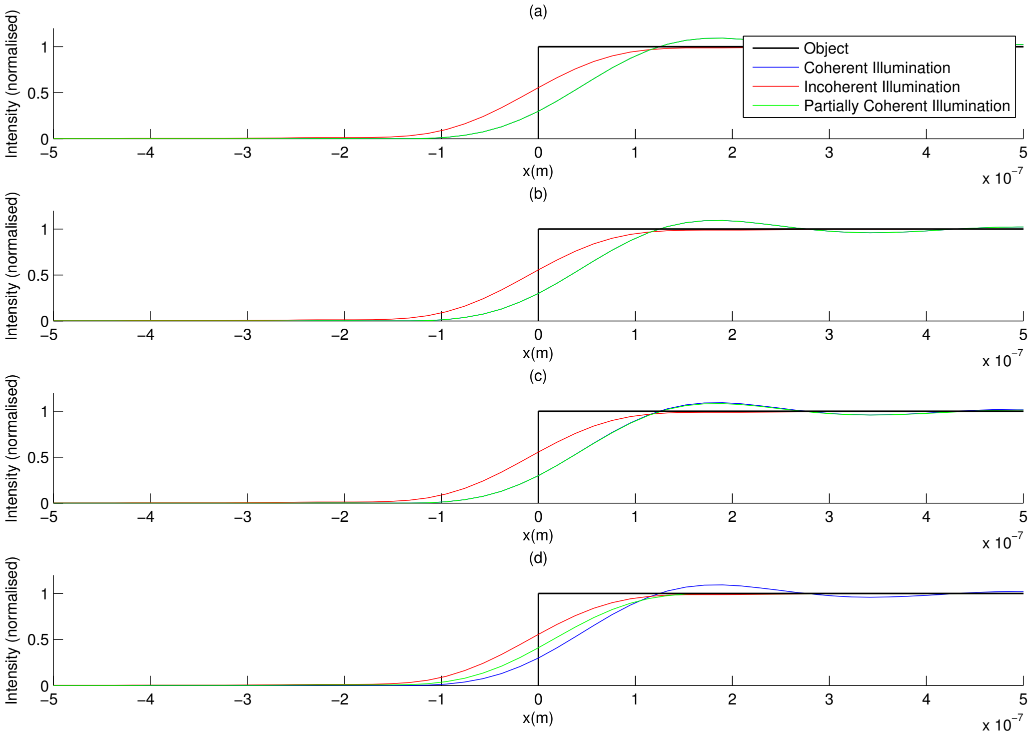

Figure 5.19: Intensity distribution in the image of a straight opaque edge illuminated by coherent, incoherent, and partially coherent light with a Gaussian intensity distribution. The intensity width is the same in all cases, $\sigma_I = 1000\ \mu m$, and the coherence width varies with (a) $\sigma_g = 100\ \mu m$, requiring 115 elementary functions, (b) $\sigma_g = 10\ \mu m$, requiring 1145 elementary functions, (c) $\sigma_g = 1\ \mu m$, requiring 11459 elementary functions, and (d) $\sigma_g = 0.1\ \mu m$, requiring 114591 elementary functions. In (a), (b) and (c), the green line representing the partially coherent data obscures the blue line representing the coherent data line which lies directly below it.

Next, the opaque edge is illuminated by the real beam in 1-D. Again, the wavelength is 248 *nm* and the system numerical aperture is 0.55. The results are given in Figure 5.20. The spatial coherence distribution in both the (a) short and (b) long axis is large compared to the psf of the imaging system, and, as we expect (given the results in Figure 5.19), the intensity in the image of an opaque edge illuminated by the real beam corresponds with the purely coherent case. As before, the data line given by coherent illumination (blue) is obscured by the partially coherent data line (green) which lies directly on top of it.

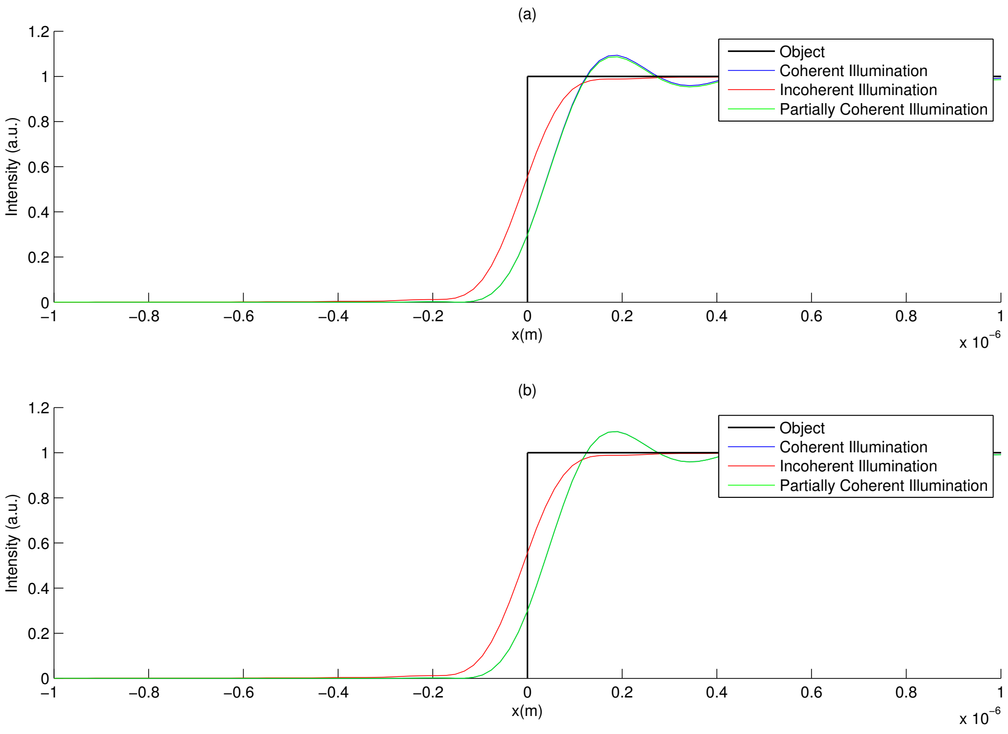

Figure 5.20: Intensity distribution in the image of a straight opaque edge illuminated by coherent, incoherent, and partially coherent light. (a) corresponds to the measured data for the short edge of the 248 *nm* beam, with the intensity width $\sigma_I = 0.003\ m$, and the coherence width $\sigma_g = 0.0007\ m$, and (b) corresponds to the measured data for the long edge of the 248 *nm* beam, with an intensity width $\sigma_I = 0.0085\ m$, and the coherence width $\sigma_g = 0.0002\ m$. As seen in Figure 5.19, the blue data line representing coherent illumination is completely obscured by the green data line representing partially coherent illumination.

At this point, the Elementary Function Method has been successfully applied to a

Gaussian Schell-model beam of various degrees of spatial coherence, and to the real 248 *nm* Braggstar excimer source. The results of imaging an opaque edge have proven useful in understanding high resolution imaging at this level. A partially coherent source with a relatively high degree of spatial coherence (i.e. the excimer source) will appear approximately coherent as long as the FWHM of the coherence distribution is greater or equal to the FWHM of the point spread function of the system. Imaging at 248 *nm* with a numerical aperture of 0.55 produces a narrow point spread function with a FWHM on the order of a micron or less. The effects of imaging with partially spatially coherent light will only become apparent when the coherence distribution becomes narrower than this. In the next Chapter, we will investigate the effects of an imaging beam homogenizer which, we expect, will highlight some of the difficulties experienced when working with light with a high degree of spatial coherence.

Chapter 6

Modelling an Imaging Beam Homogenizer

In Chapter 1, beam homogenization was introduced as an essential process for lithographic applications requiring high level beam uniformity. The raw beam emerging from an excimer laser changes its spatial intensity distribution over time and, in an effort to ensure a stable illumination source, various types of homogenizing devices are now an essential component of modern lithographic systems. In this chapter, starting in Section 6.1, we explore the method behind beam integration and its effectiveness in beam homogenization. In Section 6.1.2, we look at modelling an imaging beam homogenizer using the theory of multiple beam interference. This method is tested for coherent and incoherent sources, and, in Section 6.2, is applied to the Elementary Function Method. However, the end result appears to indicate that in practice the excimer laser source is actually so spatially coherent that very little smoothing of the intensity distribution is obtained in the homogenization process due to the coherence properties of the beam.

6.1 Beam Homogenization Optics

The homogenization of laser beams is an important issue in the areas of laser material processing, laser measuring techniques and laser analysis. In lithography, uniform illumination at the target plane is especially required, along with a well-defined energy- or power-density value. The uniformity of exposure duration is essential, but to maintain exposure latitude, the uniformity of the irradiance incident on the wafer is hugely important. Uniformity, U, for lithography is defined as the deviation of irradiance around the mean irradiance [28]:

$$\pm U = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})}, \tag{6.1}$$

where I_{max} is the maximum irradiance and I_{min} is the minimum irradiance. A beam uniformity in the range of $\pm 5\%$ is standard for laser machining applications and $\pm 2\%$ for photolithography [92]. This is measured with an 'open field' reticle that has no features present. Transmission from the reticle will vary for different features being printed.

The rectangular beam produced by most excimer lasers must usually be reshaped to match the needs of the process. Because of the shape and intensity profile of the raw excimer beam, these transformations cannot be made using the simple optical systems used for lasers with well-behaved Gaussian beams. The most popular technique is beam integration, which consists of mixing fractions of the input beam to smooth out the intensity spikes [19], [25], [26], [93], [94]. This device implicitly assumes spatial incoherence of the mixed beams.

A multi-aperture integrator system consists of two components: (a) a subaperture array component consisting of one or more lenslet arrays which divides the beam into an array of beamlets and applies a phase aberration to each beamlet, and (b) a beam integrator or focusing component which overlaps the beamlets from each subaperture at the target plane [95]. The amplitude of the irradiance arriving at the target plane is a Fourier transform of the original wavefront aberrated by the lenslet array.

