STARS

Electronic Theses and Dissertations, 2004-2019

2009

Design, Analysis, And Optimization Of Diffractive Optical Elements Under High Numerical Aperture Focusing

Toufic Jabbour
University of Central Florida

DESIGN, ANALYSIS, AND OPTIMIZATION OF
DIFFRACTIVE OPTICAL ELEMENTS
UNDER HIGH NUMERICAL APERTURE FOCUSING

by

TOUFIC G. JABBOUR
B.S. University of Balamand, 2000
M.S. Oklahoma State University, 2002

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in CREOL, The College of Optics and Photonics
at the University of Central Florida
Orlando, Florida

Summer Term
2009

Major Professor: Stephen M. Kuebler

ABSTRACT

The demand for high optical resolution has brought researchers to explore the use of beam shaping diffractive optical elements (DOEs) for improving performance of high numerical aperture (*NA*) optical systems. DOEs can be designed to modulate the amplitude, phase and/or polarization of a laser beam such that it focuses into a targeted irradiance distribution, or point spread function (PSF). The focused PSF can be reshaped in both the transverse focal plane and along the optical axis. Optical lithography, microscopy and direct laser writing are but a few of the many applications in which a properly designed DOE can significantly improve optical performance of the system.

Designing DOEs for use in high-*NA* applications is complicated by electric field depolarization that occurs with tight focusing. The linear polarization of off-axis rays is tilted upon refraction towards the focal point, generating additional transverse and longitudinal polarization components. These additional field components contribute significantly to the shape of the PSF under tight focusing and cannot be neglected as in scalar diffraction theory. The PSF can be modeled more rigorously using the electromagnetic diffraction integrals derived by Wolf, which account for the full vector character of the field.

In this work, optimization algorithms based on vector diffraction theory were developed for designing DOEs that reshape the PSF of a 1.4-*NA* objective lens. The optimization techniques include simple exhaustive search, iterative optimization (Method of Generalized Projections), and evolutionary computation (Particle Swarm Optimization). DOE designs were obtained that can reshape either the transverse PSF or

the irradiance distribution along the optical axis. In one example of transverse beam shaping, all polarization components were simultaneously reshaped so their vector addition generates a focused flat-top square irradiance pattern. Other designs were obtained that can be used to narrow the axial irradiance distribution, giving a focused beam that is superresolved relative to the diffraction limit. In addition to theory, experimental studies were undertaken that include (1) fabricating an axially superresolving DOE, (2) incorporating the DOE into the optical setup, (3) imaging the focused PSF, and (4) measuring aberrations in the objective lens to study how these affect performance of the DOE.

ACKNOWLEGMENT

The research described in this dissertation was possible due to much more than just an individual effort. This work would have been impossible to accomplish without the generous, inspiring and supportive contributions of many individuals that I met during my graduate studies. I sincerely appreciate every input, direct and indirect, to my scholarly and scientific development.

I would like to express the deepest appreciation to my advisor, mentor and friend Dr. Stephen Kuebler, who never ceased to astonish me with his intellect and expertise. With an amazing teaching style and patience, Dr. Kuebler taught me to conduct research rigorously and systematically, and to ask the right technical questions as a way to tackle scientific problems. I would not have gotten as far without his thoughtful criticism and continuous support.

I am grateful to Dr. M J Soileau, Dr. Eric Van Stryland and Dr. David Hagan for maintaining an outstanding academic program in Optics and Photonics that I am proud to have completed.

I thank my PhD committee, Dr. Demetrios Christodoulides, Dr. Florencio Eloy Hernandez and Dr. M. J. "Jim" Moharam, who provided me with unlimited time and attention to discuss my projects and provide insightful suggestions.

I would also like to acknowledge Dr. Patrick LiKamWa and his graduate students Nathan Bickel and Tony Ho for training me on the cleanroom equipment so I could fabricate the diffractive optical element. A special thanks goes to Dr. Pieter Kik for

I would be remiss not to thank the CREOL staff for their genuine readiness to assist me with administrative tasks. The staff members treated me like family and were there to help and support me in times of need. It is difficult of overstate my gratitude to Amy Perry, Denise Whiteside, Rachel Franzetta, Gail Drabczuk, Courtney Lewis, Maria Lopez, Deon Frank, Michael Torrance, Richard Zotti and Mark Wagenhauser.

I also would like to thank my colleagues at CREOL for their continuous support during the many years I spent with them: Ramy El-Ganainy, Georgios Siviloglou, Christopher Brown, Clarisse Mazuir, Supraja Murali, Leo Siiman, David Haefner, Abdullah Demir, Ozan Cakmakci, Sergiy Suntsov, and Grady Webb-Wood.

I owe a special debt to my invaluable network of caring, loyal, bighearted and loving friends without whom I could not have survived the PhD process: Feras Batarseh, Monica Dantes, Claudia Oda, Ana Dragan, Livia Dragan, Rania Ghatas, Leen Hanayneh, and Satyajeet Sharma.

Finally and most importantly, I would like to thank my father Gerges Jabbour and mother Wadad Kamal Jabbour; my brother Tony and my sister Carole for their unconditional love and support during the long years of graduate school.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
2D-FTM	Two dimensional Fourier transform method
AFM	Atomic force microscopy
APSF	Amplitude point spread function
BPSO	Binary particle swarm optimization
CCD	Charged coupled device
CZT	Chirp-z transform
DLW	Direct laser writing
DOE	Diffractive optical element
EBL	Electron beam lithography
FOV	Field of View
FWHM	Full-width at half-maximum
IL	Immersion liquid
IO	Immersion oil
IPA	Isopropanol

IPSF	Imaged point spread function
MGP	Method of generalized projections
MIBK	Methyl isobutyl ketone
mpDLW	Multi-photon direct laser writing
NA	Numerical aperture
NMSE	Normalized mean square error
NSOM	Near field scanning optical microscope
PMMA	Poly(methylmethacrylate)
PSF	Point spread function
PSO	Particle swarm optimization
S/N	Signal to noise ratio
SEM	Scanning electron microscope
TIR	Total internal reflection
TPE	Two-photon excitation

CHAPTER 1: INTRODUCTION

Beam shaping diffractive optical elements (DOEs) are becoming an essential part of many optical systems. A DOE can be used to modify the amplitude, phase, and polarization of an incident beam so that it focuses into a targeted irradiance distribution at the image plane [1]. Beam shaping can enhance performance in optical lithography, laser-based materials processing, direct laser writing, surgical applications, and optical data storage [2]. For many applications, the focal intensity distribution is obtained by focusing the laser beam via a high numerical aperture (*NA*) lens system. With such focusing systems the polarization of the incident field experiences a significant tilt with respect to the transverse plane as the light rays are refracted towards the focal point. This results in a strong longitudinal focal field component, a phenomenon known as depolarization [3]. The scalar theory of diffraction does not account for the depolarization effect and thus cannot be used to rigorously model high-*NA* systems. The most widely used and appropriate theory for modeling the focal field distribution of high *NA* systems is the electromagnetic diffraction integrals derived by Wolf [4] and extended by [5].

Several robust and efficient methods have been developed to design DOEs that can reshape the focal intensity distribution [1, 3, 6-42]. Very few of these algorithms [1, 3, 6, 7, 12, 30, 34, 38] incorporate the electromagnetic diffraction theory, so the vast majority is only valid in the paraxial domain of diffractive optics. Here, we report optimization algorithms for designing DOEs that reshape the focal intensity under high *NA* focusing and include the theory of electromagnetic diffraction. We also describe the experimental

incorporation of a DOE into a high *NA* optical system and characterize its performance in the presence of optical aberrations.

The dissertation is organized in seven chapters. Chapter two describes the application of vector diffraction theory for studying the effect of two- and three-zone annular DOEs on the three-dimensional point-spread-function (PSF) that results when linearly polarized light is focused using a high-*NA* refractory lens. Conditions are identified for which a three-zone DOE generates a PSF that is axially superresolved by 19% with minimal change in the transverse profile and sufficiently small side lobes that the intensity pattern could be used for advanced photolithographic techniques, such as multi-photon direct laser writing, as well as multi-photon imaging. Conditions are also found for which a three-zone DOE generates a PSF that is axially elongated by 510% with only 1% broadening in the transverse direction. This intensity distribution could be used for sub-micron-scale laser drilling and machining [43].

Chapter three describes a new approach for designing pupil phase DOEs that modify the focused axial PSF under high-*NA* aperture conditions. The approach is based on the method of generalized projections (MGP) with the theory of electromagnetic diffraction incorporated to account for non-paraxial focusing and the full vector character of the field. The procedure is applied to the design of a pupil filter that superresolves the axial intensity distribution with controlled side-lobe peak intensity. It is shown that the solutions obtained depend strongly on the starting pupil function. Methods are described and implemented to generate a systematic set of starting conditions that enable a more thorough search of the solutions space. Several satisfactory solutions are obtained, including one for which the central lobe of the PSF is axially narrowed by 29% while

2

maintaining the axial side lobes at or below 52% of the peak intensity. A comparative study shows how the solutions obtained also depend subtly on the starting constraints [44, 45].

In chapter four a particle swarm optimization (PSO) algorithm is described which can be used to design binary phase-only DOEs that superresolve the axially focused PSF. Similarly, the method is based on vector diffraction theory to ensure solutions are valid under high-*NA* conditions. A DOE is identified that superresolves the focal spot by 34% and maintains the side lobes below 50% of the peak intensity. The algorithm was used to obtain the *Pareto front* of the fitness-value space, which describes the achievable superresolution versus an allowed upper bound in side lobe intensity. The results suggest that the algorithm yields solutions that are global in terms of the co-optimized fitness values G and M [46].

In chapter five, an algorithm is reported for the design of a phase-only DOE that reshapes the transverse profile of a beam focused using a high-*NA* lens. The vector diffraction integrals are used to relate the field distributions in the DOE plane and focal plane. The integrals are evaluated using the chirp-z transform and computed iteratively within the MGP to identify a solution that simultaneously satisfies the beam shaping and DOE constraints. The algorithm is applied to design a DOE that transforms a circularly apodized flat-top beam of wavelength λ to a square irradiance pattern when focused using a 1.4-*NA* objective. A DOE profile is identified that generates a $50\lambda \times 50\lambda$ square irradiance pattern having 7% uniformity error and 74.5% diffraction efficiency (fraction of focused power). The diffraction efficiency and uniformity decrease as the size of the

focused profile is reduced toward the diffraction limited spot size. These observations can be understood as a manifestation of the uncertainty principle [47].

In chapter six, the vectorial theory was used to study the effect of experimental imperfections on the DOE designed in chapter three. Such imperfections include fabrication errors, surface quality variation, and optical misalignment. The analyses of superresolution properties G and M as a function of experimental errors provided tolerance margins required to properly choose the fabrication technique and the type of experimental setup to implement the experiment. The analysis also provides a theoretical basis for understanding degradation in performance due to experimental errors.

Chapter seven outlines the experimental work necessary to incorporate the axially superresolving DOE designed using the MGP algorithm into the optical system and characterize its performance. The experimental procedures included (1) fabricating the DOE, (2) integrating the DOE into the optical (3) mapping the PSF with and without the DOE and (4) characterizing the objective lens. It was observed that the presence of aberration in the optical system can significantly degrade the theoretically predicted performance of the DOE. The primary difference between the theoretical and experimental axial PSFs lies in the side-lobe regions. A 24% enhancement was achieved in the central-lobe; however, the side-lobe peak intensity was 2.6 times larger than that predicted by theory. By carefully analyzing the optical system, it was determined that this discrepancy is due to aberrations in the objective lens. A Mach-Zehnder interferometer was utilized to measure and identify the type of aberrations present in the objective. The phase front error measured at the entrance pupil was $\lambda/2.8$ at $\lambda = 532$nm. The Zernike polynomial decomposition of the measured wavefont revealed the type of

aberrations that were not corrected for in the objective lens design. A theoretical study of how the measured aberrations affect the axial PSF shows that secondary spherical aberration is the main source for the discrepancy observed between theory and experiment.

Chapter eight outlines theoretical and experimental routes in which the work in this dissertation can be extended.

CHAPTER 2: VECTOR DIFFRACTION ANALYSIS OF HIGH NUMERICAL APERTURE FOCUSED BEAMS MODIFIED BY TWO- AND THREE-ZONE ANNULAR MULTI-PHASE PLATES

2.1 Introduction

Performance can be improved in many optical applications by engineering the focused three-dimensional (3D) intensity distribution, or point-spread-function (PSF), using diffractive optical elements (DOEs). DOEs are passive components that can be placed in the pupil plane of an optical system to alter the amplitude, phase, and polarization of the light prior to focusing [1]. The resolution achieved in scanned-laser imaging techniques, such as confocal and multi-photon microscopy, is determined by the transverse and axial extent of the central (most intense) lobe of the diffraction-limited PSF [48]. DOE designs have been reported that can decrease the lateral or axial extent of the central lobe, and this is frequently referred to as superresolving [32]. In other applications it is beneficial to elongate the PSF axially so the focused intensity remains peaked over a greater depth of field [11, 37]. This enables uniform laser cutting of topographically complex work pieces, such as corrugated steel, and improved signal-to-noise in some optical data storage and read-out schemes [49].

DOE design and performance have been examined computationally and experimentally [50, 51]. Radially symmetric amplitude-only and phase-only DOEs exhibiting a quantized profile (see Fig. 2.1) have attracted the most attention because they are structurally simple and relatively easy to fabricate using commonly available manufacturing and replication techniques [52]. Rotationally symmetric phase only DOEs

offer a significant advantage over amplitude-only DOEs in that they can deliver a greater fraction of incident optical power to the sample.

Computational approaches involving various levels of approximation have been applied to study and design DOEs that affect the PSF in the vicinity of the focus. DOEs have been designed using various methods, including satisfaction of constraints [53-55], solving for zeros of the PSF [26], parameter property optimization [8, 10, 13, 35, 56], and extensive search [57]. Most investigations employ the scalar, paraxial approximation [3, 10, 11, 13, 24, 25, 27, 33, 35]. Scalar approaches are mainly based on the Fresnel model of diffraction. In some reports the second-order approximation of the focal intensity distribution was used to derive analytical expressions for the DOE-modified PSF [10]. These approximations greatly decrease computation time, but forfeit information concerning the vector character of the field. It is well known that under high numerical aperture (*NA*) conditions, rays refracted near the periphery of the limiting aperture have a non-zero longitudinal field component (component parallel to the direction of propagation). This contribution to the overall intensity distribution is unaccounted for in scalar methods, so they do not accurately model focusing and DOE performance in a high-*NA* configuration. A detailed analysis of binary- and multi-phase annular filters in the scalar limit has been reported by Sales and Morris [32].