6.1.1 Imaging Homogenizers

Beam-folding homogenizers are designed on the principle that the source beam is divided into subsections that are superimposed on top of one another. There are two types of beam-folding homogenizers: *imaging* and *non-imaging*. Both types of homog-

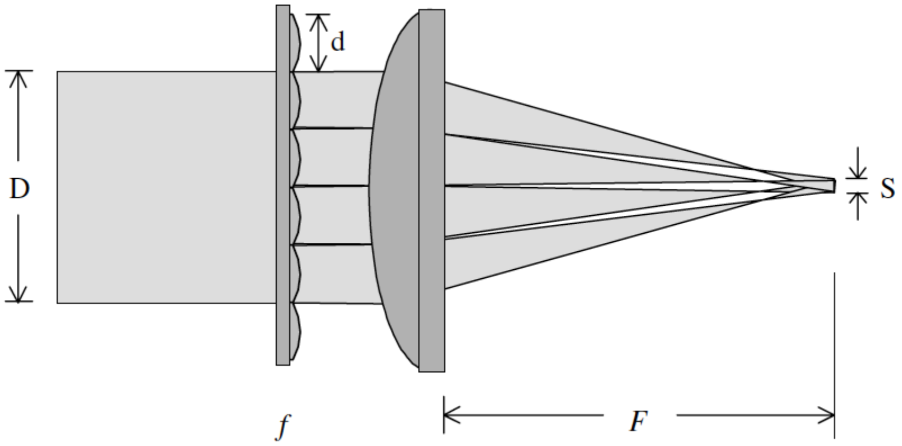

Figure 6.1: Concept of a non-imaging beam integrator to provide uniform beam intensity. A single microlens array is used to divide the beam into subapertures which are imaged using a spherical lens. Image source: [95]

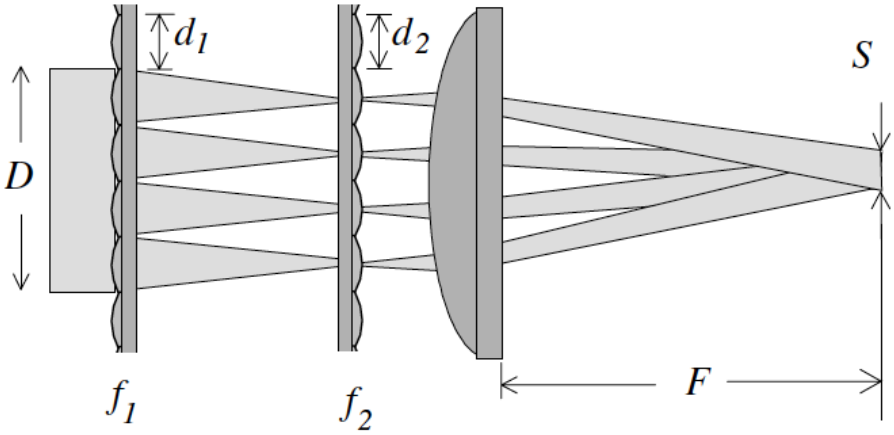

Figure 6.2: Concept of an imaging beam integrator to provide uniform beam intensity. Two identical microlens arrays are used. The first lenslet array divides the beam into subapertures. Then the second lenslet array, along with the spherical lens, images the subapertures onto the final image plane S. Source: [95]

enizers use arrays of crossed cylindrical lenses or square spherical lenses to divide the beam into beamlets. Square lens apertures generate a square, flat-top or top-hat beam profile in the Fourier plane; circular microlenses will generate a circular flat-top intensity profile. The beamlets are then passed through a spherical (Fourier) lens to be overlapped at the homogenization plane. The intensity pattern in the homogenization plane is related to the spatial frequency spectrum generated by the microlens array or arrays prior to entering the Fourier lens. The difference between the two types of homogenizers is that non-imaging homogenizers use a single lens array with a spherical lens (Figure 6.1), imaging homogenizers use two lens arrays and a spherical lens (Figure 6.2).

The beam propagation in an imaging homogenizer is based on the Köhler illumination system. In Köhler illumination the object is illuminated with a perfectly defocused image of the source ensuring uniform illumination as any intensity unevenness in the source is not seen.

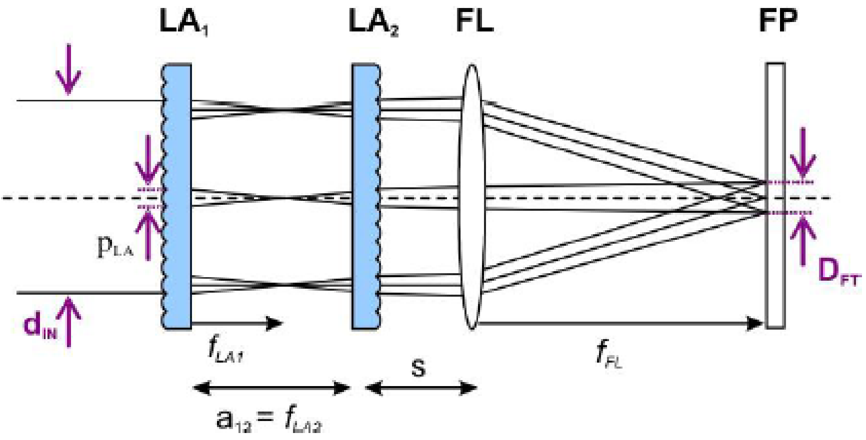

Figure 6.3: Suss MicroOptics imaging homogenizer: Two microlens arrays LA_1 and LA_2, on spherical Fourier lens FL. Image source: [92].

The choice of beam homogenizer depends on the light source and the desired application; Dickey [28] suggests the Fresnel number is used as a guideline. Using the terms shown in Figure 6.3, which represents the layout of the homogenizer in use in DIMO, we can define the Fresnel Number (FN) as

$$FN = \frac{p_{LA} D_{FT}}{4\lambda f_{FL}}, \tag{6.2}$$

where p_{LA} is the pitch of microlens array, D_{FT} is the dimension of the flat-top intensity profile in the homogenization plane, f_{FL} is the focal length of the Fourier Lens, and λ is the wavelength of the source. Higher Fresnel numbers give sharper edges and smaller variations of the flat-top profile. Non-imaging homogenizers should have FNs >10, or even FN>100 to obtain a good uniformity. Non-imaging homogenizers are suitable for large area illumination, as the flat top dimension D_{FT} is proportional to the Fresnel Number FN. For small Fresnel numbers FN<10 or high uniformity flat top requirements, the imaging homogenizer is the preferred solution.

Usually, imaging homogenizers consist of two identical lenslet arrays (same pitch p_{LA} and focal length $f_{LA1} = f_{LA2}$). The divergence is controlled by the distance between the second lenslet array and the Fourier lens - as this distance, s increases, the beam divergence increases. This arrangement is the classic fly's eye condenser. For imaging homogenizers the diameter of the individual beamlets at the second microlens array LA_2 must be smaller than the lens pitch to avoid overfilling of the lens aperture and the loss of light.

The imaging beam homogenization system that we will model here is based on a system from Suss MicroOptics, and is pictured in Figure 6.3. It has been designed for use with 248 *nm* Braggstar source in place in DIMO, and comprises two square spherical lenslet arrays and a spherical Fourier lens, arranged as shown in Figure 6.3. Inserting the specifications of this homogenizer into Equation 6.2, the Fresnel Number (FN) can be calculated as follows:

$$FN_{Suss} = \frac{0.5\ mm \times 4.59\ mm}{4 \times 248\ nm \times 150\ mm} = 15.423 \tag{6.3}$$

The Fresnel Number is just above the threshold of 10, deeming it suitable for a non-imaging integrator, but the high uniformity requirements of lithography (±2%) make the imaging homogenizer a more suitable choice for this application. Prior to entering the beam homogenization optics, the raw beam passes through a randomizing phase plate (or rotating diffuser). The purpose of the rotating diffuser is to reduce the spatial coherence by the addition of static aberration. The performance of a lenslet array is generally improved by the additions of a randomizing aberration plate. If such a randomizing phase plate is used, then only the dynamic effects in the laser will contribute to further suppression of coherent artifacts, and the static (fixed) aberrations of the laser do not matter and should not be included in the laser characterization [24].

6.1.2 Modelling an Imaging Homogenizer

In Chapter 5, the Elementary Function Method was applied successfully to the 248 *nm* Braggstar beam. A comparison between the measured beam intensity profile in both axes of the beam and the sum of the shifted, squared, and weighted elementary functions showed almost identical profiles. The next step in the beam model is to include homogenization optics, or to mimic the effect of such optics in some way.

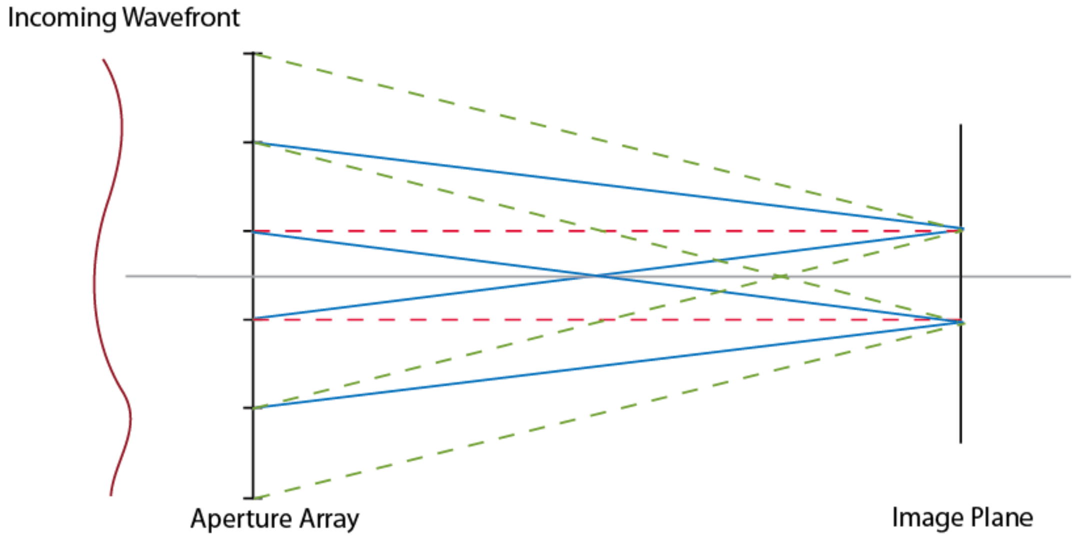

Figure 6.4: Special case of Multiple Beam Interference to imitate the effect of a lenslet array. The aperture (slit) diameter is equal to the aperture separation.

As discussed in Section 6.1.1, a wavefront arriving at the first microlens array of an imaging homogenizer is divided into an array of beamlets. The second microlens array, along with the Fourier lens, images the complex amplitude of the incident wavefront onto the image plane, where each beamlet is overlapped with a certain magnification. The superposition of beamlets averages out any intensity fluctuations in the incident wavefront if the beamlets are mutually incoherent; a greater number of lenslets results in greater uniformity in the image plane. On a basic level, for light of arbitrary coherence, this process may be compared to the phenomenon of *Multiple Beam Interference* [37], [96]. Earlier, in Section 4.2.3, the theory behind the interference pattern created by light incident on a non-redundant array was introduced. In the double-slit case, the intensity at a point in the image plane is equal to the sum of the intensities of the two beams multiplied by an interference term. If the source is spatially incoherent, the interference term is zero. If the source is partially coherent, the interference term begins to grow, becoming more significant as the coherence of the source increases. Here, we will approximate the effect of an imaging homogenizer using a special case of multiple beam interference: the aperture (or "slit") diameter and aperture separation are the same, and are equal to the diameter of a single lenslet (see Figure 6.4). With the exception of the beamlet corresponding to the centre lenslet (on the optic axis), a phase shift, dependant on distance from the optic axis, is applied. The beamlets are then added together. To test the method, the model was applied to a purely coherent source (a point source positioned both on-axis and off-axis) and to sources of decreasing coherence.

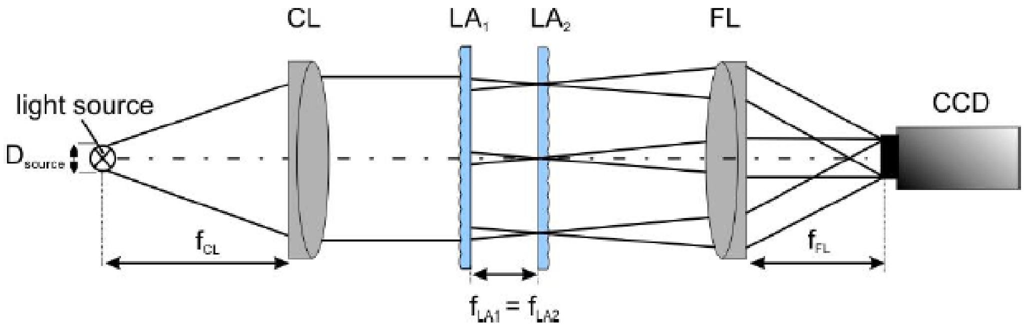

Figure 6.5: Scheme of a flys' eye condenser with two identical microlens arrays. Image source: [92].

Before proceeding with the model, we need to apply any necessary scaling changes. To illustrate the source position and the optical path of a wavefront before it arrives at the lenslet array, we refer to Figure 6.5. A wavefront leaving the source travels a distance f_{CL} to the collimating lens, and then onwards to the first lenslet array. The effect of the collimating lens can be modelled using a simple Fourier transform. However, the distance f_{CL} is important as it affects the subsequent scaling of the wavefront in the frequency domain. The scaling can be calculated as follows:

$$y = \frac{\lambda \times f_{CL}}{x} \tag{6.4}$$

where x is the width of the sampling window in the source plane, which is the product of the number of samples taken and the pixel size (unit: m), y is the scaled pixel size, λ is the wavelength of the source (248 nm), and f_{CL} is the focal length of the collimating lens. For example, a 2 μm pixel size in a source plane with 700001 sample points, and with $f_{CL} = 150\ mm$

$$y = \frac{248\ nm \times 150\ mm}{700001 \times 2\ \mu m} = 0.026\ \mu m \tag{6.5}$$

This difference in scaling between the initial wavefront leaving the source and the wavefront (post Fourier transform) arriving at the first lenslet array is taken into account in the lenslet model. It is important to ensure the lenslets are filled by the source illumination, e.g. by diverging the beam as shown in Figure 6.5. Also, in order to apply the appropriate tilt to each wavefront segment in the model (as shown in Figure 6.4), f_{FL} was taken to be 150 mm.

The interference pattern in the image plane for a coherent point source on-axis is given in Figure 6.6. The intensity distribution in the image plane is not uniform: the coherent nature of the light source results in interference between the beamlets from each aperture which manifests itself as a series of diffraction orders or Airy disks in the image plane. Similarly, the intensity in the image plane for a coherent point source positioned 2 μm off-axis is shown in Figure 6.7. In both cases (i.e. Figure 6.6 (a) and Figure 6.7 (a)), an interference pattern, spanning a width equal to the diameter of one lenslet (0.5 mm), is seen in the image plane of the Fourier lens. Zooming in on both of these plots (i.e. Figures 6.6 and 6.7, plot (b)), the individual diffraction orders are visible. For the point source on-axis (Figure 6.6), there is an Airy disk centered in the image plane; the central peak of the airy disk is centered in the image plane. In Figure 6.7, we see the effect of moving the point source just one pixel off-axis (i.e. 2 μm). As the point source moves off-axis, the diffraction orders shift. An extended incoherent source should give a resultant intensity that is the sum of many such distributions, each displaced from each other. The results of Figures 6.6 and 6.7 show that, with spatially coherent illumination, this "homogenizer" produces an extremely non-homogeneous intensity, because of the process of interference.

To investigate the effect of decreased spatial coherence, the width of the source was increased by adding more point sources (delta functions) either side of the central, on-axis point source. Increasing the width of the source decreases the spatial coherence of the beam combination process. In the first instance, one extra delta function is placed either side of the on-axis point source, to create a source width of 3 pixels, which corresponds to 6 μm (1 pixel = 2 μm). Each delta function is imaged separately, and the intensity in the image plane is summed. The result of this test is given in Figure 6.8 (b). The source width is continually increased by 2 pixels in each plot (c) - (u) and again in Figure 6.9 (a) - (u). Each of these plots depict the effect of increasing the width of the source: the interference effects in the image plane are gradually less-

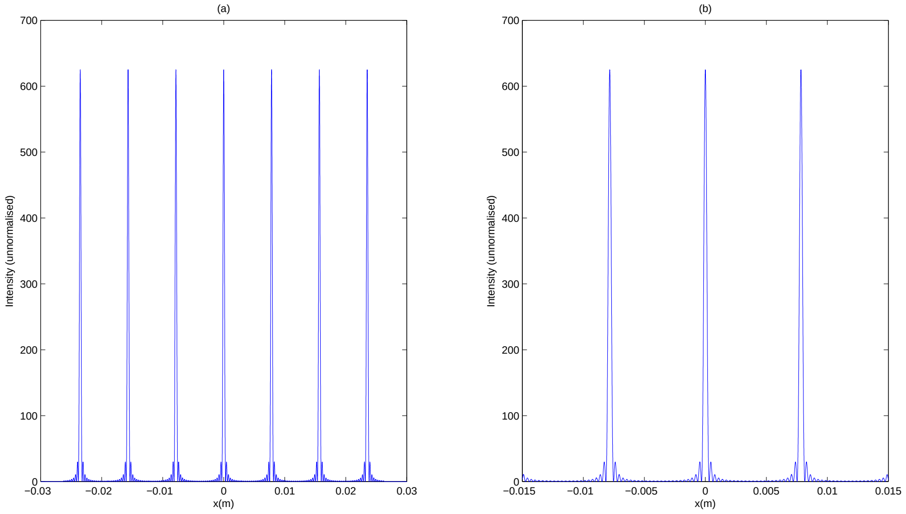

Figure 6.6: (a) Simulated intensity in the image plane of an imaging homogenizer illuminated by a coherent point source on-axis, and (b) zoomed in central part of (a) showing single airy disks of the diffraction orders.

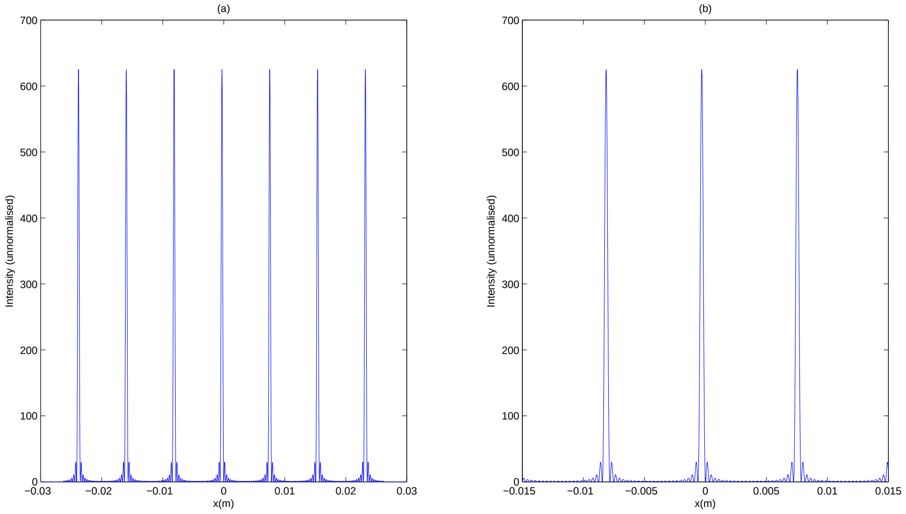

Figure 6.7: (a) Simulated intensity in the image plane of an imaging homogenizer illuminated by a coherent point source off-axis, and (b) zoomed in central part of (a) showing single airy disks of the diffraction orders. Moving the point source off-axis by a single sampling point (0.019 μm) results in a shift in the diffraction orders, visible particularly in (b).

ened, corresponding to a decrease in spatial coherence. Comparing the result from a coherent point source on-axis (shown again in Figure 6.8 (a)) to a source 3 pixels wide (Figure 6.8 (b)), we see the diffraction orders (Airy disks) become broader as the source width increases. The diffraction orders continue to widen with increasing source width (Figure 6.8(c) - (m). In Figure 6.8 (m) we see the diffraction orders are at their widest (with very little space between adjacent orders).

In Figure 6.8 (n), a "flip" occurs and the diffraction orders become narrow once again. Once again the diffraction orders widen with increasing source width, until reaching a maximum width (Figure 6.9 (f), and the orders become narrow once again (Figure 6.9 (g)). This cycle continues as the source width is increased, but with a noticeable decrease in amplitude of the diffraction orders each time. The spatial coherence of the source is decreasing, and the intensity is gradually becoming more uniform.