Accurate modeling of high-*NA* focusing can be achieved using vector diffraction theory, which is equally well known as electromagnetic diffraction theory. The vector diffraction integrals derived by Wolf [4], and later extended to radially symmetric
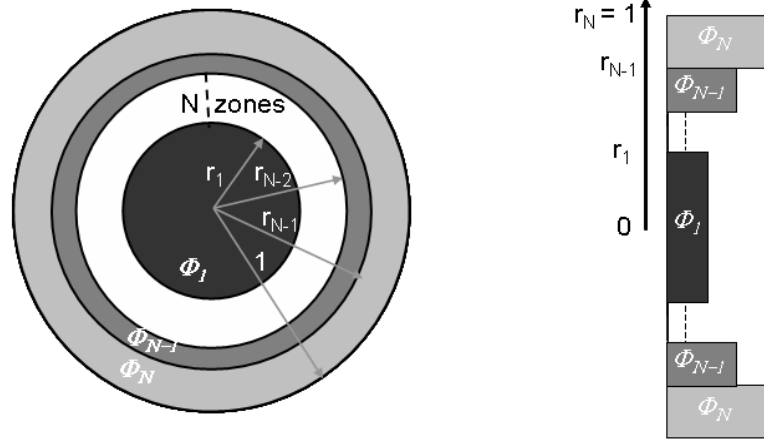
Figure 2.1. Front (left) and profile (right) views of an annular multi-phase DOE. The parameters Φ_i and r_i represent the differential phase transmittance and fractional radius of the ith annular zone, respectively.

systems by Richards and Wolf [5], provide a means for directly computing the intensity distribution around the focus for an optical system that includes a phase aberration. Vector diffraction methods have been employed to examine some specific DOE configurations. Sheppard *et al*. applied this method to study how an amplitude DOE alters the transverse intensity distribution [34]. Martínez-Corral *et al*. used the vector diffraction method to study axial superresolution achieved using amplitude-only DOEs [3].

The DOE-modified PSF has commonly been regarded as superresolved in the axial/transverse direction if the separation between local minima adjacent to the primary lobe is decreased relative to that of the diffraction limited intensity distribution. Particularly in studies using vector diffraction methods, little attention has been paid to changes in the relative intensities of the main lobe and adjacent side lobes. It is known that DOEs alter the PSF such that the intensities of side lobes and local minima adjacent

to the central lobe can become large and non-negligible in many applications. This is particularly true when the photo-response depends upon an absolute threshold intensity, as is the case for photolithographic techniques. Thus, more complete knowledge of the full 3D-PSF is required before a particular DOE design can be regarded as useful for a given application.

In this work, the Richard-Wolf integrals are applied to evaluate the entire solutions space of two- and three-zone DOEs and their effect on the PSF generated under high-NA refractive focusing. Emphasis is placed on characterizing changes in the axial extent of the central lobe and changes in the relative intensity of side lobes. These characteristics are most relevant to multi-photon imaging techniques, multi-photon 3D microfabrication, and optical data storage and read-out schemes.

2.2 Method and theory

Richards and Wolf formulated an integral representation of the electromagnetic field formed in the image space of an aplanatic optical system that images a point source located at infinity in the object space [5]. This theory is well suited for modeling the effect of DOEs on the focused PSF under high-NA conditions. The optical geometry is depicted in Fig. 2.2. An N-zone DOE and an aberration free lens (or lens system) are positioned such that their optical axes are collinear with the z-axis of a cylindrical coordinate system whose origin is located at the Gaussian focus of the lens. The numerical aperture is $NA = 1.4$ in all calculations, unless otherwise stated. Monochromatic linearly polarized plane waves, with electric field vector parallel to the x-axis, propagate along the z-axis, passing through the DOE and entering the pupil of the lens. The light focuses into a medium of refractive index $n = 1.5$. In the absence of the

DOE, the situation is consistent with common applications of high-*NA* oil-immersion objective lenses.
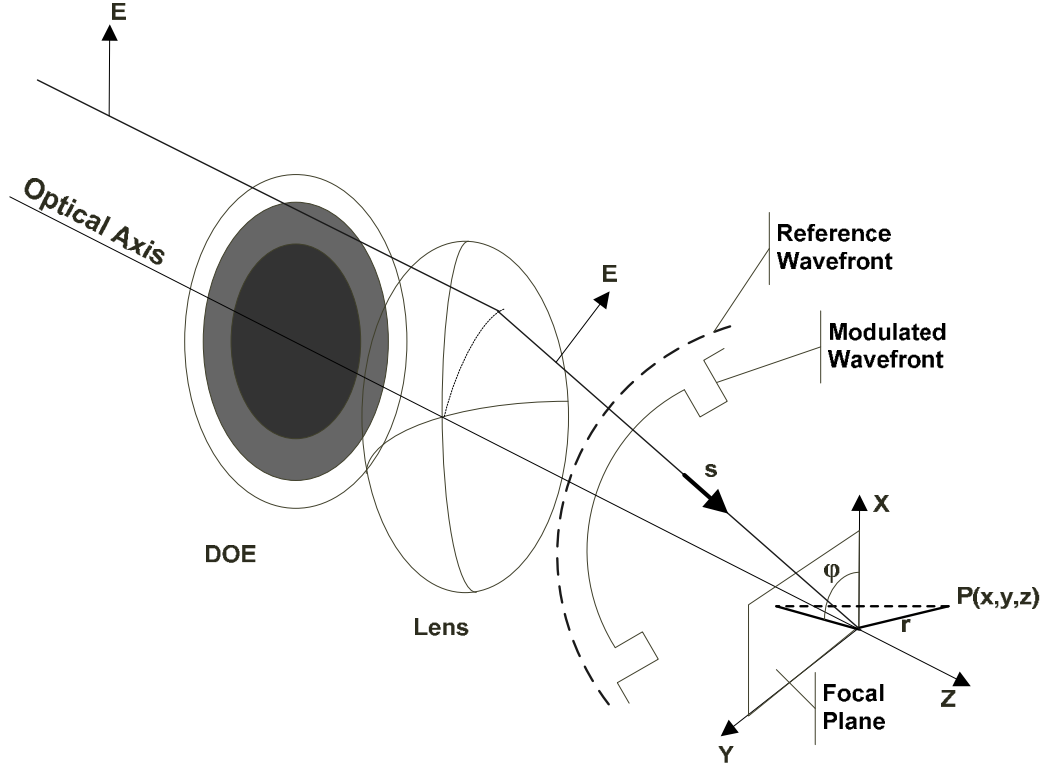
Figure 2.2. Optical geometry in which an DOE is used to modify the phase front and resulting PSF of a focused optical beam.

The electric field at point $P(x, y, z)$ in the neighborhood of the focus may be expressed in the cylindrical optical coordinate system $[u, v, \varphi]$ as

$$
\begin{aligned}
E_x(u,v,\varphi) &= -iA\left(I_0 + I_2 \cos 2\varphi\right) \\
E_y(u,v,\varphi) &= -iAI_2 \sin 2\varphi \\
E_z(u,v,\varphi) &= -2AI_1 \cos \varphi
\end{aligned}
\tag{2.1}
$$

The intensity at P is $I \propto |E_x + E_y + E_z|^2$, and the PSF is a spatial map of intensity for all values of $[u, v, \varphi]$ about the focus. φ is defined as the angle subtended by the electric

10

field vector of the incident field and the meridional plane in which the field is calculated. The constant $A = \pi l_0 f / \lambda$ is defined in terms of the focal length, f, the wavelength within the medium, λ, and l_o, which describes the amplitude distribution of the incident field. It is assumed that uniform amplitude plane waves impinge on the lens, so l_o is set to unity.

Equation (2.1) is expressed in terms of the cylindrical optical coordinates u and v:

$$u = kz \sin^2 \alpha$$
$$v = kr \sin \alpha$$

(2.2)

where z and r are the radial and axial coordinates, respectively, of the point in the original coordinate system. The maximum aperture angle, $\alpha = \arcsin(NA/n)$, is determined by the numerical aperture of the lens. The wave number $k = 2\pi/\lambda$. $I_{0,1,2}$ are integrals evaluated over the aperture half-angle θ as

$$I_0(u,v) = \int_0^\alpha t(\theta)\sqrt{\cos\theta}\,\sin\theta(1+\cos\theta)J_0\left(\frac{v\sin\theta}{\sin\alpha}\right)\exp\left(\frac{iu\cos\theta}{\sin^2\alpha}\right)d\theta$$

(2.3)

$$I_1(u,v) = \int_0^\alpha t(\theta)\sqrt{\cos\theta}\,\sin^2\theta J_1\left(\frac{v\sin\theta}{\sin\alpha}\right)\exp\left(\frac{iu\cos\theta}{\sin^2\alpha}\right)d\theta$$

(2.4)

$$I_2(u,v) = \int_0^\alpha t(\theta)\sqrt{\cos\theta}\,\sin\theta(1-\cos\theta)J_2\left(\frac{v\sin\theta}{\sin\alpha}\right)\exp\left(\frac{iu\cos\theta}{\sin^2\alpha}\right)d\theta$$

(2.5)

The DOE spatially modifies the phase of the wave front according to the spatial phase transfer function $t(\theta)$. In applying these formulae, the following approximations are implicit. (1) All inhomogeneous waves are ignored. (2) The Kirchoff boundary conditions are imposed, which is appropriate for DOEs having macroscopic features, as

11

considered here. (3) The Debye approximation is also applied, so only rays falling within the numerical aperture of the lens are considered [5, 58]. Note that the electric field distribution along the optical axis, $E(u,v = 0)$, only depends upon $I_0(u,v = 0)$ giving:

$$E(u, v = 0) = -iA \int_0^\alpha t(\theta)\sqrt{\cos\theta} \sin\theta(1 + \cos\theta)\exp\left(\frac{iu\cos\theta}{\sin^2\alpha}\right)d\theta \qquad (2.6)$$

The two- and three-zone DOEs investigated are comprised of a set of N concentric annular zones each having constant differential phase transmittance Φ_i (Fig. 2.1). The radial extent of the DOE, R, is matched to the limiting aperture of the lens. The radius of the ith zone may be expressed as a dimensionless fraction of R using $r_i = \sin(\theta_i)/\sin(\alpha)$, where θ_i is the aperture half-angle of the ith zone. The optical characteristics of a DOE are determined by the radius and relative phase of each zone. As such, the innermost zone may always be set to $\Phi_1 = 0$, and the others may be varied independently over the interval $[0, 2\pi]$. The number of independent degrees of freedom is then two for a two-zone DOE (r_1 and Φ_2, where $0 < r_1 < 1$), and it is four for a three-zone DOE (r_1, r_2, Φ_2, and Φ_3, where $0 < r_1 < r_2 < 1$).

The two- and four-dimensional solutions spaces associated with a two- and three-zone DOE, respectively, were discretized and the PSF was calculated using Eqs. (2.1) - (2.5) for each unique combination of zone radii and relative phases. The solutions space was evaluated using a coarse discretization of $\Delta\Phi = 2\pi/20$ and $\Delta r = 0.05$. Specific regions of interest were studied in greater detail as needed by decreasing $\Delta\Phi$ and Δr. The PSF for each set of DOE parameters was characterized relative to the diffraction limited pattern in

terms of (1) the axial (transverse) width of the central lobe; (2) the peak intensity; and (3) the intensity of the largest side lobe(s). The peak in the PSF pattern having the highest intensity was regarded as the central lobe. Under this definition, the central lobe is not necessarily centered at the Gaussian focus. The axial (transverse) extent of the central lobe was quantified using a superresolution factor, G, defined as the full-width at half-maximum (FWHM) of the central lobe divided by the same in the diffraction limited pattern. The axial Strehl ratio, S, is defined as the peak intensity of the central lobe normalized to that of the diffraction limited pattern. The relative intensity of the largest axial side lobe is quantified using the parameter, M, which is defined as the peak-intensity of the side lobe divided by that of the central lobe.

2.3 Results and discussion

Figure 2.3 summarizes the characteristic changes to the axial PSF that result when a two-zone DOE is placed in front of the lens. G, M, and S all exhibit the greatest variation as a function of r_1 along the line $\Phi_2 = \pi$, and the plots are symmetric about this line. G varies from a maximum of 2.31 ($\Phi_2 = \pi$ and $r_1 = 0.54$) to a minimum of 0.90 ($\Phi_2 = \pi$ and $r_1 = 0.76$). Thus, a two-zone DOE could be used to elongate the central lobe by as much as a factor of two. Where $G = 0.90$, the axial intensity distribution is comprised of two partly overlapping lobes of equal peak intensity, so $M = 1$. Given that there are two lobes in the intensity distribution, the PSF cannot reasonably be regarded as superresolved. This finding is consistent with that reported by Sales, who evaluated superresolution in terms of the separation of minima in the axial PSF in the confocal mode [33]. The absolute intensity of this central lobe pair is reduced relative to the central lobe of the diffraction limit in the amount $S = 0.35$.

13

Simulations of the PSF over the complete multi-dimensional solutions space for two- and three-zone DOEs both show that the largest variation in G and the greatest superresolution occurs when successive zones of the DOE differ in phase by π. The overall appearance of the PSF is determined by the vector sum of the electric field component of rays that converge near the focus. The greatest overall variation can be
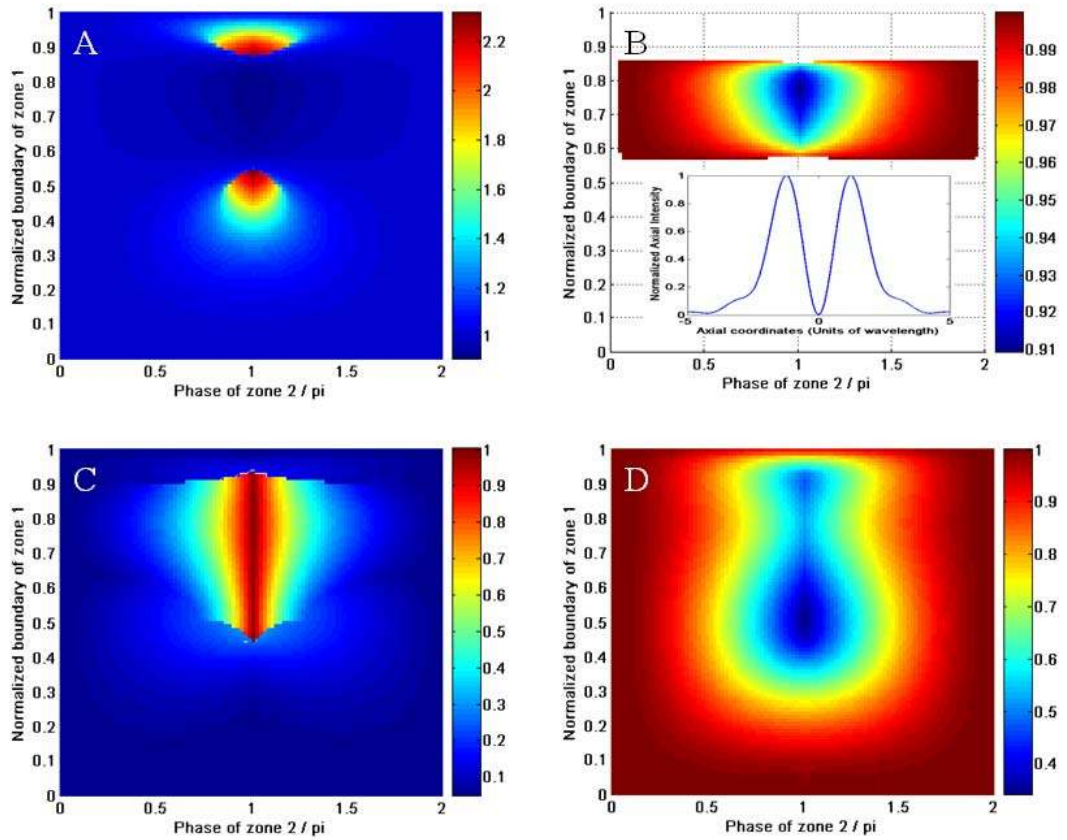
Figure 2.3. Characteristic changes to the axial PSF affected by a two-zone DOE. Shown are (**A**) the super-resolution factor G, (**B**) the sub-space of $G < 1$ on an expanded scale, (**C**) the side lobe intensity M, and (**D**) the Strehl ratio S versus $[r_1, \Phi_2]$. The two-dimensional solutions space was discretized by intervals of $\Delta\Phi = 2\pi/100$ and $\Delta r = 0.01$. The inset to B shows the normalized double-peaked axial distribution that results for $r_1 = 0.7$ and $\Phi_2 = \pi$.

expected then when rays recombine with the highest degree of destructive interference, or

when they successively differ in phase by π.