Zimmermann et. al. reported similar findings in their 2007 paper [92]. Assuming each light source point with axial distance x_{source} forms a laterally shifted spot array with a shift of $\Delta x = \theta_{in} f_{FL} = x_{source} f_{FL} / f_{CL}$, then *"if the diameter D_{source} of the light source is such that the lateral shift Δx of all light source points is just one period Λ_{FP} of the spot array with $\Lambda_{FP} = \lambda f_{FL} / p_{LA}$, a smooth homogeneous intensity will be obtained in the focal plane of lens FL (assuming all light source points emit identically). However, if the diameter of the light source increases further there will be again high fluctuations of the intensity pattern."*. Zimmermann et. al. present the following relations

$$\frac{D_{source}}{f_{CL}} = \frac{\lambda}{p_{LA}} \tag{6.6}$$

for the case of smooth homogeneous intensity, and

$$\frac{D_{source}}{f_{CL}} = \frac{1.5\lambda}{p_{LA}} \tag{6.7}$$

for high fluctuations. Eventually the source diameter becomes so large that the spot arrays of the different light source points are shifted by many periods, and the intensity is almost homogeneous even if the light source diameter changes by a few percent.

In this Section, the method has been applied to a purely spatially coherent point source, and to incoherent sources of increasing width. Physically, the intensity in the image plane for each source seems reasonable. In the purely coherent case, the greatest amount of interference is obtained. We see the effect of the light emerging from

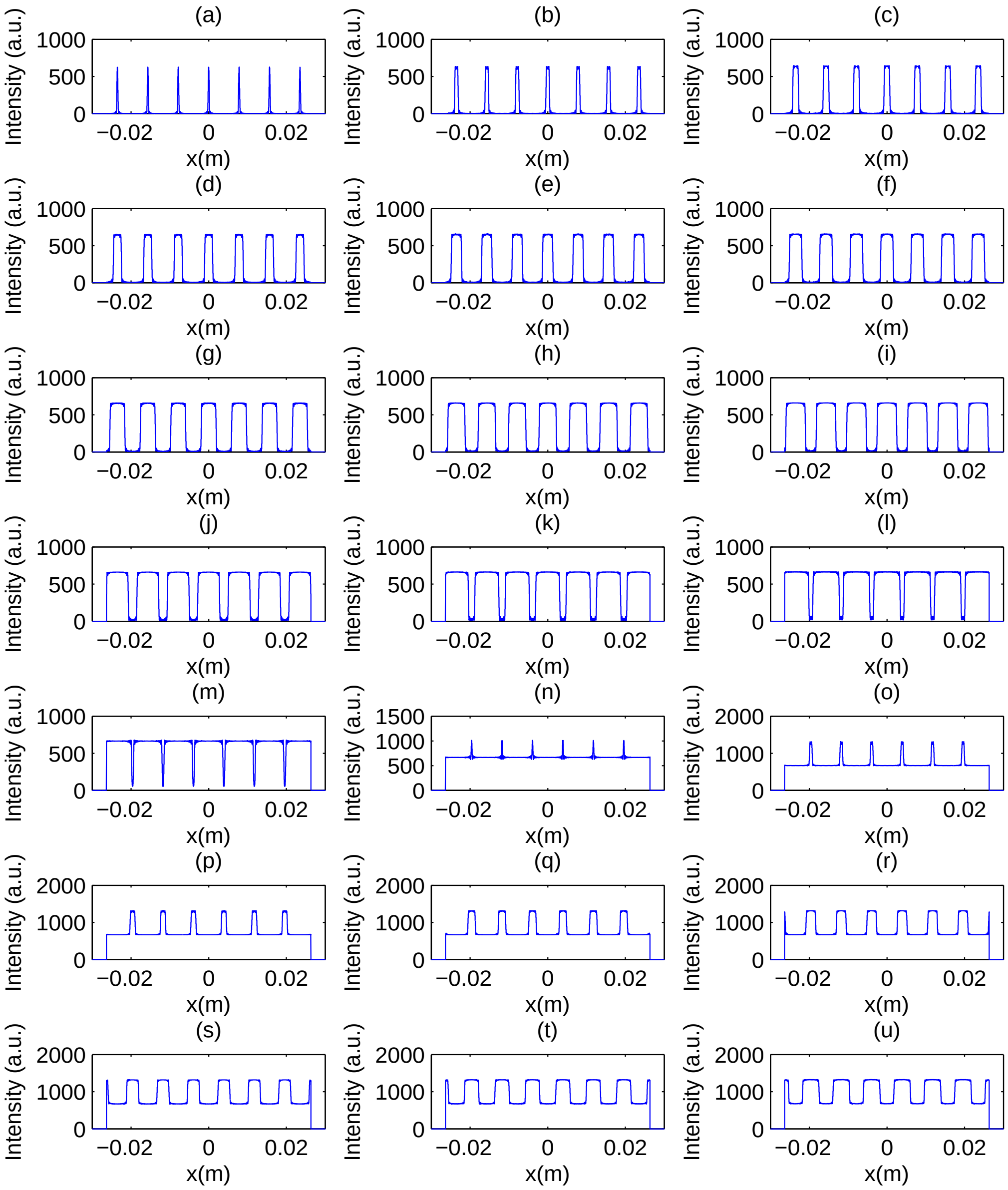

Figure 6.8: Simulated intensity in the image plane of an imaging homogenizer illuminated by sources of increasing width. In (a), the source has a width of 1 pixel (2 μm), in (b), the source width is 3 pixels (6 μm), in (c), the source width is 5 pixels (10 μm), and so on in (d) - (u), with the source width increasing by 2 pixels each time. In (u), the source width is 41 pixels, corresponding to 82 μm.

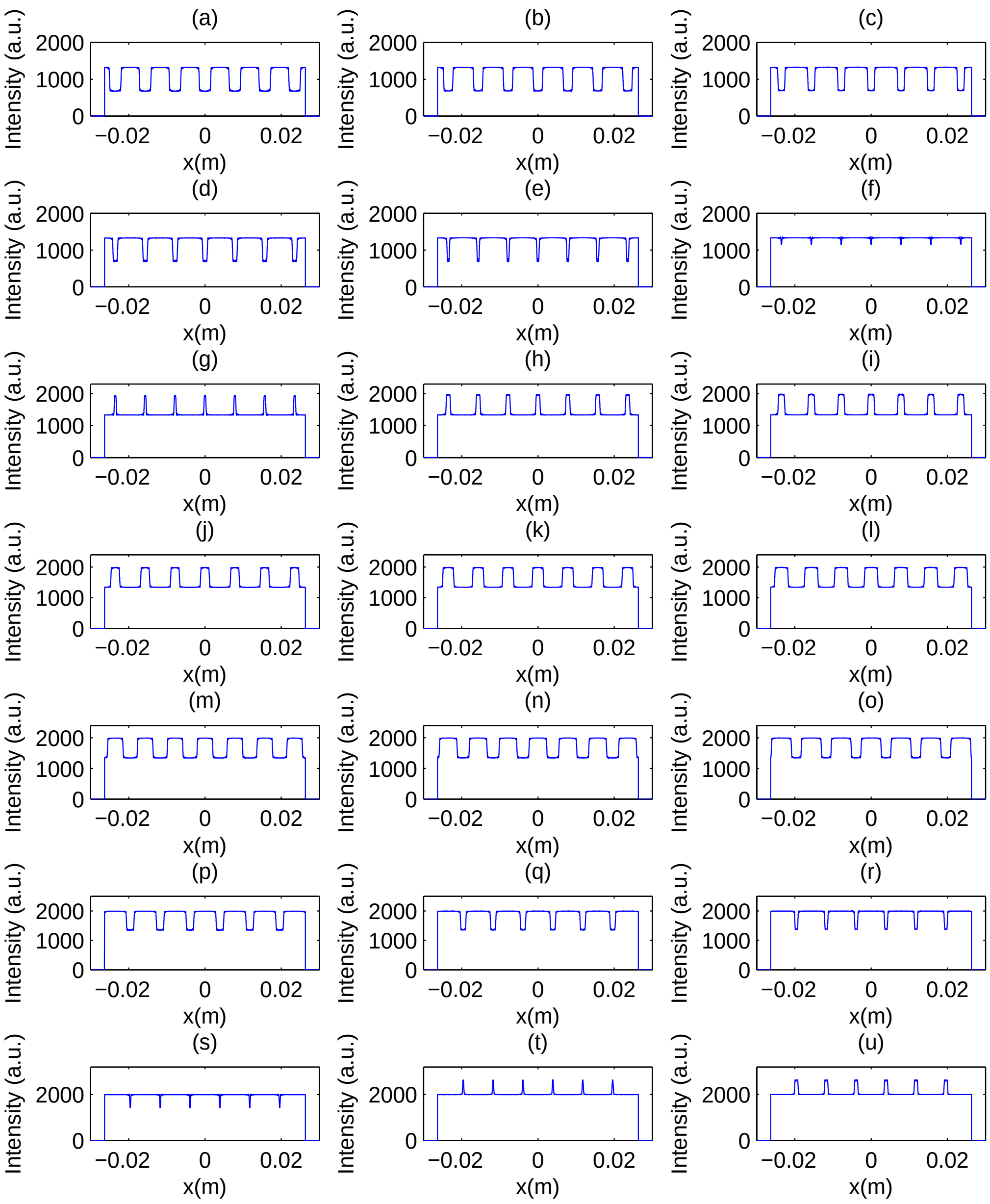

Figure 6.9: Simulated intensity in the image plane of an imaging homogenizer illuminated by sources of increasing width, and thus, decreasing spatial coherence. This is a continuation of Figure 6.8. In (a) the source has a width of 43 pixels (86 μm). The source width is increased in steps of 2 pixels (4 μm) in (b) - (u). The final width in (u) is 83 pixels (166 μm).

25 lenslets and interfering in the focal plane of the Fourier lens (Figure 6.6). Zooming in on this interference pattern enlarges the various diffraction orders (Airy disks). Moving the coherent point source off-axis by a single pixel results in a shift of the the diffraction orders (Figure 6.7), and as we expect, this shift is much less than the width of a single airy disk. The effect of increasing the source width is to decrease the interference effects, as seen in Figures 6.8 and 6.9. Most importantly we note that the intensity in the image plane is highly non-uniform, thus showing that the beam homogenizer does not work at all well with a coherent source. As the source width is increased, and thus the spatial coherence of the source decreases, the intensity of the source becomes more uniform. The next step is to apply the Elementary Function method to this imaging homogenizer approximation, where the delta functions modelled here will be replaced by elementary functions.

6.2 The Elementary Function Method applied to an Imaging Homogenizer

As in earlier stages of this study, the initial proof of concept of the method is performed using a Gaussian Schell-model beam. The model is applied in a similar way to the coherent case, except the delta functions are replaced by elementary functions. Each shifted elementary function is propagated through the imaging homogenizer using the special case of multiple beam interference, squared (to give the intensity) and then weighted according to its position in the source. The weighted intensities are then summed. The Elementary Function Method was tested in two ways: firstly the coherence width (σ_g) is defined and is unchanged as the intensity width (σ_I) is varied, secondly the coherence width is varied as the intensity width is fixed.

Figure 6.10 shows the simulated intensity in the image plane of an imaging homogenizer illuminated by Gaussian Schell-model source of various widths. The coherence width (σ_g) is fixed at 0.1 mm and the intensity width (σ_I) is set at 0.1 mm in Figure 6.10 (a) and increased gradually in plots (b) - (f). In each case, a greater number of elementary functions must be propagated to fully represent the source. In Figure 6.10 (a), the coherence width is equal to the intensity width, resulting in a highly spatially coherent source. According to our sampling criterion, 13 elementary functions are required to represent the source. As the width of the intensity distribution increases, the degree of spatial coherence of the source decreases, and as a result, the fluctuations in the intensity profile gradually decrease to produce a more uniform intensity

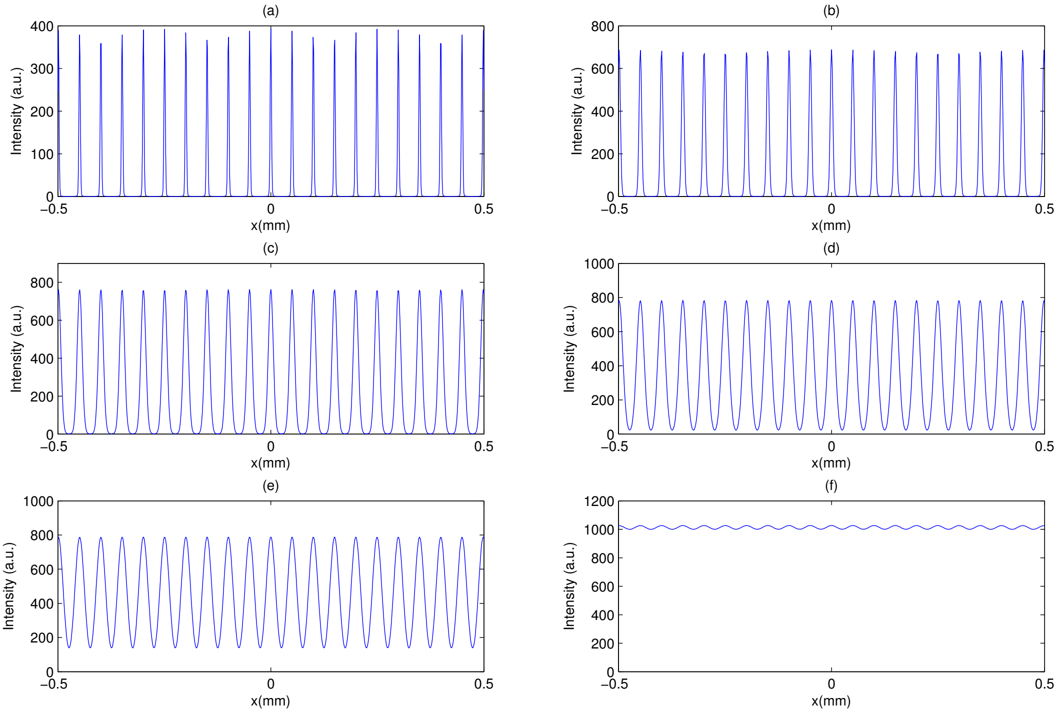

Figure 6.10: Simulated intensity in the image plane of an imaging homogenizer illuminated by Gaussian Schell-model sources of increasing intensity width. In each case, the coherence width is the same ($\sigma_g = 0.1\ mm$) and the intensity width (σ_I) is varied. (a) $\sigma_I = 0.1\ mm$, requiring 13 elementary functions, (b) $\sigma_I = 0.25\ mm$, requiring 29 elementary functions, (c) $\sigma_I = 0.5\ mm$, requiring 57 elementary functions, (d) $\sigma_I = 0.75\ mm$, requiring 87 elementary functions, (e) $\sigma_I = 1.0\ mm$, requiring 115 elementary functions, and (f) $\sigma_I = 2.25\ mm$, requiring 257 elementary functions..

distribution. In Figure 6.10 (f), $\sigma_g = 2.25\ mm$, requiring 257 elementary functions to be propagated. In this case, the area of the coherence distribution occupies just 4 % of the intensity profile, producing a source with a low degree of spatial coherence. Thus, the interference effects in the image plane have decreased considerably. This trend is similar to that observed in Figures 6.8 and 6.9 for an increasing source width.

In Figure 6.11, the intensity in the image plane of an imaging homogenizer is shown for Gaussian Schell-model sources of fixed intensity width ($\sigma_I = 0.1\ mm$) and various coherence widths (σ_g). In effect, the degree of coherence of the source is decreasing in each step (a) - (f) in Figure 6.11 as the width of the coherence distribution becomes increasingly small relative to the width of the intensity profile of the source. We observe the same gradual decrease in interference effects as in Figure 6.10. Figure 6.10 (a) and Figure 6.11 (a) are the same: in each case $\sigma_I = 0.1\ mm$ and $\sigma_g = 0.1\ mm$, and 13 elementary functions are required. From this point, in Figure 6.11 (b) - (f), the coherence width σ_g is decreased, thus decreasing the coherence of the source and increasing the number of elementary functions to be propagated.

From these two Figures it is clear that the relationship between the coherence and intensity widths is key to the effects (if any) seen in the image plane of the imaging homogenizer. The number of elementary functions required gives a good indication of the degree of spatial coherence of the source, and can be a useful predictor of the outcome of imaging such a source through a homogenizer. Furthermore, in both cases, the non-uniformity in the intensity profile in the image plane of the homogenizer as a result of the coherence properties of the source is not trivial.

The next step is to apply the model to the data gathered from the 248 *nm* excimer source. We will do this in three stages. Firstly, we will examine the effect produced when the imaging homogenizer is illuminated by a spatially incoherent source with the same intensity width as the excimer beam. Secondly, we will look at a Gaussian Schell-model source with the same measured intensity and coherence values as the Braggstar excimer beam. Finally, we will model the effect of illuminating the imaging homogenizer with the real excimer beam.

Figure 6.12 shows the simulated intensity in the image plane of an imaging homogenizer illuminated by a spatially incoherent source with the same intensity width as (a) the short axis and (b) the long axis of the Braggstar 248 *nm* beam. The intensity width in (a) is 3.27 *mm*, corresponding to a width of 1551 pixels, and in (b) the intensity width is 8.964 *mm*, corresponding to a width of 4247 pixels. These widths are large, particulary in comparison to the source widths examined in Section 6.1.2, and

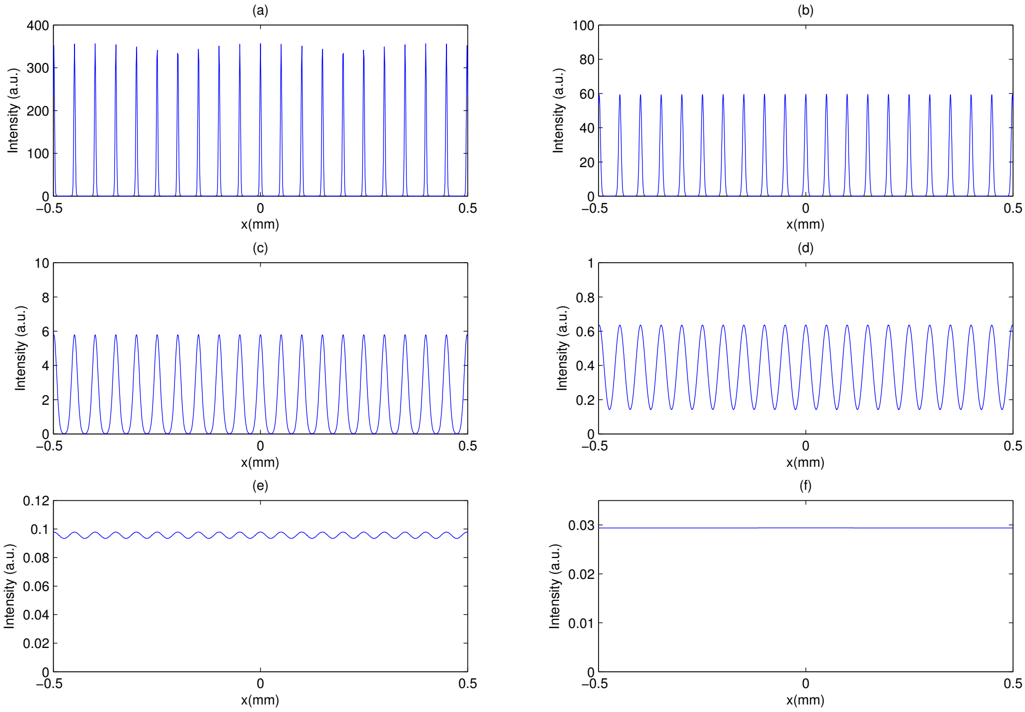

Figure 6.11: Simulated intensity in the image plane of an imaging homogenizer illuminated by Gaussian Schell-model sources of increasing intensity width. In each case, the intensity width is the same ($\sigma_I = 0.1$ mm) and the coherence width (σ_g) is varied. (a) $\sigma_g = 0.1$ mm, requiring 13 elementary functions, (b) $\sigma_g = 0.05$ mm, requiring 23 elementary functions, (c) $\sigma_g = 0.025$ mm, requiring 47 elementary functions, (d) $\sigma_g = 0.01$ mm, requiring 115 elementary functions, (e) $\sigma_g = 0.005$ mm, requiring 229 elementary functions, and (f) $\sigma_g = 0.003$ mm, requiring 287 elementary functions.