To study the axial superresolution that can be achieved with a 3-zone DOE in greater

detail, the zone phases were fixed to $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$, and the axial PSF was

simulated with the $[r_1, r_2]$ space discretized by $\Delta r = 0.01$. The corresponding plots of G,

M, and S versus r_1 and r_2 are shown in Fig. 2.4. The PSF characteristics are only defined
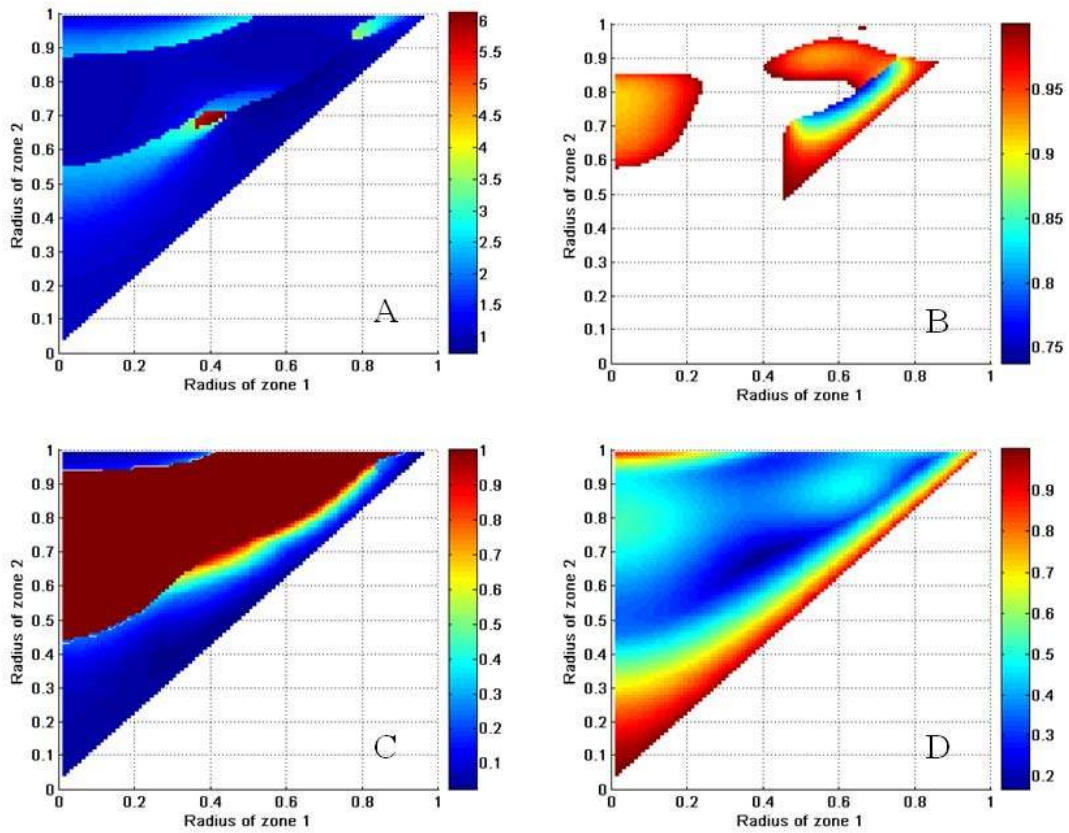
Figure 2.4. Characteristic changes to the axial PSF affected by a three-zone DOE having $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$ as a function of radial zone boundaries r_1 and r_2. Shown are (**A**) the superresolution factor G, (**B**) the sub-space $G < 1$ on an expanded scale, (**C**) the side lobe intensity M, and (**D**) the Strehl ratio, S. The two-dimensional solutions space $[r_1, r_2]$ was discretized by intervals of $\Delta r = 0.01$.

15

for the upper-left half of the $[r_1, r_2]$ space due to the constraint $0 < r_1 < r_2 < 1$. G takes a minimum value of 0.73 at $r_1 = 0.60$ and $r_2 = 0.77$; however, M is approximately unity under these conditions because the axial intensity distribution near the focus actually consists of three lobes having nearly the same peak intensity. This is similar to the circumstances under which G is minimized for a two-zone DOE.

The vector diffraction calculations show that in general increased axial superresolution of the central lobe is strongly correlated with a decrease in the Strehl ratio and an increase in the intensity of the side lobes. Similar conclusions have been reported previously for scalar studies of axial superresolution [23, 33]. A similar trade-off is known for changes in the transverse intensity distribution in the focal plane [32]. On the basis of conservation of power, the intensity of side lobes in the transverse direction of the focal plane must also increase in those situations for which the axial Strehl ratio is observed to decrease. This implies then that the axial and transverse PSF are coupled, and the axial and transverse spot size cannot be separately engineered. Thus, in designing an DOE the complete 3D PSF must be considered within the context of a given application and with regard for how the photo-activated process depends upon the 3D intensity distribution about the focal region.

To augment this point, let us consider the application of PSF engineering to a photolithographic technique known as three-dimensional microfabrication (3DM). In 3DM, complex microstructures can be fabricated by patterned scanning of a tightly focused pulsed laser beam within the volume of a multi-photon-excitable medium [59]. The structure resolution is determined by the axial and transverse size of the photo-processed volume element or voxel generated at the focus. For most photo-induced

processes, the material response to the local intensity is nonlinear (whether one- or multi-photon induced), so the voxel is defined by those points in the PSF for which the local intensity exceeds the photo-response threshold of the medium [59-61]. The average focused power can be adjusted so that a selected isophote matches the threshold intensity. It is practical then to consider the situation that results when the threshold matches the 50% isophote, for which the resulting voxel shape would match the 50% isophote surface. In this case, the axial (transverse) size of the voxel as a function of DOE configuration would be given by the axial (transverse) superresolution factor G, if only the central lobe has appreciable intensity. Note, however, that if the side lobe intensity exceeds 50%, then the photo-processed volume will be comprised by multiple features.

Previous studies of superresolution have defined the spot-size in terms of the separation between local minima adjacent to the central lobe. Such a definition is appropriate for imaging applications, but poorly suited to lithographic processes (like 3DM) because the minima may or may not correspond to points at which the intensity is zero and they do not alone indicate the shape of the PSF with respect to the photo-response threshold. Clearly, side lobe intensity and the Strehl ratio are relevant when considering the superresolution that can be achieved in a given application.

The superresolving performance of a three-zone DOE can be evaluated using criteria that account more thoroughly for overall changes in the 3D-PSF. Figure 2.5 is a plot of axial G for all points in the $[r_1, r_2]$ space for which $G < 1$ and $M < 0.5$. This represents the sub-set of three-zone DOEs that yield an axially superresolved focus and for which the intensity of the side lobes remains below 50% of the peak intensity. Under these criteria, the maximum axial superresolution occurs for $r_1 = 0.58$, $r_2 = 0.73$, $\Phi_1 = 0$, $\Phi_2 =$
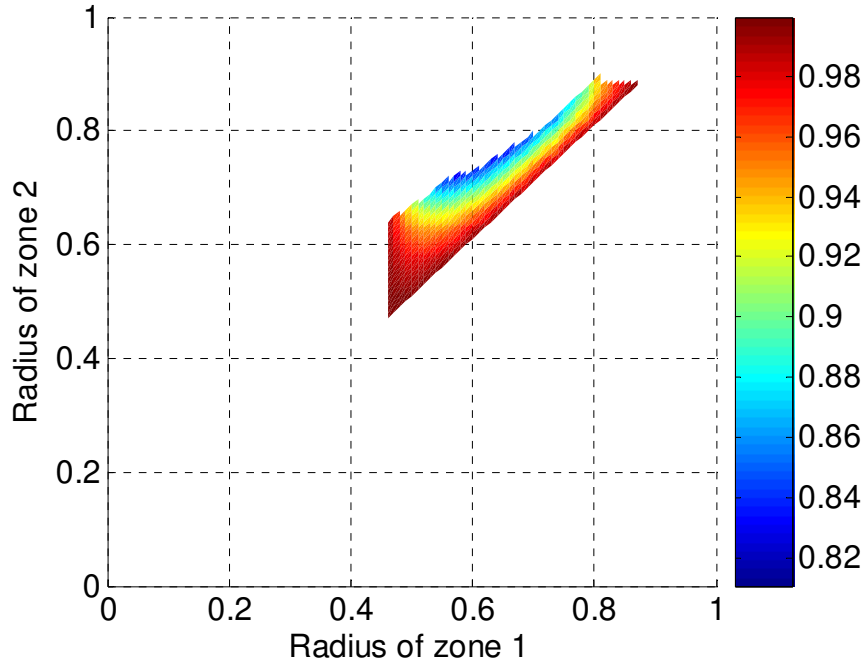
Figure 2.5. Sub-space of G versus $[r_1, r_2]$ for a three-zone DOE having $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$ for which axial superresolution is achieved ($G < 1$) and side lobe intensity remains below 50% of the peak value ($M < 0.5$).

π, and $\Phi_3 = 0$, at which point $G = 0.81$, $M = 0.47$, and $S = 0.38$. The effect of this DOE on the 3D-PSF is shown in Fig. 2.6. The intensity distribution along the optic axis shows clearly that superresolution is achieved at the expense of higher side lobes. The transverse intensity distribution also shows that some power is re-distributed into weak side lobes, which is consistent with the decrease in the Strehl ratio. This DOE could be used for 3DM, 3D optical data storage/read-out, or any other application that requires a co-minimized axial and transverse intensity distribution and minimized side lobes.

In certain regions of the PSF solutions space, the DOEs elongate the PSF so that the intensity along the optic axis remains high over a greater distance from the focal plane ($G_{axial} > 1$). This can also be viewed as an extended depth of focus. Similar findings
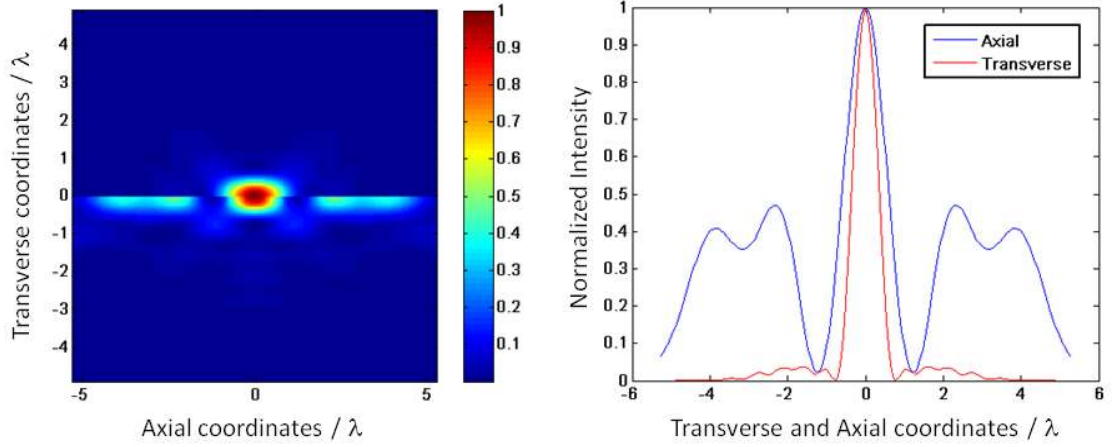
Figure 2.6. Comparison of the focused PSF generated when a three-zone DOE having $\Phi_1 = 0$, $\Phi_2 = \pi$, $\Phi_3 = 0$, $r_1 = 0.58$, and $r_2 = 0.73$ is placed before the lens. (**Left-top**) Normalized axial and transverse intensity distribution within the plane of polarization (*xz*-plane) in the diffraction-limit (no DOE) and (**Left-bottom**) when the three-zone DOE is present. (**Right**) Axial and transverse intensity distribution of the DOE-modified beam alone. The DOE-generated PSF is axially super-resolved by $G_{axial} = 0.81$, with $M = 0.47$ and $S = 0.38$, whereas the transverse intensity distribution of the central lobe is minimally broadened by $G_{trans} = 1.01$.

have been reported for annular phase DOEs [12, 37]. In the case of the two-zone DOE, G_{axial} takes a maximum value of 2.31 at $\phi_2 = \pi$ and $r_1 = 0.54$. Under these conditions the lateral extent of the central lobe is $G_{trans} = 1.00$, as measured in the transverse plane that contains the peak axial intensity. Even greater PSF elongation can be obtained with a three-zone DOE, which produces a maximum value of $G_{axial} = 6.1$ for $r_1 = 0.43$, $r_2 = 0.69$, $\phi_1 = 0$, $\phi_2 = \pi$, and $\phi_3 = 0$. The axial and transverse intensity distribution in the plane of polarization (*xz*-plane) is shown in Fig. 2.7. The elongated PSF appears to be the result of close overlap between a focal-plane centered lobe and four adjacent axial side lobes. The outer side lobes of the set attain the same peak intensity as the focal-plane centered lobe, and the intensity between lobes decreases to only ~60% of the peak value. It is

19

noteworthy that the transverse width of the central lobe remains nearly invariant along the full length of the five-lobe set ($G_{trans} = 0.99$ in the Gaussian focal plane). This intensity profile could be used for laser drilling and laser machining applications in which sub-diffraction-limited features are created over an axial distance of several microns. It
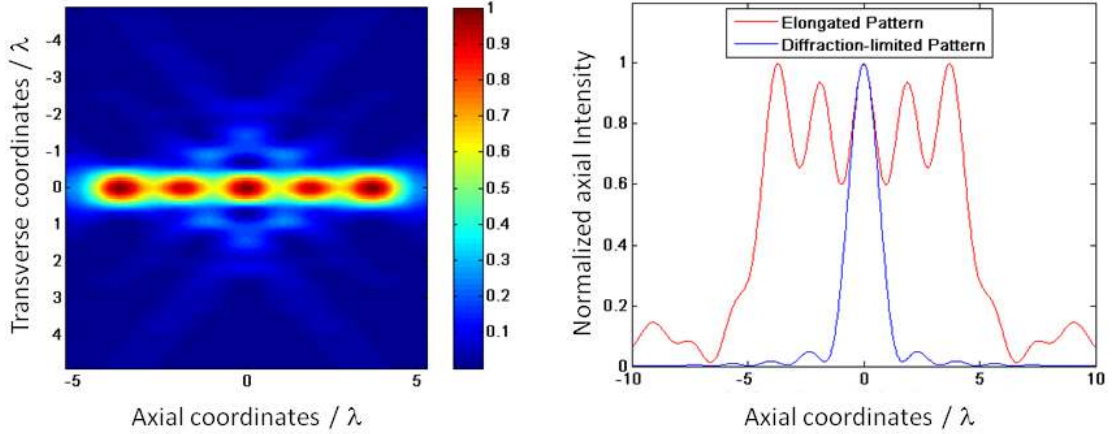
Figure 2.7. (**Left**) Normalized axial and transverse intensity distribution in the plane of polarization (*xz*-plane) resulting when a three-zone DOE having $r_1 = 0.43$, $r_2 = 0.69$, $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$ is placed before the lens. The PSF is axially elongated by a factor of $G_{axial} = 6.1$ yet remains diffraction limited in the transverse direction ($G_{trans} = 0.99$). (**Right**) DOE-modified axial intensity distribution (red trace) versus that computed for diffraction limited focusing (no DOE, blue trace).

could also be used in microscopy and imaging applications for achieving sub-diffraction-limited resolution over an extended depth of field.

The effect of annular DOEs on G_{trans} was considered in some earlier studies of axial superresolution [24, 26, 35]. There, it is shown that Eq. (2.6) expressed in the scalar approximation can be re-written as a one-dimensional Fourier transform of the pupil function *t* through the change of variables $\zeta = [(\cos\theta - \cos\alpha)/(1 - \cos\alpha)] - 0.5$. In these

works, centrosymmetric DOEs – those for which $t(\zeta)$ is an even function – are shown to leave $G_{\text{trans}} = 1$. However, the subject does not appear to have been explored to a level that one may conclude $G_{\text{trans}} = 1$ *if and only if* the DOE is centrosymmetric. We note that the DOEs discussed in the present work are not exclusively centrosymmetric (*e.g.* the DOEs corresponding to Figs. 2.6 and 2.7 are non-centrosymmetric). These results suggest that minimal change to the transverse PSF can also be achieved with certain non-centrosymmetric DOE configurations.
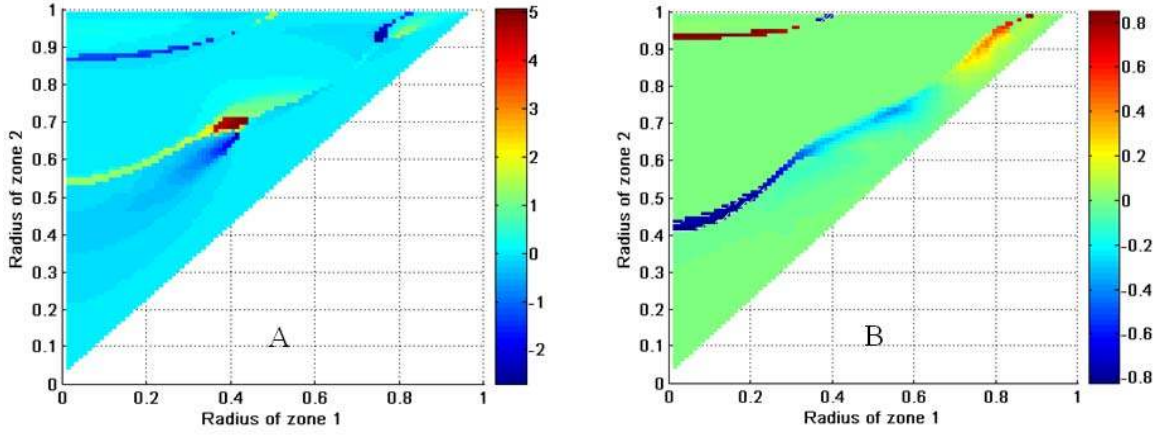
Figure 2.8. Comparison of the axial PSF parameters G and M as calculated using vector diffraction and scalar theory for three-zone DOEs having $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$. (**A**) $G_{\text{vector}} - G_{\text{scalar}}$ and (**B**) $M_{\text{vector}} - M_{\text{scalar}}$ versus $[r_1, r_2]$.

Although it is commonly agreed that high-NA focal field distributions are not accurately described by scalar theory or methods that employ the paraxial approximation, the magnitude of the discrepancy has not been widely examined. This subject was investigated quantitatively by using both vector diffraction and scalar theory [4] to

compute axial PSFs generated with three-zone DOEs having $\Phi_1 = \Phi_3 = 0$ and $\Phi_2 = \pi$ and

then plotting the differences in the characteristic parameters G_{vector} - G_{scalar} and M_{vector} -

M_{scalar} versus $[r_1,\ r_2]$ (Fig. 2.8). It was found that both levels of theory predict
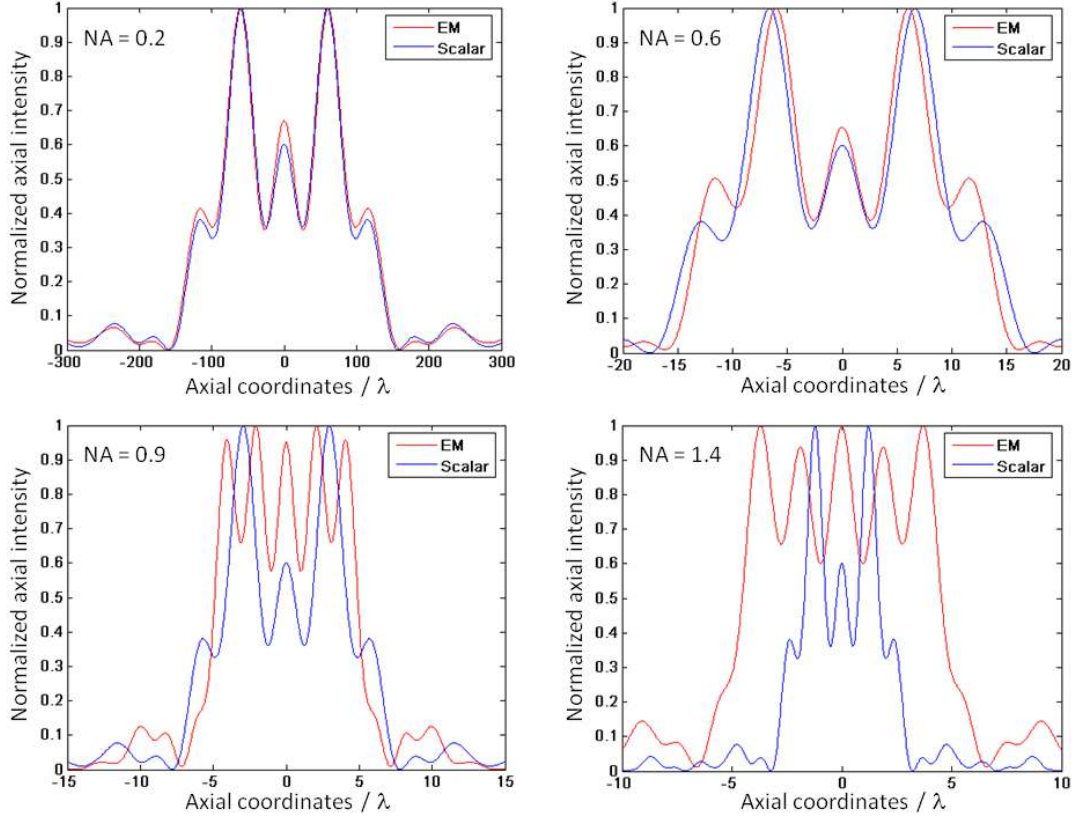
Figure 2.9. Comparison of the normalized axial intensity distribution in the plane of polarization (*xz*-plane) calculated using vector diffraction (EM) and scalar theory at four values of *NA* for the case in which a three-zone DOE having $r_1 = 0.43$, $r_2 = 0.69$, $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$ is placed before the lens.

quantitatively similar changes in the PSF as a function of DOE configuration and the

values of the characteristic PSF parameters *G* and *M* are similar. Yet they differ most in

those situations for which the PSF undergoes extreme axial change, be that

superresolution or elongation (compare Figs. 2.8 and 2.4). To illustrate the point further,

22

Fig. 2.9 shows the evolution of the axial intensity distributions calculated using the vector diffraction and scalar methods at four values of NA for a three-zone DOE having $r_1 = 0.43$, $r_2 = 0.69$, $\Phi_1 = 0$, $\Phi_2 = \pi$, and $\Phi_3 = 0$, which produces the axially stretched PSF shown in Fig. 2.7 (at $NA = 1.4$). Although both levels of theory predict that the PSF is axially elongated, the patterns differ significantly as the NA increases. Notably, we find that $G_{\text{vector}} - G_{\text{scalar}} = 5.04$ at $NA = 1.4$. Thus, scalar theory may be useful for rapid, qualitative assessment of DOEs under high-NA conditions, but vector diffraction theory appears essential for accurate simulation of the PSF.

2.4 Conclusion

Vector diffraction theory was used to examine the effect of two- and three-zone DOEs on the 3D-PSF generated under high-NA focusing of linearly polarized incident light. A systematic approach was adopted in which PSFs were calculated and compared for all possible combinations of phase and zone radius within the discretized two- and four-dimensional space associated with two- and three-zone DOEs, respectively. Two-zone DOE configurations were identified that marginally decrease the axial width of the central lobe, but this is accompanied by a large increase in the intensity of adjacent side lobes that make the achievable intensity distributions unsatisfactory for most applications. Conditions were found for which a three-zone DOE yields an axial intensity distribution that is superresolved by 19% with minimal change in the transverse profile and sufficiently small side lobes that the intensity pattern could be used for micro-lithographic and micro-imaging applications. Interestingly, conditions were also identified for which the axial PSF is elongated by 510% with only 1% change along the transverse direction. This intensity distribution could be used for sub-micron-scale laser

23

drilling and machining. A comparison of intensity distributions calculated under high-*NA*

conditions using vector and scalar theories shows that the latter is suitable for identifying

qualitative changes in the PSF, but the detailed intensity distribution can differ markedly

from that computed using the more accurate vector diffraction method.

CHAPTER 3: DESIGN OF AXIALLY SUPERRESOLVING PHASE FILTERS USING THE METHOD OF GENERALIZED PROJECTIONS

3.1 Introduction

In chapter 2, we reported the use of exhaustive search to investigate superresolution using two- and three-zone phase-only annular DOEs, and we showed that the vector character of the field cannot be neglected under high-*NA* focusing [43]. This study was limited to simple DOEs because applying exhaustive search to pupil filters having a larger number of zones is computationally impractical.

To explore more general DOE profiles under high-*NA* conditions, we developed an algorithm for designing phase-only pupil filters that control the axial PSF based on the Method of Generalized Projections (MGP) [62]. The MGP belongs to the family of iterative algorithms in which the field is propagated repetitively forward and backward between the pupil and focal domain as constraints are applied in both regions. This process is continued until the algorithm converges to a satisfactory solution or a fixed number of iterations is completed. The MGP is particularly well suited for designing pupil filters under high-*NA* focusing due to the following features. The algorithm is independent of the propagation operator, so it does not require any kind of approximation to the vector diffraction integral and high-*NA* focusing can be treated rigorously. Further, the MGP can accommodate nonlinear, non-convex and inconsistent/non-physical constraints, which is one of the major problems in synthesizing 3D and/or axial field distributions [53]. Most importantly, it enables constraints to be defined that superresolve the axial intensity while keeping the sidelobes below a threshold limit.

Although axial superresolution is advantageous to many applications, it is always accompanied by the detrimental effect of higher side-lobe energy. In certain applications, such as DLW [59], the presence of high intensity side-lobes would render the improvement in axial resolution useless. Therefore, approaches to DOE design should enable constraints to be defined that superresolve the axial intensity while keeping the side-lobes below a threshold limit. The concern with such a problem, as with any field synthesis problem, is that it is not known *a priori* whether such an axial field distribution satisfies the wave equation. In this context, the problem involves finding a solution that satisfies the constraints as closely as possible while still conforming to the physical principles of diffraction. Previous studies suggest that the set theoretic approaches and vector-projection type algorithms like MGP are most efficient for these types of problems [53].

In our MGP-based approach, the electromagnetic diffraction integral derived by Richards and Wolf [5] is used to calculate the axial field generated by a given DOE. This process is referred to hereafter as "propagating [the field] forward". The converse, referred to as "propagating backward", is considerably more challenging, in that an arbitrary field distribution one may devise is not necessarily a solution to the wave equation. Kant reported a formulation for calculating the annular complex (phase and amplitude) DOE that would most closely generate a given axial PSF. We apply this formulation for backward propagation, using a phase-only constraint applied in the object field to retain a phase-only DOE function for subsequent iterations.

As with many optimization algorithms, the MGP is sensitive to the starting conditions, and different solutions can result depending upon the initial configuration of

the pupil filter. Such behavior results due to the presence of ''traps'' and ''tunnels'' [62], which hinder the MGP from finding a global solution. Traps and tunnels are common in PSF engineering problems because the constraints are commonly non-convex and inconsistent. To study the solutions space more rigorously and in turn avoid traps and tunnels, routines are incorporated that systematically generate a large number of starting pupil filters (3442 in all). The solutions obtained were analyzed manually and the best were selected based on a compromise between superresolving the central-lobe and increasing intensity of the axial sidelobes.

In the present chapter we provide a detailed account of our MGP-based algorithm and demonstrate how the method may be used to design phase-only DOEs that reshape the axial field under high-*NA* focusing. We show that by properly choosing constraints, one can design a phase-only DOE that superresolves the axial PSF while maintaining side-lobes below a specified limit. Because binary phase DOEs are generally more effective for axial superresolution (see Sect. 2.3) and are more easily fabricated than analog phase DOEs, we have also incorporated an iterative binarization algorithm [63] that transforms the resulting analog DOE profile into a $0/\pi$ binary phase profile.