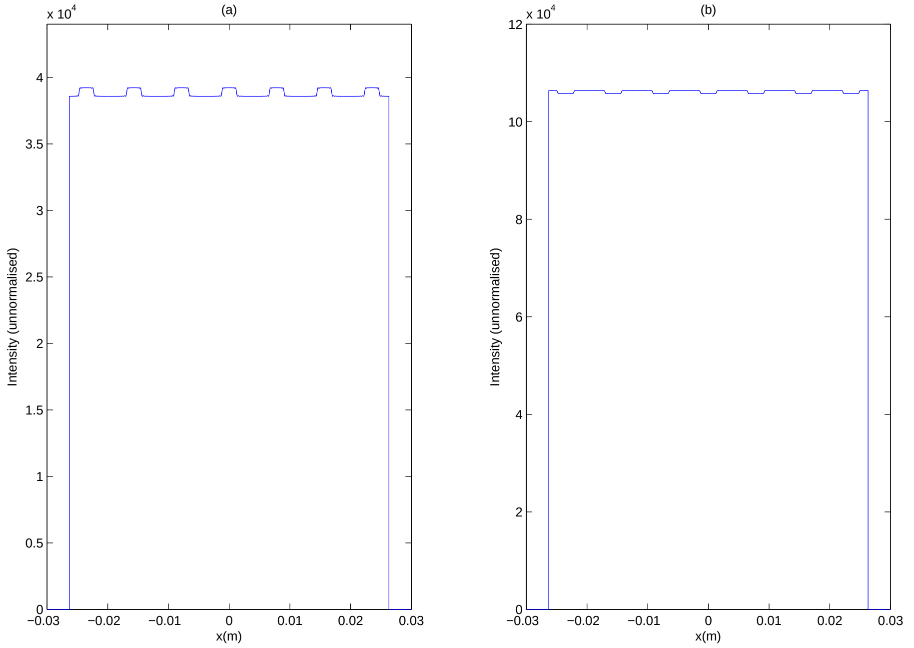

Figure 6.12: Simulated intensity in the image plane of an imaging homogenizer illuminated by an incoherent source with the same intensity width as the (a) short axis and (b) long axis 248 *nm* beam. The intensity width in (a) is 3.27 *mm* and in (b) the intensity width is 8.964 *mm*.

in both cases, the fluctuations in the intensity profile are very low, as we would expect to see from a spatially incoherent source.

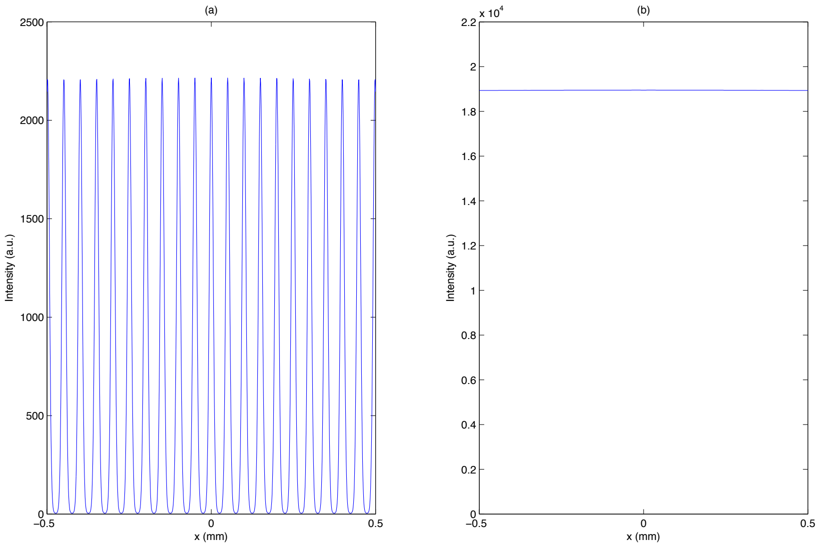

Figure 6.13: Simulated intensity in the image plane of an imaging homogenizer illuminated by a Gaussian Schell-model source with the same intensity and coherence widths as the (a) short axis and (b) long axis 248 *nm* beam. The intensity and coherence widths in (a) are 3.27 *mm* and 0.67 *mm* respectively, and in (b) the intensity and coherence widths are 8.946 *mm* and 0.214 *mm* respectively.

Next, the intensity and spatial coherence data for the 248 *nm* beam is applied to a Gaussian Schell-model beam, which is modelled using the Elementary Function Method. The intensity widths (σ_I) and coherence widths (σ_g) are assumed to refer to the FWHM's of Gaussian distributions, and the elementary function and sampling distribution is calculated as before. Figure 6.13 shows the intensity in the image plane for the 248 *nm* beam assuming illumination along (a) the short axis of the beam, and (b) the long axis of the beam. The intensity and coherence widths in (a) are 3.27 *mm* and 0.67 *mm* respectively, resulting in 57 elementary functions to be propagated, and in (b) the intensity and coherence widths are 8.946 *mm* and 0.214 *mm* respectively, requiring 479 elementary functions. In Figure 6.13 (a), the fluctuations in the image plane are considerable and the intensity profile is highly non-uniform. Figure 6.13 (b) shows a smoother intensity profile in the image plane, due to the much lower level of spatial coherence along this axis. This result is reasonable considering the coherence measurements for this source. Along the short axis of the beam, the width of the

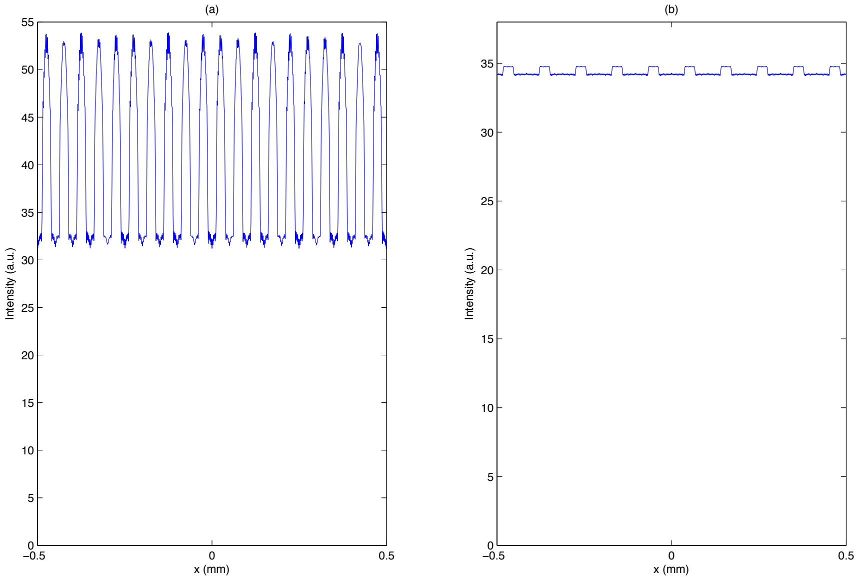

Figure 6.14: Simulated intensity in the image plane of an imaging homogenizer illuminated by the (a) short axis and (b) long axis of the 248 *nm* excimer source. The intensity and coherence widths in (a) are 3.27 *mm* and 0.67 *mm* respectively, and in (b) the intensity and coherence widths are 8.946 *mm* and 0.214 *mm* respectively.

coherence distribution corresponds to 20 % of the width of the intensity profile of the source; along the long axis of the beam, the coherence distribution covers just over 2 %.

The final part of this study involves the application of the homogenizer model to the real 248 *nm* source data using the Elementary Function Method. Figure 6.14 shows the simulated intensity in the image plane of an imaging homogenizer illuminated by the partially coherent excimer beam. As in Figure 6.13, a pronounced interference effect is seen when the homogenizer is illuminated by the short axis of the beam. The long axis of the excimer beam has a much lower degree of spatial coherence and thus the intensity profile produced by the imaging homogenizer is much smoother. It is clear from both Figure 6.13 and 6.14 that the use of an imaging homogenizer with a source with the level of spatial coherence present in the short axis of the beam does not smooth the final intensity (as we showed earlier for the fully coherent source). The spatial coherence of the source manifests itself in dramatic interference effects in the image plane of the homogenizer, and thus the source is not ideal for applications involving high resolution imaging. The imaging beam homogenizer is really only suitable for use with spatially incoherent sources, and the coherence properties of this excimer beam will produce dramatic interference effects in the image plane of the homogenizer.

Chapter 7

Conclusions

7.1 Summary of Thesis

Despite recent advancements in the technology of lithographic imaging, excimer lasers have maintained a popularity in the lithography field due to their low spatial coherence, high brightness, and cost efficiency. Excimer lasers are spatially partially coherent sources and, as such, are complex to model. The cross-spectral density - the quantity used to describe a partially coherent source in terms of its intensity distribution and its degree of coherence - is difficult to propagate mathematically. The propagation calculations are tedious and complex, and place large demands on computer memory. The purpose of this work was to develop a simple model of a partially coherent source. This model can then be applied to an imaging beam homogenizer, an example of a key element of modern lithographic systems which require high level beam uniformity.

The work in this thesis will be discussed under two headings: experimental and simulation.

7.1.1 Experimental Results

The standard method for measuring spatial coherence of a light field is based on the double-aperture interferometer (the Young's slits experiment). Spatial coherence data is gathered through lateral movement of the double-aperture mask across the beam, along with variation of aperture separation. However, this method requires several measurements across the beam and several aperture separations to get a full range of values across the coherence distribution of the source. The method used here is based on the non-redundant array: an array of points on a regular lattice such that the vector difference between any two points in the array is unique. The array is illuminated by the partially coherent source, and the spatial coherence is found from the Fourier spectrum of the resulting interferogram. In our mask design, one aperture separation occurred twice (i.e. a redundant frequency separation), which was a necessary compromise to ensure sufficient data was gathered within the coherence length of the source. This measurement of a redundant separation did not contribute to the final coherence measurements, only the unique aperture separations produced useful data. For each measurement, the array of apertures was placed within the area of highest intensity in the beam. This area was chosen as, in general, any work carried out using the raw beam would be centred on this area. While the spatial coherence measurements were in good agreement with those provided by the laser manufacturer, more data sets including other areas of the beam could produce a more in-depth and accurate representation of the spatial coherence across the beam. A redesign of the pinhole array used in the measurements to remove the one degree of redundancy would make the data gathering more efficient, and eliminate the need to reject any data, but could reduce the amount of useful measurements made within the coherence length of the source.

Recently, González and Mejía have published results based on the non-redundant array method previously published by them in 2007 [85] but extended to two dimensions [97]. Their design features a two dimensional grid of apertures, and is used to measure the spatial coherence of a laboratory-generated partially coherence source and a laser diode source. Future work in this area could involve simultaneous two-dimensional measurement of the spatial coherence of an excimer source.

7.1.2 Simulation Results

The Elementary Function Method

The Elementary Function Method, first introduced by Wald et al in 2005 [64], and developed further by Burvall et al in 2009 [65], was applied numerically to the 248 *nm* excimer source. Existing numerical methods to propagate partially coherent light have proven costly with respect to computer memory and time. Initially, before applying the method to the excimer beam, the model was applied numerically to a Gaussian

Schell-model beam. This test beam was a case study in [65], where the model was applied analytically to the partially coherent source. In Chapter 5, we presented the successful numerical application of the method to a Gaussian Schell beam. A comparison between the analytical and numerical results provided an initial proof of concept of the Elementary Function Method.