3.2 Inverse problem of vector diffraction

Consider a rotationally symmetric optical system with a DOE placed at the entrance pupil of an aberration-free high-*NA* objective lens, as shown in Fig. 3.1. The DOE introduces aberration that modulates the PSF in the region about the Gaussian focus. The incident plane wave is assumed to be linearly polarized with spatially uniform amplitude. According to the vector diffraction theory of Richards and Wolf [5], the electric field distribution along the optical axis is given by Eq. (2.6).
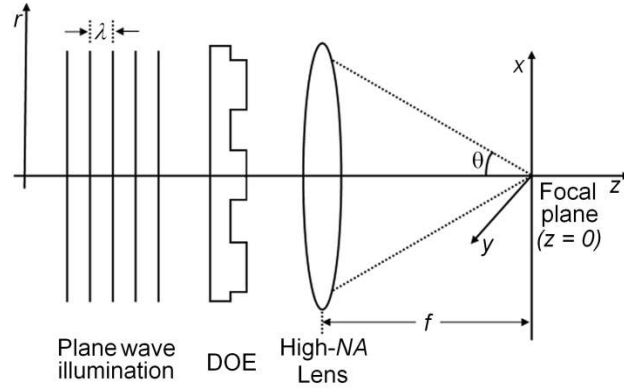
Figure 3.1. Optical configuration for modifying the focused point-spread function (PSF) using a DOE.

By setting the radial coordinate $v = 0$ in Eq. (2.6), the axial field distribution is given by

$$E_{axial}(z) = iA \int_{q(0)}^{q(\alpha)} T(q) \sqrt{q}(1+q) \exp(ikzq) \, dq \qquad (3.1)$$

where $q = \cos\theta$. The maximum aperture angle, $\alpha = \arcsin(NA/n)$, is determined by the numerical aperture of the lens and the medium refractive index, n. The wave number $k = 2\pi/\lambda$. The variable q is related to the normalized radius of the aperture, r, by $r = n(1 - q^2)^{1/2}/NA$. The axial intensity distribution can be calculated as $I(z) = |E_{axial}(z)|^2$. The DOE spatially modifies the wave front according to the complex transfer function $T(q)$. The modified wave front can then be propagated forward with the help of Eq. (3.1) giving the resulting axial intensity distribution, E_{axial}.

To propagate backward it is necessary to solve the inverse of the vector diffraction problem represented by Eq. (3.1); we want to be able to calculate the DOE transfer function $T(q)$ that would generate a *given* axial field distribution $E_{axial}(z)$. This type of problem is similar to solving the Fredholm integral of the first kind [64] and can

28

be accomplished by applying the following formulation developed by Kant [65]. The kernel of Eq. (3.1) is expanded as a summation of Gegenbauer polynomials, $C_s^{(1/2)}$, (Appendix B.1) as:

$$T(q)\sqrt{q}(1+q) = \sum_{s=0}^{\infty} a_s C_S^{1/2}(q). \tag{3.2}$$

Substituting the right hand side of Eq. (3.2) into Eq. (3.1) enables the whole integral to be evaluated as a finite summation of spherical Bessel functions of the first kind (Appendix B.2):

$$E_{axial}(z) = \sum_{s=0}^{\infty} a_s i^S j_S(kz), \tag{3.3}$$

where j_s denotes spherical Bessel functions of the first kind. The infinite sum may be replaced by a finite sum of N terms whose number depends upon the required accuracy. Expression (3.3) provides a set of N algebraic equations which can be solved in terms of E_{axial} to obtain N complex coefficients a_s. The transfer function that generates E_{axial} can then be computed by substituting a_s into Eq. (3.2) and solving for $T(q)$ as:

$$T(q) = \frac{\sum_{s=0}^{\infty} a_s C_S^{1/2}(q)}{\sqrt{q}(1+q)}. \tag{3.4}$$

It should be noted that an arbitrary axial PSF is not necessarily a solution to the wave equation, so coefficients a_s obtained by solving Eq. (3.3) may only define a field distribution that most closely matches E_{axial}. Collectively, Eqs. (3.1) - (3.4) provide a means that may be implemented in the MGP for rigorously propagating forward and

backward between the DOE transfer function and the axial field distribution under non-paraxial conditions.

3.3 Method of Generalized Projections

Given a set of constraints C_γ ($\gamma = 1,2,\ldots,\eta$), the MGP seeks a solution function S by iteratively projecting onto constraints according to $S_{\xi+1} = P_1 P_2 \ldots P_\eta S_\xi$ [62], where ξ is an iteration index, S_0 is an arbitrary initial function, and P_γ is a projection operator that maps S onto its nearest neighbor in C_γ. When the constraints are inconsistent (meaning all cannot be satisfied simultaneously), the MGP yields a solution that most nearly satisfies the constraints, as quantified by the summed-distance error, SDE [62]. SDE gives the sum of distances of S_ξ from the constraint set C_γ. The set-distance reduction property states that if one or more constraints is non-convex, the SDE is only guaranteed not to increase when η is limited to two. Yet if $\eta = 2$ and the SDE converges to a non-zero minimum, the solution obtained may yet correspond to a local minimum, or a "trap", and thus may not be the optimum solution globally. A trap may be thought of as a solution that lies at equal distance from all constraints while satisfying none. Traps exist only when nonconvex constraints are involved, as is most often the case in arbitrary PSF engineering.

3.4 Performance parameters

Performance-based metrics, such as the normalized mean square error (NMSE), are closely related to SDE but better suited to the present problem, given that we are primarily interested in achieving a prescribed axial intensity profile. The NMSE can be defined as [66]:

$$\text{NMSE} = \frac{\sum\limits_{z}\left(\rho I_\xi - I_{\text{target}}\right)^2}{\sum\limits_{z} I_{\text{target}}^2} \qquad \rho = \frac{\sum\limits_{z} I_{\text{target}} I_\xi}{\sum\limits_{z} I_\xi^2}, \qquad (3.5)$$

where I_ξ is the axial PSF generated in iteration ξ, and I_{target} is the desired axial PSF. The factor ρ is chosen to scale I_ξ to the same average value as I_{target}.

The performance of the DOE was characterized using G and M as defined in the previous chapter.

3.5 Starting conditions and design constraints

A key to implementing the MGP algorithm successfully is applying appropriate starting conditions and constraints. As demonstrated by Wyrowski [63], proper engineering of the applied constraints can result in a tremendous improvement in the performance of the achieved result. Starting conditions refers specifically to the DOE transfer function that is applied at the first iteration. Our calculations are based on a vacuum wavelength of $\lambda_v = 800$ nm and a limiting aperture diameter of 9.3 mm set by the entrance pupil of a 1.4 NA lens, which focuses through a medium having $n = 1.516$. In keeping with the Kirchoff boundary condition, the minimum width of a DOE zone was limited to $100\lambda_v$, which corresponds to a maximum of 58 phase-zones across the radius of the limiting aperture. Given that we are ultimately interested in obtaining a binary phase DOE, there are $2^{57} \approx 1.44 \times 10^{17}$ unique possible DOEs. This corresponds to such a vast phase space that any solution obtained by the MGP is likely to correspond to a local minimum in the NMSE (a trap) and not necessarily the optimum solution globally. Additionally, any solution obtained will depend upon the starting conditions, particularly

given that any set of arbitrary constraints is likely nonconvex. In addition to traps, MGP algorithms can also generate apparent solutions called "tunnels" that are in fact paths to solutions for which the gradient of the NMSE is so small that the algorithm appears to have already converged [62]. Global optimization algorithms, such as simulated annealing [67], avoid tunnels and traps by incorporating random variation into the algorithm. We have found that MGP results can vary considerably with different starting conditions and constraints. In what follows we discuss how the constraints were chosen and how a set of variable yet deterministic starting DOEs were constructed to enable a more thorough search of the multi-dimensional solutions space.

3.5.1 Design constraints

It is essential to build degrees of freedom into the constraints that permit the algorithm to converge to a solution that is physical yet most closely approximates the idealized PSF [68]. Degrees of freedom may also be incorporated into the axial field distribution by not restricting values of amplitude, phase, and/or absolute scale. Allowing the field amplitude to vary arbitrarily outside the region of interest is commonly referred to as amplitude freedom. Phase freedom is even less restrictive in that it involves allowing the phase to achieve any value at all points along the axial field, not just outside the region of interest. Phase freedom is particularly useful when the goal is to achieve a specific intensity distribution, without concern for the phase itself. Finally, scaling freedom may be applied when the overall shape of the field pattern is of greater importance than the absolute intensity at any point. When scaling freedom is included, a scaling factor is incorporated into the constraints to specify the *relative* intensity of the central-lobe relative to side-lobes.

For the present problem, axial superresolution is achieved by defining the following constraints. The algorithm is constrained to seek a phase-only DOE by applying projection operator P_1, defined as $P_1[T_\xi(q)] = \exp[j\,\Phi_\xi(q)]$, where $\Phi_\xi(q)$ is the phase component of the DOE complex profile in the ξ-th iteration. With respect to the axial field distribution, the current problem's requirements imply two constraints: (i) sharpening the primary lobe and (ii) maintaining the side-lobe amplitudes below the specified limit A_{SL}. To take advantage of the set-distance reduction property, requirements (i) and (ii) are combined into a single constraint P_2 that projects a target amplitude distribution solely onto the field-domain as

$$P_2\big[E_\xi(z)\big] = \begin{cases} \frac{1}{2}\big[1 + \cos(z/a)\big]\exp\big[i\phi_\xi(z)\big] & |z| \le Z_{PL} & \text{Primary-lobe region,} \\ 0 & Z_{PL} < |z| \le (Z_{PL} + Z_{ZR}) & \text{Zero-region,} \\ \min\big[|\,E_\xi(z)|, A_{SL}\,\big]\exp\big[i\phi_\xi(z)\big] & |z| > (Z_{PL} + Z_{ZR}) & \text{Side-lobe region.} \end{cases} \qquad (3.6)$$

The primary lobe amplitude is shaped to one cycle of a DC-offset cosine whose frequency and extent are set by a and Z_{PL}, respectively. To encourage the primary lobe to narrow, the field amplitude is forced to zero in the "zero-region". In the "side-lobe region", $|E_\xi(z)|$ is allowed to assume any amplitude less than or equal to A_{SL}. This provides a degree of amplitude freedom. The axial phase distribution, $\phi_\xi(z)$, is retained unchanged in each iteration. This measure introduces phase freedom into the procedure. Scaling freedom is employed by normalizing the axial field prior to applying P_2. In our implementation the following values were used: $a = 0.2$, $Z_{PL} = 0.5\lambda$, $Z_{ZR} = 1.5\lambda$ and $A_{SL} = 0.4 \times A_{peak}$, where A_{peak} is the maximum of the central-lobe.

3.5.2 Initial conditions

To explore the solutions space more thoroughly, and thereby avoid traps and tunnels, the MGP algorithm was initiated multiple times using a systematic set of starting binary DOEs. The best resulting PSFs and corresponding starting DOEs were selected based on comparison of G, M, and NMSE achieved from among all starting conditions. Starting DOEs were constructed from a sequence of phase-only $0/\pi$ zones characterized by a small set of parameters, principal among which are the zone-frequency, F, and duty cycle, DC. The period $D = 1/F$ defines the width of two consecutive DOE zones, and DC defines the width of the inner zone relative to D. Figure 3.2 illustrates how F and DC are related to a DOE phase profile for which F is held constant. The maximum permissible frequency was set by the Kirchhoff boundary condition. The minimum frequency was limited by the radius of the entrance pupil, so a starting DOE consists of only two zones when $1/F$ equals the radius of the entrance pupil. Starting DOEs were constructed using several different variations of F along the radial direction. These included (i) constant F, (ii) geometrically decreasing F, (iii) Gaussian F, and (iv) starting DOEs having N_Z equal-area zones. This approach is somewhat like simulated annealing, in that many different starting conditions within the solutions space are considered, but it has the merit of being fully deterministic and repeatable.

Here we describe construction of a starting DOE for which F decreases geometrically along r. The resulting DOE and parameters are illustrated in Fig. 3.3(A). The minimum period of the first zone-pair, D_1, was computed for a given DC based on the Kirchoff constraint:

$$D_1 = \frac{100\lambda_v}{DC}.$$ (3.7)
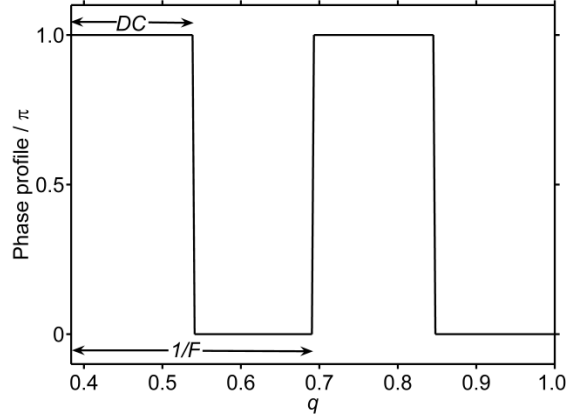
Figure 3.2. Multi-zone binary DOE represented by two parameters only: a constant zone-frequency F and duty cycle DC.