While the Elementary Function Method provides a way to reduce the complexity of propagation calculations for a spatially partially coherent source, some assumptions and approximations are necessary. The first assumption was that the initial cross-spectral density was assumed to be real. This puts limitations on the applicability, but it could be fulfilled for most beams. Divergent or convergent beams will have complex cross-spectral densities but can be modelled as part of the impulse response function instead. In terms of a lithographic system model, this assumption is reasonable as the beam divergence is low in excimer beams, and it is usually collimated by the illumination optics. Secondly, it was assumed that the elementary function $f(\mathbf{r})$ is real and even. For real and even cross-spectral densities this should not cause any problems. The third assumption was that $\hat{f}^2(\mathbf{u})$ has no zeros. This may be the most restricting assumption, limiting the method to cross-spectral densities made up of smooth envelope functions such as Gaussians or super-Gaussians. For the case of the sources modelled in this project these assumptions are not debilitating. Other source types may stretch the boundaries of the method, but testing this was beyond the scope of this study. A standard test to generate the intensity in the image of an opaque edge produced results in the order of that expected. The point spread function of the system was such that the partially coherent source produced an image consistent with that expected from a spatially coherent source. However, a model of a partially coherent test beam (Gaussian Schell-Model beam) showed that as the spatial coherence of the source decreased, the intensity in the image of an opaque edge moved towards the expected partially coherent result. We can conclude that the model developed using the Elementary Function Method has provided a good approximation of a partially coherent excimer source, but that sampling issues are still not fully understood.

Beam Homogenization

The second part of the simulation work involved development of a model of an imaging homogenizer. Homogenization optics are used in lithographic systems due to the uneven distribution of intensity in the raw excimer beam and the need for high level uniformity in lithographic exposures. Beam homogenization by means of integration involves splitting the input beam into subapertures. These subapertures are focused by microlenses in a common focal plane. The spherical waves originating from the focal spots are re-collimated by a primary lens and overlapped in the target plane. An imaging homogenizer consists of two microlens arrays (identical, in this case) and a Fourier lens. The incident wavefront is divided into beamlets by the first microlens array, and the second lens array, along with the Fourier lens, images the beamlets onto the image plane of the system where they are overlapped. This process was simplified using a special case of multiple beam interference, where the aperture size and separation are the same, and are set equal to the pitch of the lenslets.

This method provided a good approximation of the physical process taking place in a beam homogenizing system. A coherent point source on-axis was imaged through the system, and a high level of interference was observed in the image plane, consistent with what is expected from a coherence source. Increasing the width of the source led to a decrease in interference effects consistent with decreasing spatial coherence, which is also physically realistic. The Elementary Function Method was applied to the homogenizer model in two different ways. By keeping the coherence width σ_G fixed and slowly increasing the intensity width σ_I, the degree of coherence of the source was effectively decreased, and the interference effects in the image plane of the homogenizer gradually decreased to produce a more uniform intensity profile. Similar interference effects were observed when the intensity width of the source was fixed, and the coherence width was slowly decreased. The relationship between the FWHM of the coherence distribution and the intensity distribution is key to predicting the level of fluctuations in the intensity profile post-homogenization.

The spatial coherence data and the intensity data from the 248 *nm* beam were tested in three different ways. Firstly, a source (filled with delta functions) was defined with a width corresponding to the intensity width of the excimer source along the x- and y- axes. The simulated intensity in the image plane of the Fourier lens was quite uniform, with only slight fluctuations visible in the intensity distribution (see Figure 6.12). Secondly, the model was applied to a Gaussian Schell-model beam with the same intensity and coherence widths as the 248 *nm* beam. The fluctuations in the intensity distribution in the homogenizer image plane were considerably large for the short, and more coherent, axis of the beam. A more effective smoothing was seen along the less coherent long axis. Finally, the model was applied to the real 248 *nm* excimer beam. The result was similar to the Gaussian Schell-model case: the coherence properties along the short axis of the beam produce dramatic interference effects in the image plane of the homogenizer, while a smoother intensity profile is produced when the homogenizer is illuminated with the long axis of the beam. From these simulations it is clear that the coherence properties of the excimer beam give rise to undesirable interference effects in the image plane of an imaging homogenizer, which, without further averaging, would greatly hinder high resolution circuit printing.

While the Elementary Function Method has proven to be a simple and efficient tool in the task of partially coherent light propagation, it is possibly not as useful when applied to an imaging homogenizer model. Ideally, an imaging homogenizer should be used on a completely incoherent source; even a small level of spatial coherence in the source produces significant interference effects in the homogenizer image plane. The partially coherent excimer source modelled here has quite a large degree of coherence, and use of the Elementary Function Method to propagate this source through the imaging homogenizer is more complicated than necessary. An estimate of the interference effects one can expect at the output of an imaging homogenizer could be provided equally well by propagating a series of delta functions through the system. However, the Elementary Function Method could prove useful in other areas of imaging with partially spatially coherent sources.

Bibliography

[1] International technology roadmap for semiconductors: http://www.itrs.net/.

[2] G. E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38 (8), 1965.

[3] K. Jain, C.G. Willson, and B. J. Lin. Ultrafast high-resolution contact lithography with excimer lasers. *IBM J. Res. Develop.*, 26(2):151–159, 1982.

[4] W. Muckenheim and B. Ruckle. Excimer laser with narrow linewidth and large internal beam divergence. *J. Phys. E: Sci. Instrum*, 20:1394–1396, 1987.

[5] R. Patzel, I. Bragin, J. Kleinschmidt, U. Rebhan, and D. Basting. Excimer laser with high repetition rate for DUV lithography. *Microelect. Eng.*, 30:165–167, 1996.

[6] M. Rothschild, A. R Forte, M. W. Horn, R. R. Kunz, S. C. Palmateer, and J. H. C. Sedlacek. 193-nm lithography. *IEEE (Quantum Elect)*, 1(3):916–923, 1995.

[7] A. K. Bates, M. Rothschild, T. M. Bloomstein, T. H. Fedynyshyn, R. R. Kunz, V. Liberman, and M. Switkes. Review of technology for 157-nm lithography. *IBM J. Res. & Dev.*, 45(5):605–614, 2001.

[8] Harry J. Levinson. *Principles of Lithography (Second Edition)*. SPIE Press, 2005.

[9] K. Jain, C.G. Willson, and B.J. Lin. Ultrafast deep UV lithography with excimer lasers. *IEEE Electron Device Letters*, EDL-3 (3):53–55, 1982.

[10] K. Jain. Lithography and photoablation systems for microelectronics and optoelectronics: Importance of laser beam shaping in system design. *Proc. SPIE*, 4770:1–12, 2002.

[11] W. B. Veldkamp. Laser beam profile shaping with interlaced binary diffraction gratings. *Applied Optics*, 21(17):3209–3212, 1982.

[12] W. B. Veldkamp. Technique for generating focal-plane flat-top laser-beam profiles. *Rev. Sci. Instr.*, 53(3):294–297, 1982.

[13] P. K. Katti and P. C. Mehta. Noise elimination technique in holography. *Applied Optics*, 15(2):530–534, 1976.

[14] W. H. Lee. Method for converting a Gaussian laser-beam into a uniform beam. *Opt. Comms.*, 36(6):469–471, 1981.

[15] Y. Ohtsuka. Acousto-optic intensity modification of a Gaussian laser beam. *Opt. Comms*, 39(1-2):70–74, 1981.

[16] A. Penzkofer and W. Frohlich. Apodizing of intense laser-beams with saturable dyes. *Opt. Comms.*, 28(2):197–201, 1979.

[17] M. Wagner, H. D. Geiler, and D. Wolff. High-performance laser beam shaping and homogenization system for semiconductor processing. *Meas. Sci. Technol.*, 1:1193–1201, 1990.

[18] C. Kopp, L. Ravel, and P. Meyrueis. Efficient beamshaper homogenizer design combining diffractive optical elements, microlens array and random phase plate. *J. Opt. A: Pure Appl. Opt*, 1:398–403, 1999.

[19] J. Turunen, P. Paakkonen, M. Kuittinen, P. Laakkonen, J. Simonen, T. Kajava, and M. Kaivola. Diffractive shaping of excimer laser beams. *J. Mod. Opt.*, 47(13):2467–2475, 2000.

[20] T Kajava, M Kaivola, J Turunen, P Paakkoned, M Kuittinen, P Laakkonen, and J Simonen. Excimer laser beam shaping using diffractive optics. *Laser and Electro-Optics Europe*, Conference Digest 2000:1, 2000.

[21] Y. Ozaki and K. Takamoto. Cylindrical fly's eye lens for intensity redistribution of an excimer laser beam. *Applied Optics*, 28(1):106–110, 1989.

[22] P. Dainesi, J. Ihlemann, and P. Simon. Optimization of a beam delivery system for a short-pulse KrF laser used for material ablation. *Applied Optics*, 36(27):7080–7085, 1997.

[23] X. Deng, X. Liang, Z. Chen, W. Yu, and R Ma. Uniform illumination of large targets using a lens array. *Applied Optics*, 25(3):377–381, 1986.

[24] Y. Lin, G. N. Lawrence, and J. Buck. Characterization of excimer lasers for application to lenslet array homogenizers. *Appl. Opt.*, 40(12):1931–1940, 2001.

[25] E. Tefouet Kana, S. Bollanti, P. Di Lazzaro, and D. Murra. Beam homogenization: theory, modelling and application to an excimer laser beam. *Proc. SPIE*, 5777:716–724, 2005.

[26] E. Tefouet Kana, S. Bollanti, P. Di Lazzaro, D. Murra, O. Bouba, and M. Boyomo Onana. Laser beam homogenization: Modelling and comparison with experimental results. *Optics Comms*, 264:187–192, 2006.

[27] J. L. Bunis, C. C. Abele, J. D. Campbell, and G. F. Caudle. Producing a uniform excimer laser beam for materials processing applications. *Proc. SPIE*, 1377:30–36, 1990.

[28] Fred M. Dickey, Scott C. Holswade, and David L. Shealy. *Laser Beam Shaping Applications*. Taylor & Francis, 2006.

[29] R. T. Kerth, K. Jain, , and M. R. Latta. Excimer laser projection lithography on a full-field scanning projection system. *IEEE Elect. Dev. Letters*, 7(5):299–301, 1986.

[30] K. Ogawa, M. Sasago, M. Endo, H. Nakagawa, Y. Tani, and T. Ishihara. Advanced KrF excimer laser stepper for half micron lithography. *J. Electrochem. Soc.: Solid State Sci. and Tech.*, 135(9):2347–2352, 1988.