This sets the maximum starting frequency $F_1 = 1/D_1$. The period for subsequent zone-pairs was calculated using a fixed period-multiplier, PM:

$$D_{i+1} = PM \times D_i = PM^i \times D_0 \tag{3.8}$$

This causes the zone widths to steadily increase across the DOE width. The process was ended when the i-th period exceeded the remaining width of the DOE, and the final zone was extended to the outer edge of the DOE. PM was varied between 1 and 3 in steps of 0.05. For every value of PM, the DC value was varied between 10% and 90% with a step of 10%. Geometric starting DOEs were also constructed by shifting the minimum-width zone from the DOE periphery to a more central position q_c, as illustrated in Fig. 3.3(B). The frequency of flanking zones were decreased to the left and right of q_c by geometric factors PM and $m \times PM$, respectively. In the case of $m = 1$, the zone frequency decreased symmetrically on either side of q_c. The highest frequency zone was positioned at $q_c =$

35

0.55, 0.63, and 0.76, and for each q_c starting DOEs were generated for all combinations of $PM = 1$ to 3 with a step of 0.1; $m = 1$ to 3.5 with a step of 0.5, and DC = 10% to 90% with a step of 10%.
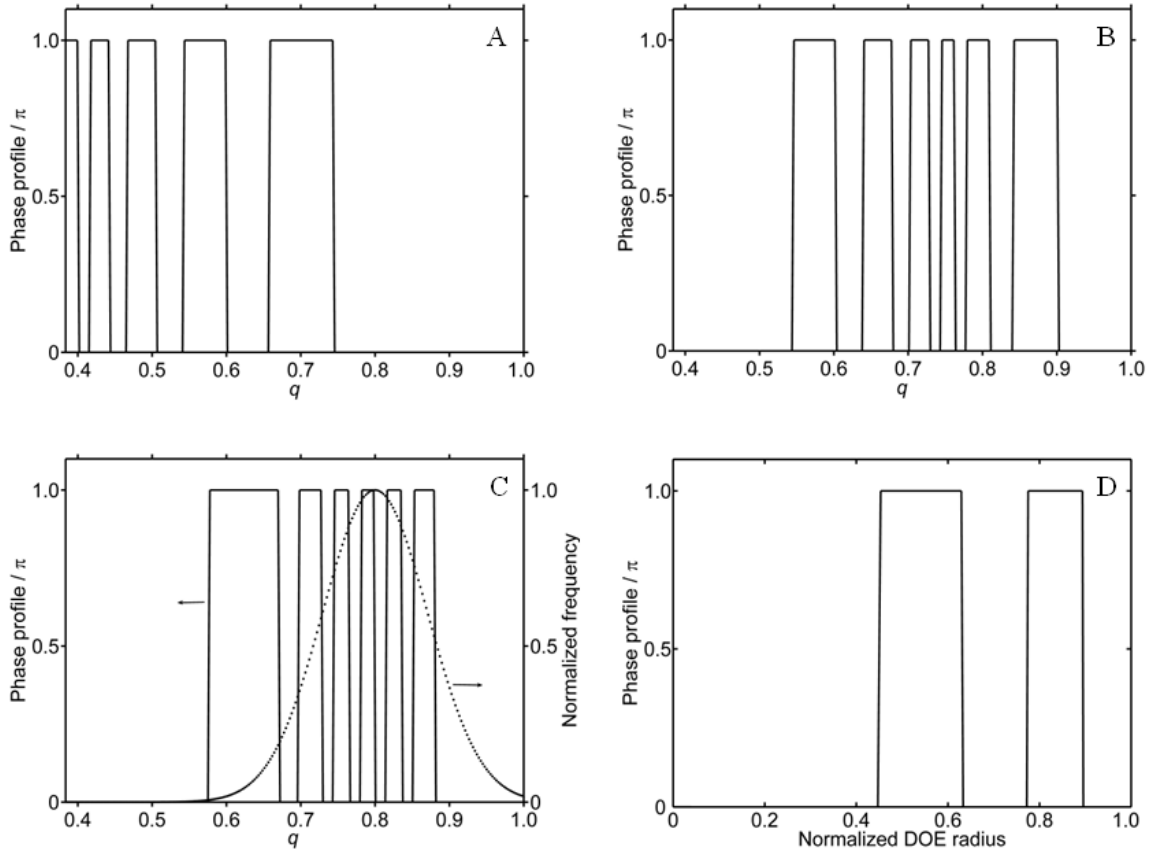
Figure 3.3. Construction of systematic starting DOEs used to initiate MGP calculations. Individual zone widths are specified in terms of F and DC, where F (shown as the dotted line) varies with DOE radial coordinate q according to (A) geometric-, (B) asymmetric geometric-, and (C) Gaussian functions. (D) MGP simulations were also started using DOEs having Nz equal-area zones where $Nz = 1$ to 29 (five-zone DOE shown).

Gaussian starting DOEs like that in Fig. 3.3(C) were constructed using zone-frequencies given as an analytic function of q by

$$F(q) = \exp\left[-\left(\frac{q - q_{\text{peak}}}{\Delta_{1/e}}\right)^2\right]$$

(3.9)

Starting DOEs were generated using all combinations of q_{peak} = 0.5 to 0.9 with a step of 0.05; $\Delta_{1/e}$ = 0.02 to 0.2 with a step of 0.02; and DC = 25%, 50%, and 75%. Constant-frequency DOEs were created with F varied from $1/D_1$ as defined in Eq. (3.7) down to the frequency for which only a single period fit within the DOE radius. For each value of F, DC was varied from 10% to 90% with a step of 10%. Equal-area DOEs were constructed by calculating the boundaries along r that yielded N_z zones having equal area in real space. The maximum value of N_z was set to 29 by the Kirchoff boundary condition. An example of an equal-area DOE is shown in Fig. 3.3(D).

3.6 DOE binarization

To obtain binary DOEs, a binarization algorithm based on the "soft-quantization" concept [63] was incorporated into the MGP algorithm. Binarization is implemented after the NMSE minimizes by continuing MGP iterations with a modified P_1 that progressively forces $\Phi_\xi(q)$ toward the closer of values 0 or π over all q. The DOE phase space is folded into the domain $[0, 2\pi]$ by taking the modulus of Φ_ξ with respect to 2π. Within $[0, 2\pi]$ two sub-domains of width 2Δ are defined as $[\pi/2-\Delta, \pi/2+\Delta]$ and $[3\pi/2-\Delta, 3\pi/2+\Delta]$. With each iteration the sub-domains are widened by increasing Δ, and for all q lying within a sub-domain $\Phi_\xi(q)$ is set to the nearest bounding value of the sub-domain. When Δ reaches $\pi/2$, Φ_ξ equals 0 or π for all q. The binarized profile can then be transformed to a function of r. Wyroski has shown that forcing an immediate binary

37

solution causes the MGP algorithm to stagnate because the pupil phase function changes too abruptly between iterations [63]. The soft-quantization process implemented here within the MGP prevents stagnation yet forces the algorithm to converge to a binary DOEs.

3.7 Algorithm flow

The MGP algorithm iterates in the following order. Beginning with a starting DOE, the complex axial field $E_\xi(z)$ is calculated over the range $-25\lambda \leq z \leq +25\lambda$ using Eq. (3.1). The axial intensity constraints of P_2 are then applied and NMSE is calculated using Eq. (3.5). The resulting axial field is back-propagated to the DOE plane using Eqs. (3.3) and (3.4) to obtain a revised complex transfer function $T(q)$. Constraint P_1 is then applied to $T(q)$, and the resulting phase-only DOE provides a starting point for the subsequent iteration. The process is repeated for a set number of iterations or until the NMSE minimizes.

3.8 Results and discussion

Figure 3.4 illustrates convergence of the MGP algorithm to a solution given a single starting DOE. In this example the starting phase element has a constant frequency binary $0/\pi$ phase-only profile with $DC = 20\%$. Figure 3.4(B) shows the calculated axial PSF generated when the starting phase element is positioned at the entrance pupil of the optical system. Compared with the diffraction-limited intensity profile, the modified intensity distribution does not exhibit any superresolution in the main-lobe ($G = 1$) but does have strong side-lobes. As the algorithm iterates, the phase profile is modified and the NMSE steadily decreases (Fig. 3.4(D)). This indicates that the axial PSF generated

38

by the evolving DOE increasingly satisfies constraints P_1 and P_2 and is thus reshaped towards the targeted profile. The NMSE value also plateaus between iterations 6 and 21, which corresponds to a tunnel, during which the algorithm converges only slowly toward the solution. The simulation was regarded as converged when the NMSE decreased by less than 0.0001 over 25 consecutive iterations. The resulting analog-phase DOE generated an axial PSF characterized by $G = 0.75$ and $M = 0.59$. Following convergence, the DOE was binarized, giving the 17-zone $0/\pi$ phase-only DOE shown in Fig. 3.4(C). This solution DOE generates the superresolved axial PSF of Fig. 3.4(E) having $G = 0.69$ and $M = 0.61$. In this instance, binarization further improved superresolution with minimal increase in side-lobe intensity relative to the analog-phase solution DOE, although this was not observed with every starting DOE.

Initiating the MGP algorithm using the systematic set of starting DOEs yielded a wide range of solutions with varying degrees of axial superresolution. Taken as a whole, all solutions suggest that strong axial superresolution and weak axial side-lobes are mutually exclusive characteristics. Solutions offering G as low as 0.66 were accompanied by large side-lobes having $M \cong 1$. Solutions having smaller side-lobes were also obtained ($M < 0.1$), but these offered much less axial superresolution ($1 > G > 0.9$).

Solutions offering high superresolution with $G < 0.66$ were also be found, but these produced axial side lobes that were much more intense that the central lobe located at (or near) the geometric focus. Such solutions could in fact be useful for certain applications, such as confocal and multiphoton imaging, for which optical signal originating from all points outside of the superresolved geometric focus can be suppressed by a confocal aperture located in front of the detector.
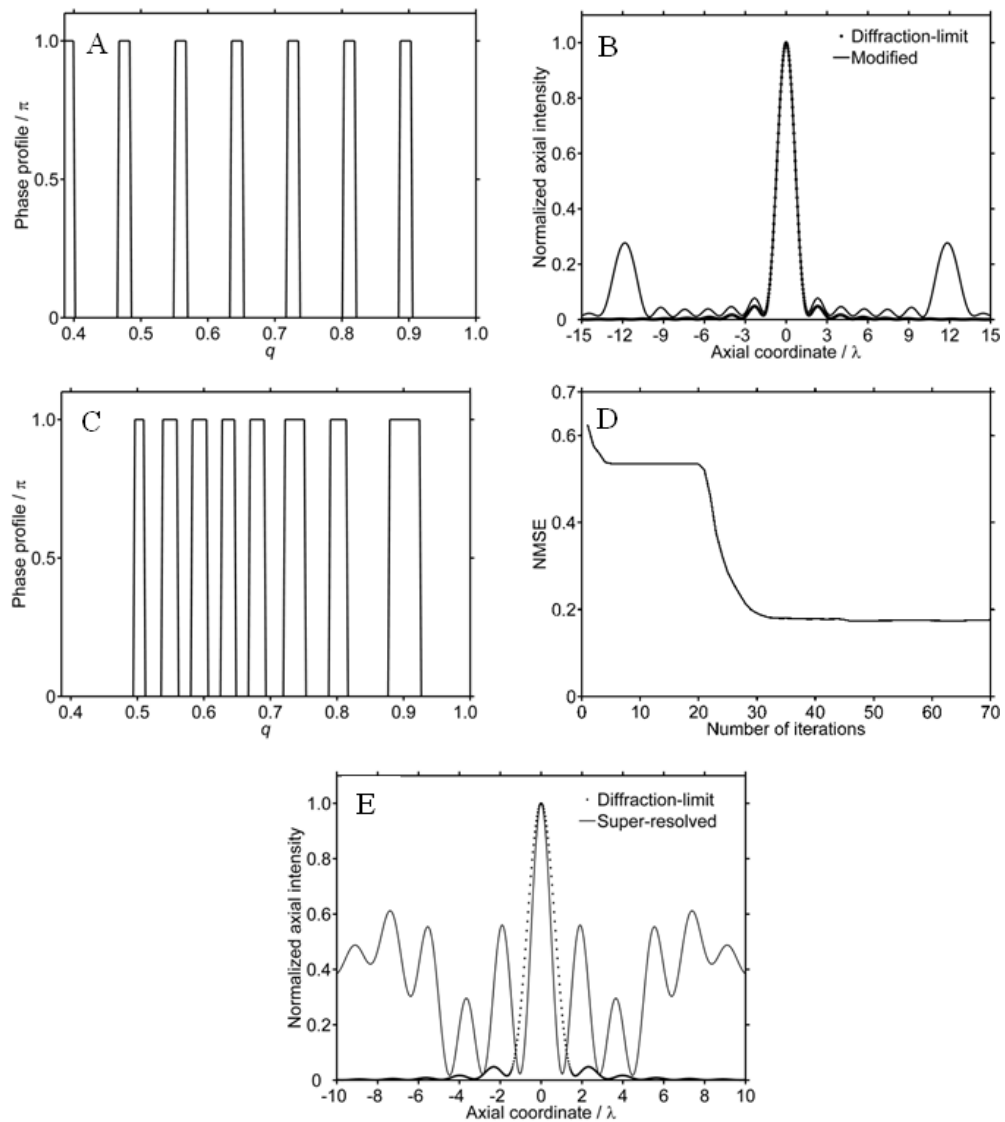
Figure 3.4. Example of the vector diffraction MGP algorithm converging to a solution from a starting DOE. (A) Phase profile of the starting DOE versus coordinate q (0/π phase only, constant zone-frequency, DC = 20%). (B) Axial PSF resulting when the starting DOE in (A) is positioned at the entrance pupil of a 1.4-*NA* objective lens (*G* = 1, *M* = 0.24). (C) Phase profile versus q of the binarized super-resolving DOE to which the algorithm converged. (D) NMSE versus iteration when the algorithm is started using the DOE in (A). (E) Axial PSF generated when the solution DOE in (C) is placed before a 1.4-*NA* objective.

Solutions involving strong superresolution and greatly increased side-lobe intensity are not appropriate for certain applications -- like multi-photon direct laser writing, for which photo-patterning near the geometric focus would be dominated by the lobe(s) having the highest intensity, irrespective of proximity to the geometric focus. The goal of this work was to develop and apply an MGP-based algorithm that could identify binary phase-only DOEs that produced strong axial superresolution with minimal increase in side-lobe intensity. This is a much more stringent requirement than is commonly applied in PSF engineering problems, but its successful implementation here demonstrates the generality and broad utility of the present method.

Table 3.1 presents the parameters for some solutions that were identified as optimum in terms of both achievable superresolution and controlled side-lobe intensity. These results were selected from among all MGP solutions by inspecting G, M, and NMSE, and they represent the sub-set which offers a satisfactory compromise between decreasing G and maintaining small M. Also included in the table are the parameters for the starting DOE used to achieve each solution; the values of G, M, and NMSE before binarization; and the corresponding values of G and M after binarizing the DOE (listed as G_B and M_B). These data show that the binarization algorithm successfully transforms the analog phase-only DOE into a $0/\pi$ phase-only DOE with minimal change in the values of G and M in most cases. Close inspection of those situations for which G changed significantly revealed that the central lobe did remain superresolved, but its intensity decreased below that of the side lobes. According to the conventions applied here this circumstance has $M = 1$ and G_B ends up being computed based on the width of the side-lobes, as these are the most intense peaks in the axial PSF.

Table 3.1. Summary of parameters used to construct starting DOEs that yielded optimum performance in terms of both G and M, before and after binarization. The symbol "-" indicates that the parameter is not relevant for that DOE type.