[31] G. L. T. Chiu and J. M Shaw. Optical lithography: Introduction. *IBM J. Res. Develop.*, 41(1/2):3–6, 1997.

[32] A.C. Liu and B.J. Lin. A study of projected optical images for typical IC mask patterns illuminated by partially coherent light. *IEEE Trans. Elect. Dev.*, 30(10):1251–1263, 1983.

[33] W. T. Welford. Microlithography and the ultraviolet: experiments with an excimer laser. *Proc. R. Soc. Lond. A*, 419:173–180, 1988.

[34] X. Chen and S. R. J. Brueck. Imaging interferometric lithography: approaching the resolution limits of optics. *Optics Letters*, 249(3):124–126, 1999.

[35] T. Ito and S. Okazaki. Pushing the limits of lithography. *Nature*, 406:1027–1031, 2000.

[36] T. A. Brunner. Why optical lithography will live forever. *J. Vac. Sci. Technol. B*, 21(6):2632–2637, 2003.

[37] Max Born and Emil Wolf. *Principles of Optics*. Cambridge University Press, 1999.

[38] D. C. Cole, E. Barouch, and U. Hollerbach. Derivation and simulation of higher numerical aperture scalar aerial images. *Jpn. J. Appl. Phys.*, 31:4110 – 4119, 1992.

[39] K. Adam, Y. Granik, A. Torres, and N. Cobb. Improved modeling performance with an adapted vectorial formulation of the hopkins imaging equation. *Proc. SPIE*, 5040:78–91, 2003.

[40] M. Mansuripur. Distribution of light at and near the focus of high-numerical-aperture objectives. *J. Opt. Soc. Am. A*, 3 (12):2086 – 2093, 1986.

[41] M. S. Yeung, D. Lee, R. Lee, and A. R. Neureuther. Extension of the hopkins theory of partially coherent imaging to include thin-film interference effects. *Proc. SPIE*, 1927:452 – 463, 1993.

[42] D. G. Flagello. Theory of high-na imaging in homogeneous thin films. *J. Opt. Soc. Am. A*, 13 (1):53 – 64, 1996.

[43] E. Verdet. *Ann. Scientif. l'Ecole Normale Superieure*, 2:291, 1865.

[44] J. W. Strutt (Lord Rayleigh). On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Philos. Mag.*, 10:73–78, 1880.

[45] F. Zernike. The concept of degree of coherence and its application to optical problems. *Physica*, 5:785, 1938.

[46] L. Mandel and E. Wolf. Coherence properties of optical fields. *Rev. Mod. Phys.*, 37 (2):231–287, 1965.

[47] E. Wolf. New theory of partial coherence in the space-frequency domain. Part I: spectra and cross spectra of steady state sources. *J. Opt. Soc. Am.*, 72 (3):343 – 351, 1982.

[48] Leonard Mandel and Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995.

[49] Joseph W. Goodman. *Statistical Optics*. John Wiley & Sons, 2000.

[50] M. J. Bastiaans. A frequency-domain treatment of partial coherence. *Optica Acta*, 24 (3):261 – 274, 1977.

[51] A. T. Friberg and E. Wolf. Relationships between the complex degrees of coherence in the space-time and in the space-frequency domains. *Optics Letters*, 20 (6):623 – 625, 1995.

[52] A. C. Schell. A technique for the determination of the radiation pattern of a partially coherent aperture. *IEEE. Trans. Antennas Propag.*, 15 (1):187 – 188, 1967.

[53] B. E. A. Saleh and M. Rabbani. Simulation of partially coherent imagery in the space and frequency domains and by modal expansion. *Appl. Opt.*, 21(15):2770–2777, 1982.

[54] I. P. Christov. Propagation of partially coherent light pulses. *Optica Acta*, 33 (1):63 – 72, 1986.

[55] R. Castañeda and F. F. Medina. Partially coherent imaging with Schell-model beams. *Optics & Laser Technology*, 29 (4):165 – 170, 1997.

[56] R. Castañeda and Z. Jaroszewicz. Determination of the spatial coherence of Schell-model beams with diffraction gratings. *Opt. Comms.*, 173:115–121, 2000.

[57] J. Garcia-Sucerquia and R. Castañeda. Full retrieving of the complex degree of spatial coherence: theoretical analysis. *Opt. Comms.*, 228:9–19, 2003.

[58] A. Starikov and E. Wolf. Coherent-mode representation of Gaussian Schell-model sources and their radiation fields. *J. Opt. Soc. Am.*, 72:923 – 928, 1982.

[59] J. Deschamps, D. Courjon, and J. Bulabois. Gaussian Schell-model sources: an example and some perspectives. *J. Opt. Soc. Am.*, 73 (3):256 – 261, 1983.

[60] F. Gori, M. Santarsiero, R. Simon, G. Piquero, R. Borghi, and G. Guattari. Coherent-mode decomposition of partially polarized, partially coherent sources. *J. Opt. Soc. Am. A*, 20 (1):78–84, 2003.

[61] Y. Cai and Q. Lin. Partially coherent flat-topped multi-Gaussian-Schell-model beam and its propagation. *Opt. Comms.*, 239:33 – 41, 2004.

[62] J. Turunen and P. Vahimaa. Independent-elementary-field model for three-dimensional spatially partially coherent sources. *Optics Express*, 16 (9):6433–6442, 2008.

[63] E. Wolf. *Introduction to the Theory of Coherence and Polarisation of Light*. Cambridge U. Press, 2007.

[64] M. Wald, M. Burkhardt, A. Pesch, H. Gross, and J. Greif. Design of a microscopy illumination using a partial coherent light source. *Proc. SPIE*, 5962:420–429, 2005.

[65] A. Burvall, A. Smith, and C. Dainty. Elementary functions: Propagation of partially coherent light. *Journal of Optical Society of America A*, 26(7):1721–1729, 2009.

[66] Joseph W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1996.

[67] H. H. Hopkins. The concept of partial coherence in optics. *Proc. Royal Soc. London (A)*, 208:263–277, 1951.

[68] H. H. Hopkins. On the diffraction theory of optical images. *Proc. Royal Soc. London (A)*, 217:408–432, 1953.

[69] F. Gori. Directionality and spatial coherence. *Optica Acta*, 27 (8):1025 – 1034, 1980.

[70] E Wolf. New spectral representation of random sources and of the partially coherent fields that they generate. *Opt. Commun.*, 38:3–6, 1981.

[71] P Vahimaa and J Turunen. Finite-elementary source model for partially coherent radiation. *Opt. Express*, 14(4):1376–1381, 2006.

[72] F. Gori and C. Palma. Partially coherent sources which give rise to highly directional light beams. *Opt. Commun.*, 27:185–187, 1978.

[73] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749 – 759, 1932.

[74] A. Walther. Radiometry and coherence. *J. Opt. Soc. Am.*, 58(9):1256–1259, 1968.

[75] M. J. Bastiaans. Wigner distribution functoin and its application to first-order optics. *J. Opt. Soc. Am.*, 69 (12):1710 – 1716, 1979.

[76] M. J. Bastiaans. The Wigner distribution function of partially coherent light. *J.Mod. Opt.*, 28 (9):1215 – 1224, 1981.

[77] M.J. Bastiaans. Application of the Wigner distribution function to partially coherent light. *J. Opt. Soc. Am. A*, 3:1227–1238, 1986.

[78] P. Vahimaa and J. Turunen. Independent-elementary-pulse representation for non stationary fields. *Opt. Express*, 14(12):5007–5012, 2006.

[79] P. Vahimaa and J. Turunen. Finite elementary-field model for three-dimensional spatially partially coherent sources. *Opt. Express*, 16:6433–6442, 2008.

[80] A T Friberg, H Lajunen, and V Torres-Company. Spectral elementary-coherence-function representation for partially coherent light pulses. *Opt. Express*, 15(8):5160–5162, 2007.

[81] M Unser. Sampling - 50 years after Shannon. *Proc. IEEE*, 88:569–587, 2000.

[82] K A Nugent and J E Trebes. Coherence measurement technique for short-wavelength light sources. *Rev. Sci. Instrum.*, 63(4):2146–2151, 1992.

[83] R. H. Dicke. Scatter-hole cameras for x-rays and gamma rays. *Astrophys. J.*, 153 (2P2):L101, 1968.

[84] E. E. Fenimore and T. M. Cannon. Coded aperture imaging with uniformly redundant arrays. *Applied Optics*, 17 (3):337 – 347, 1978.

[85] Y Mejía and A I González. Measuring spatial coherence by using a mask with multiple apertures. *Optics Communications*, 273:428–434, 2007.

[86] E Wolf. A macroscopic theory of interference and diffraction of light from finite sources. *Proc. Roy. Soc. Lond. A*, 225:96–111, 1954.

[87] M. J. E. Golay. Point arrays having compact, nonredundant autocorrelations. *J. Opt. Soc. Am.*, 61 (2):272 – 273, 1971.

[88] A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analy. Chem.*, 36 (8):1627 – 1639, 1964.

[89] C. Solomon and T. Breckon. *Fundamentals of Digital Image Processing*. Wiley-Blackwell, 2011.

[90] B.M. Watrasiewicz. Theoretical calculations of images of straight edges in partially coherent illumination. *Journal of Modern Optics*, 12(4):391–400, 1965.

[91] H. H. Hopkins. Applications of coherence theory in microscopy and interferometry. *J. Opt. Soc. Am.*, 47(6):508–526, 1957.

[92] M. Zimmermann, N. Lindlein, R. Voelkel, and K. J. Weible. Microlens laser beam homogenizer - from theory to application. *Proc. SPIE*, 6663:666302, 2007.

[93] K. Mann, A. Hopfmüller, P. Gorzellik, R. Schild, W. Stöffler, H. Wagner, and G. Wolbold. Monitoring and shaping of excimer laser beam profiles. *Proc. SPIE*, 1834:184–194, 1992.

[94] Alexander Büttner and Uwe D. Zeitner. Wave optical analysis of light-emitting diode beam shaping using microlens arrays. *Opt. Eng.*, 41 (10):2393–2401, 2002.

[95] D. M. Brown, F. M. Dickey, and L. S. Weichman. *Laser beam shaping theory and techniques*. Marcel Dekker Inc, 2000.

[96] E. Hecht. *Optics*. Addison Wesley, 2002.

[97] A I González and Y Mejía. Nonredundant array of apertures to measure the spatial coherence in two dimensions with only one interferogram. *J. Opt. Soc. Am. A*, 28 (6):1107 – 1113, 2011.