Type	F	DC %	N_Z	PM	m	q_c	q_{peak}	$\Delta_{1/e}$	G	M	G_B	M_B
Constant	5	30	-	-	-	-	-	-	0.73	0.60	0.69	0.70
	6	70	-	-	-	-	-	-	0.73	0.55	0.72	0.50
	10	70	-	-	-	-	-	-	0.69	0.82	0.73	0.52
	11	80	-	-	-	-	-	-	0.72	0.61	0.72	0.61
Geometric (starting at edge)	-	10	-	1.2	-	-	-	-	0.73	0.53	0.82	1.00
	-	40	-	2.2	-	-	-	-	0.71	0.65	0.70	0.61
	-	65	-	1.35	-	-	-	-	0.71	0.62	0.67	0.97
	-	75	-	2.75	-	-	-	-	0.70	0.66	0.82	1.00
Geometric (starting at q_c)	-		-	1.0	1.5	0.70	-	-	0.71	0.67	0.70	0.91
	-	50	-	1.7	1.5	0.84	-	-	0.70	0.67	2.35	1.00
	-	50	-	2.4	1.5	0.84	-	-	0.69	0.69	0.75	0.59
	-	80	-	1.9	1.5	0.90	-	-	0.70	0.67	1.42	1.00
Gaussian	-	75	-	-	-	-	0.60	0.12	0.69	0.65	0.73	1.00
	-	75	-	-	-	-	0.75	0.06	0.71	0.64	0.70	0.64
Equal-area zones	-	-	5	-	-	-	-	-	0.73	0.57	0.71	0.52

Figure 3.5 shows the PSF for the result offering the best compromise between superresolution and suppressed side-lobes. This solution was found by initiating the MGP using a $0/\pi$ phase-only profile having five equal-area zones (Fig. 3.3(D)). The algorithm converged to an analog-phase DOE after only 40 iterations. Binarizing the analog-phase solution yielded the final $0/\pi$ phase-only DOE shown in Fig. 3.6. The binary DOE is comprised of 11 zones having different widths, unequal areas, and non-periodic spacing. Most of the zones are located on the outer periphery of the DOE radius. The calculated PSF is superresolved by 29% with $G = 0.71$, and the normalized intensity of the side-lobes is held to 0.52 or less. It is noteworthy that the solution DOE has minimal effect on the transverse PSF, as the central lobe transverse FWHM increases by only 4% relative to the diffraction limit.

The performance of the resulting axial PSF compares well with that reported for other axial sectioning techniques. For example, it is well known that a confocal pinhole is not strictly required in two-photon-excitation (TPE) imaging because it is inherently an optical sectioning technique. Nonetheless, Higdon *et al*. [69] have shown that including the pinhole does improve the axial resolution by ~30%; however, this is done at the expense of the signal-to-noise ratio. Nearly the same degree of axial superresolution could be achieved by placing the DOE of Fig. 3.6 before the objective, but now with no degradation of the signal-to-noise ratio because all available signal would reach the detector. Further, this axial resolution enhancement is comparable to that obtained with more complicated multi-arm detection systems [70]. Improving the axial superresolution without significantly widening the transverse distribution gives the PSF a more spherical 3D distribution about the geometric focus. A spherical 3D PSF is highly desirable in
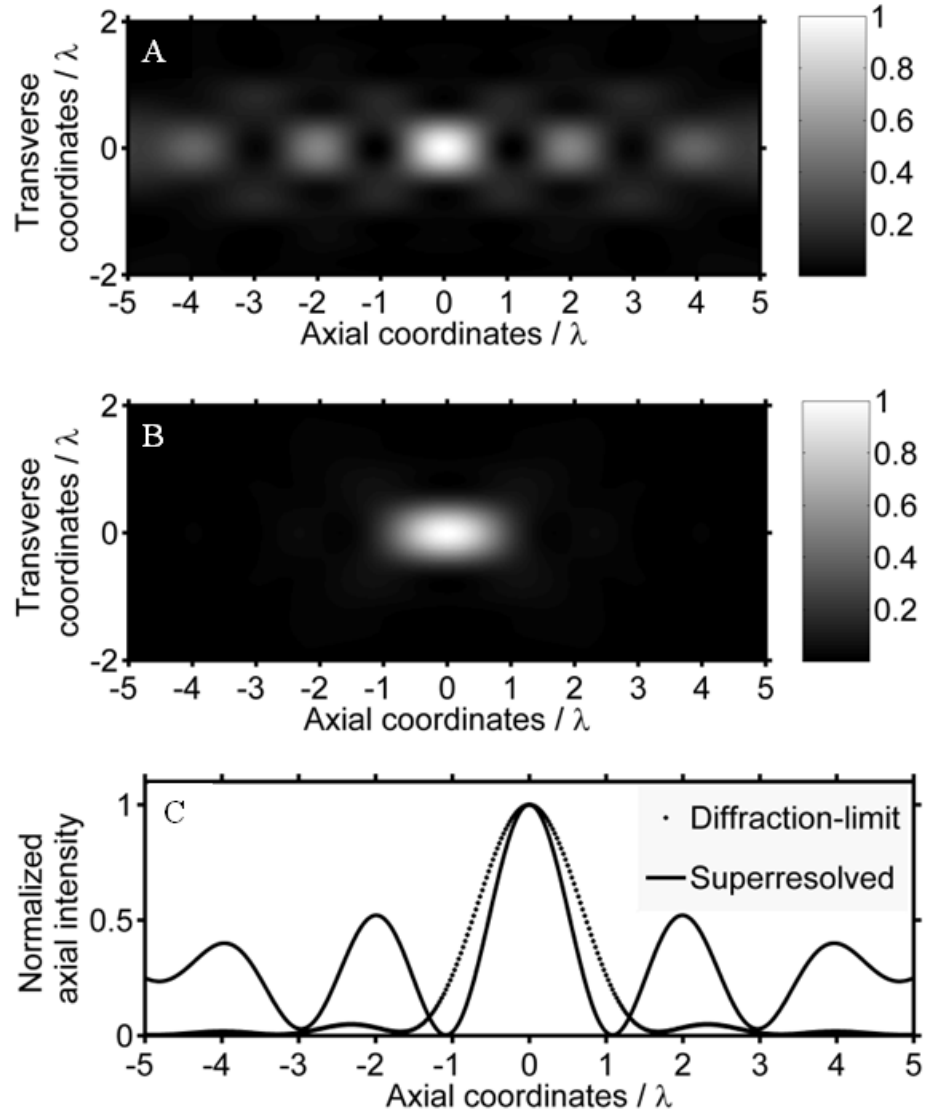
Figure 3.5. Calculated PSFs obtained (a) when the 11-zone DOE of Fig. 3.6 is placed before a 1.4-*NA* lens and (b) under diffraction-limited focusing (no DOE). The plots show the normalized axial and transverse intensity distribution about the geometric focus within the plane of polarization (xz-plane). (c) Intensity along the optical axis.

imaging, direct laser writing, and laser-tweezer applications. The quasi-spherical PSF shown in Fig. 3.5(A) is similar to that obtained in a TPE system employing a shaded ring filter [50] or 3R amplitude filter [51] with a lower *NA* = 1.2 water immersion objective lens.
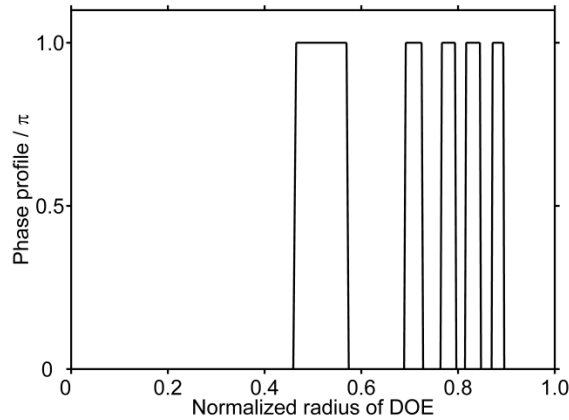
Figure 3.6. Phase profile of the 11-zone binary phase-only DOE that generates the PSF shown in Fig. 3.5(A). The phase is plotted versus the normalized real-space DOE radius.

Several different constraints were considered in this work; however, those described by expression (3.6) yielded the best results overall. Although not every result satisfied the required constraint completely, a superresolved central lobe was achieved in *circa* 95% of simulations. This is an indication of the robustness of the applied constraints, independent of the starting conditions.

It should be noted that the result to which the MGP algorithm converges is very sensitive to the shape of the applied constraint. To illustrate this point we show how simply excluding the zero-region from the constraints dramatically impacts the resulting solutions. The MGP algorithm was initiated twice using these two possible constrains, depicted graphically in Figs. 3.7(A and B), and otherwise all other conditions were identical. The starting DOE was the five-zone equal-area DOE shown in Fig. 3.3(D).

The axial intensity distributions generated by the binarized solution DOEs are shown in Figs. 3.7(C and D). It is apparent that the zero-region plays an essential role in forcing

45
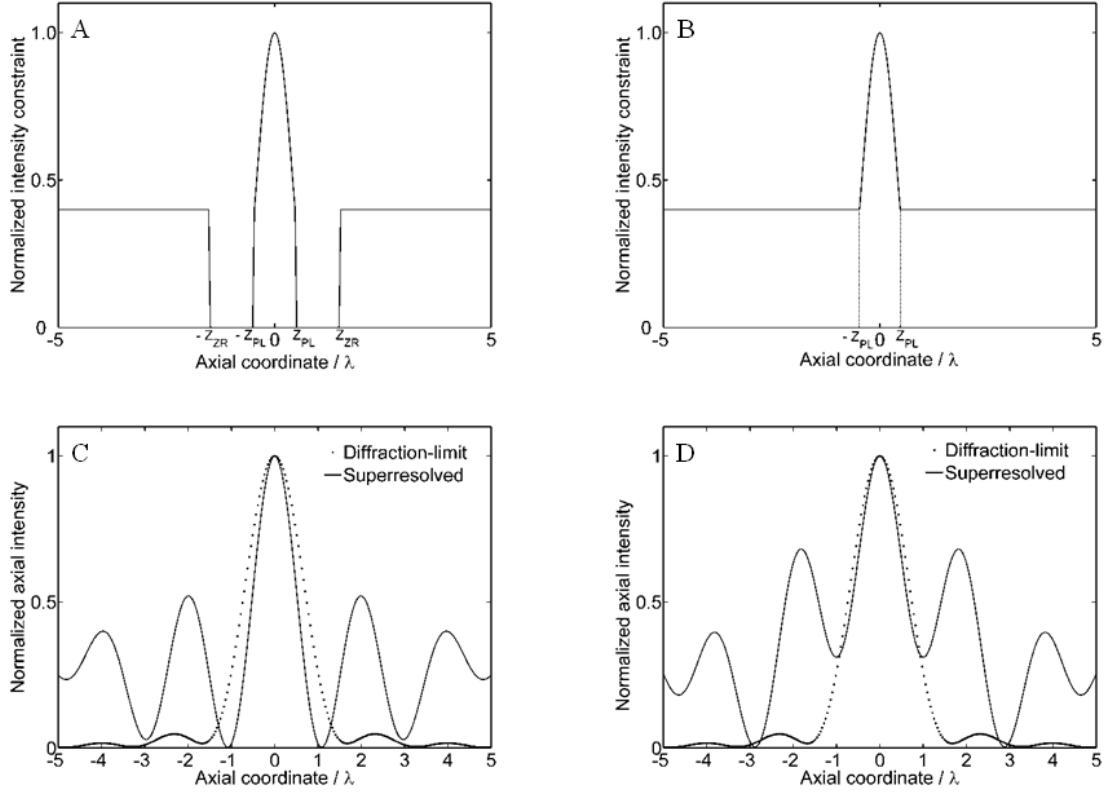
Figure 3.7. Effect of the zero-region constraint on the achieved axially super-resolved intensity distribution. Profile of the field domain constraint (A) with and (B) without the zero-region. (C and D) Axial intensity distribution achieved by applying the constraints in (A) and (B), respectively.

the minima surrounding the central-lobe down to zero as well as keeping them drawn in near the geometric focus. This type of constraint then is fundamental to achieving axial superresolution. The width of the zero-region set by values Z_{PL} and Z_{ZR} was found to be equally important. Studying the role of constraint parameters a, Z_{PL}, and Z_{ZR} in defining the axial intensity modulation is a subject of future work.

3.9 Conclusion

The method of generalized projections was adapted to vector diffraction theory to develop an algorithm for synthesizing diffractive optical elements that controllably

46

modify the axial intensity distribution under high numerical aperture focusing. The algorithm identifies solutions that most closely satisfy a set of potentially inconsistent constraints defined in terms of a targeted axial intensity distribution. A binarization procedure is incorporated into the algorithm to transform initial analog-phase solutions into binary $0/\pi$ phase profiles that are simpler to fabricate and in some cases offer improved performance. Application of the algorithm is demonstrated in the synthesis of DOEs that superresolve the focused axial intensity distribution with minimal increase in axial side-lobe intensity. It is shown that a systematic set of starting conditions can be used to initiate the computation so that multiple paths in the solutions space are explored thereby avoiding traps and tunnels associated with non-global solutions. The algorithm converges for 95% of the starting conditions, which indicates that the approach is robust. The solution identified as offering the best compromise between superresolution and side-lobe intensity is characterized by $G = 0.71$ and $M = 0.52$ and exhibits increase in the transverse spot size relative to the diffraction limit. This solution can provide not only an enhanced axial resolution but also a more isotropic focal intensity distribution, which is useful in several focused-beam applications. The algorithm is general and may be used to synthesize phase-only DOEs that generate other axial field distributions including extended depth of focus and multi-focal points.

CHAPTER 4: PARTICLE SWARM OPTIMIZATION OF AXIALLY SUPERRESOLVING BINARY PHASE DIFFRACTIVE OPTICAL ELEMENTS

4.1 Background

The MGP optimization algorithm discussed in chapter 3 can generate axially superresolving DOEs; however, it cannot guarantee an optimal solution. This is due to the presence of traps and tunnels which hinders the migration of the solution towards the global minimum. The MGP updates its variables without any randomness and thus it lacks the means to escape the local minima. The most utilized routines for solving multidimensional nonlinear optimization problems incorporate some random or heuristic strategies. Among these routines are the genetic and simulated annealing algorithms [71]. In an iterative optimization algorithm, a random modification to the current optimal value can force it to change place in the solution space; thus, exploring more possibilities. Such an operation allows the algorithm to escape local solutions. In this chapter we show how Particle Swarm Optimization (PSO) [72, 73], a population based stochastic optimization technique, can be used to design axially superresolving DOEs characterized by a performance superior to that obtained by the MGP. The results suggest that the algorithm returns optimal results.

4.2 Performance parameters

The DOE performance is characterized by the superresolution factor G and the relative side lobe intensity M of the resulting axial PSF as defined in previous chapters. Designing superresolving DOEs is a multi-objective problem in which G and M are co-minimized. All possible DOE solutions can be mapped onto a fitness-parameter space of

G versus *M* that is bound in the lower-left corner by a curve called the *Pareto front* [74].

The Pareto front defines a family of solutions that are global in terms of minimizing *G* for

a given *M*.

4.3 Problem formulation

The focused axial intensity distribution $I(z) \propto |E_{axial}(z)|^2$ generated by a rotationally

symmetric DOE can be calculated using the non-paraxial diffraction integral of Eq. (3.1),

which is reproduced here as:

$$E_{axial}(z) = iA \int_{q(0)}^{q(\alpha)} T(q)\sqrt{q}(1+q)\, e^{ikzq}\, dq \,, \tag{4.1}$$

The DOE modifies the wavefront according to complex transfer function $T(q)$. Equation

(4.1) may be expanded for an *N*-zone binary phase DOE as

$$E_{axial}(z) = iA \left[\pm \int_{q_0}^{q_1} \sqrt{q}(1+q)\, e^{ikzq}\, dq \; \pm \int_{q_1}^{q_2} \sqrt{q}(1+q)\, e^{ikzq}\, dq... \pm \int_{q_{N-1}}^{q_N} \sqrt{q}(1+q)\, e^{ikzq}\, dq \right]. \tag{4.2}$$

The sign of a given integral is positive (negative) when the phase shift of the

corresponding zone is 0 (π), so any DOE can be succinctly represented by an *N*-bit binary

position vector $X = \{x_1, x_2,..., x_N\}$. Axial superresolution can be viewed as a

combinatorial problem in which fields originating from each zone are added or subtracted

to give E_{axial} having minimum *G* for a fixed limit on side lobe intensity M_{lim}.

4.4 Particle swarm optimization (PSO)

PSO is a nature-inspired method for optimizing nonlinear functions motivated by the

idea that individuals in a population can evolve based on information gathered through

their own experience and that of the group [72]. The individuals and the group are referred to as *particles* and the *swarm*, respectively. During optimization a randomly generated swarm searches the solutions space for the "best" solution. Each iteration, solutions are compared using a fitness parameter, and the position and velocity of the *i*-th particle are updated based on the best solution it found, b_i, and the overall best position b_G found by the swarm. The comparison and update are applied to all particles and repeated over many iterations. The update process is then an aggregated acceleration of the *i*-th particle towards the best position identified by the ensemble.

4.5 Binary PSO applied to DOE design

We seek a DOE having a binary phase-only profile. As such, binary PSO (BPSO) [73] is best suited for the current problem. In BPSO each particle's position is represented by an *N*-bit binary vector. The particle velocity is a vector of *N* reals that can be thought of as giving the *probability* that bits change state as the position vector updates. In our adaptation of BPSO, each particle represents a candidate DOE in the *N*-zone solutions space using a position vector of form *X*. The modified discretized position and velocity update equations are [75, 76]

$$v_{ij}^{m+1} = w \left| v_{ij}^{m} \right| + C_1 r_1 (b_{ij} \oplus x_{ij}^{m}) + C_2 r_2 (b_{Gj} \oplus x_{ij}^{m}), \qquad (4.3)$$

$$if \ [r_3 < S(v_{ij}^{m})] \ then \ x_{ij}^{m+1} = \bar{x}_{ij}^{m} \ else \ x_{ij}^{m+1} = x_{ij}^{m}. \qquad (4.4)$$

Index m counts the iteration number. The best solutions found by the i-th particle and the swarm are given by b_i and b_G, respectively, which are also position vectors of form X. The symbols "$\oplus$" and "$\bar{}$" (overscore) indicate the *exclusive-or* and *not* operations, respectively. The terms in Eq. (4.3) are a weighted sum v_{ij}^{m+1} which gives the probability that the j-th bit changes state in iteration $m+1$, based on the magnitude of the particle's current velocity $|v_{ij}^{m}|$ and a comparison of its current position to b_i and b_G. The relative contribution of these terms is set by the "acceleration coefficients" C_1 and C_2 and the "inertia weight" w, which span the real numbers. To ensure the solutions space is adequately explored, coefficients r_1, r_2, and r_3 are random numbers that update in each iteration and take any real value equally distributed in the range [0, 1]. After evaluating Eq. (4.3), the constraint $v_{ij}^{m+1} \leq V_{max}$ is applied to control the range explored by the particle and thus the convergence rate. In Eq. (4.4) the sigmoid function S maps velocity v_{ij}^{m} onto [0, 1] for comparison to r_3. The outcome of this comparison determines the updated value of the j-th bit.

4.6 Multi-objective optimization with BPSO

Given that G and M must be co-optimized, the BPSO algorithm is structured as a multi-objective optimization [74]. The algorithm is initialized by defining M_{lim}, generating a swarm having random starting positions and velocities, and setting b_i to x_i for all particles. The solutions space is then iteratively searched as follows. The associated DOEs and resulting PSFs are calculated using Eq. (4.2) along with the corresponding fitness values $G(x_i)$ and $M(x_i)$. If both $M(x_i) < M_{lim}$ and $G(x_i) < G(b_i)$, then the current solution is regarded as superior to the particle's previous best, so b_i is set to x_i.

51

Similarly, b_G is replaced by any b_i for which $M(b_i) < M_{lim}$ and $G(b_i) < G(b_G)$. The velocities and positions are updated using Eqs. (4.3) and (4.4). The process is repeated for a fixed number of iterations. Because G and M cannot be simultaneously minimized, the swarm migrates toward solution b_G for which G is minimized at $M \rightarrow M_{lim}$.

4.7 PSO parameters selection

The values of the weighting and acceleration coefficients affect the convergence and final results of a PSO algorithm [77]; however, previously reported studies of this dependence could not be directly applied to the modified algorithm reported here. Through empirical study we found that good solutions are obtained when $w = -1$, $C_1 = 4$, $C_2 = 4$, and $V_{max} = 6$. Other values yield satisfactory results, but the structure of Eqs. (4.3) and (4.4) require $w < 0$ and positive C_1 and C_2. This may be understood as follows. Because a sigmoid function is used in Eq. (4.4), a bit is most likely to remain unchanged when its velocity is most negative. Consider the case in which the j-th bit's state matches that of best solutions b_i and/or b_G. The probability for inverting this bit should then remain low. Because the $\oplus$ operator is used in Eq. (4.3), the second and/or third terms will vanish, so the bit only remains unchanged when $w < 0$. Conversely, if the current bit differs from that of b_i and/or b_G, the velocity will become more positive, increasing the probability for bit inversion.

4.8 Results and discussion

The BPSO algorithm was applied to optimize a DOE positioned immediately before an aberration-free 1.4-NA oil-immersion objective ($n = 1.516$) having a 9.3-mm diameter entrance pupil. The DOE is uniformly illuminated by linearly polarized monochromatic

plane waves having a vacuum wavelength of 800 nm. A swarm of 40 particles was updated through 10,000 iterations with $N = 100$. Equation (4.2) was derived with Kirchhoff's boundary condition applied to the DOE, which is satisfied when all zones have a lateral width of at least 20λ [78]. This requirement is easily satisfied given that the 100-zone DOE has a minimum zone width of 35λ.

4.8.1 Simulation example

Figure 4.1(A) shows the PSF generated by a DOE optimized with $M_{lim} = 0.5$. The central lobe FWHM is decreased by 34% relative to the diffraction-limited pattern [Fig. 4.1(B)], and the relative side lobe intensity is held below 0.5. To our knowledge this is the highest single-beam axial superresolution calculated for a phase DOE with the given limit on side lobe intensity. Interestingly, the transverse FWHM of the central lobe increases by only 5% with respect to the diffraction limit, so lateral resolution is not sacrificed. The combined effect of axial superresolution and minimal transverse broadening causes the central lobe to become more spherical. The ratio of the transverse to axial FWHM is 0.78. A more spherical PSF is desirable for many focused laser applications, such as multi-photon imaging and direct laser writing.

In chapter 3, we reported the first vectorial algorithm for optimizing G and M based on the Method of Generalized Projections (MGP). Although that method yields excellent results, it is not guaranteed to find the global solution because MGP is susceptible to "traps" and "tunnels" that can cause the algorithm to stagnate in local minima. The best DOE found using MGP offers $G = 0.71$ and $M = 0.52$. It is noteworthy that BSPO outperforms MGP by finding a solution that offers both higher superresolution and

smaller side lobes. This can be attributed to the well known ability of PSO to avoid becoming trapped in local minima [72, 73].
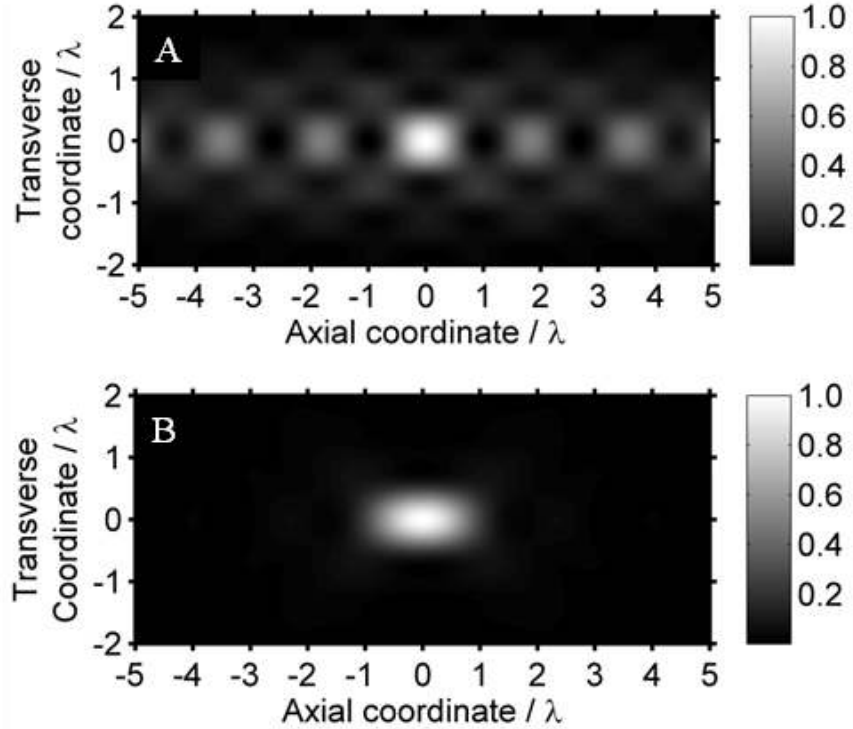
Figure 4.1. PSF within the plane of incident polarization resulting (a) with a BSO-designed superresolving DOE ($G = 0.66$, $M = 0.50$) and (b) for diffraction limited focusing (no DOE).

4.8.2 Pareto front

An analytic expression does not exist for the minimum G that can be achieved given M_{lim}. However, this relationship can be obtained numerically by running the BPSO using a range of M_{lim} values. Figure 4.2 shows G versus M obtained when M_{lim} was varied from 0.05 to 1.00 by steps of 0.05. The PSFs vary from being highly superresolved ($G = 0.59$) with strong side lobes ($M = 0.99$) to having minimal superresolution ($G = 0.90$) and weak side lobes ($M = 0.049$). The G-M pairs define a curve that lies in the lower left

corner of the fitness space, as expected for the Pareto front of a co-minimization problem. "Neighborhood search" [76] was also incorporated in each iteration by examining G and M resulting as each bit of a given particle was inverted. Similar results were obtained with this modification, yet the convergence rate was significantly decreased. This new BPSO algorithm can be used then to design axially superresolving DOEs in which both G and M are co-optimized for high-NA applications.
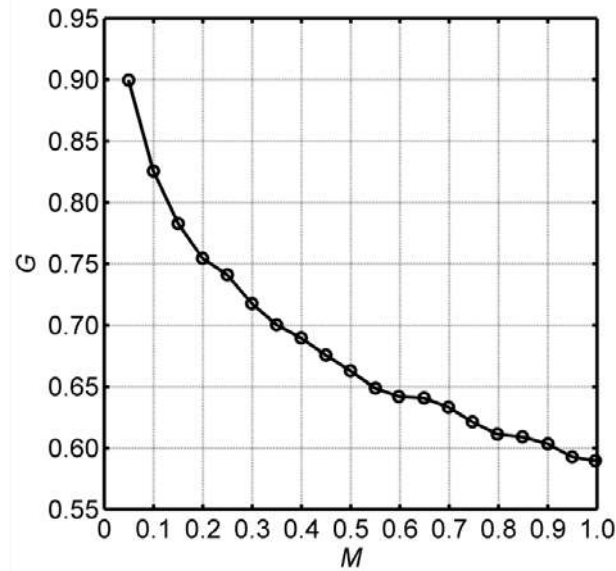
Figure 4.2. Pareto front of the G-M fitness space for axially superresolving binary phase DOEs.

4.9 Conclusion

The particle swarm optimization (PSO) algorithm was used to design binary phase-only diffractive optical elements (DOEs) that superresolve the axially focused PSF. PSO is relatively an easy algorithm to code and yet can achieve optimal results. To the best of

our knowledge, the performance of the DOE designs achieved with PSO is better than any reported in the literature.

CHAPTER 5: VECTORIAL BEAM SHAPING

5.1 Background

In this chapter, the DOE design is applied to modify the irradiance distribution at the focal plane of a high-NA focusing optical system instead of along the optic axis. Transverse beam shaping can enhance performance in optical lithography, laser-based materials processing, direct laser writing, surgical applications, and optical data storage [16, 79-81]. For many applications the optimum irradiance distribution consists of a flat-top profile having a defined geometry within the focal plane. Such irradiance patterns can be characterized in terms of the diffraction efficiency, κ, and uniformity error, δ. The diffraction efficiency quantifies the fraction of total optical power directed into the targeted region of interest and the uniformity error provides a measure of flatness in the irradiance distribution across that region.

Many excellent scalar techniques have been reported for designing beam shaping DOEs. These approaches are based on methods that include geometric mapping [82, 83], analytical solution [84], iterative processes [63, 85-87], and genetic optimization [88]. Although exceptional results have been achieved with these algorithms, they are all based on scalar diffraction theory and as such are only valid in the paraxial domain of diffractive optics [89]. For systems with high NA, depolarization effects are significant [3], so vectorial diffraction theory must be used in the DOE design process. This becomes particularly challenging because the overall beam shape is determined by the summed irradiance of the x-, y-, and z-polarized electric fields. Although the field components are orthogonal, they are not entirely independent because each is reshaped

by a common DOE. As a result, the DOE must be designed to reshape the field components collectively so the irradiance of the *total* field matches the targeted beam shape. One report of vectorial beam shaping has appeared, but the method was not applied to high-*NA* systems [90]. Given that high-*NA* systems are being increasingly employed in frontier technologies, further applications of beam shaping will be stymied unless accurate methods for vectorial beam shaping are developed.

In this work we report a vectorial beam shaping algorithm that can be used to design phase-only DOEs for use under high-*NA* conditions. The algorithm was developed by incorporating the vector diffraction integrals [4] into the Method of Generalized Projections (MGP) [62]. The diffraction integrals are used to interrelate the DOE phase profile and the resulting vectorial electric field in the focal plane. The integrals are evaluated using the chirp-z transform [91] to improve computational speed and accuracy. Iterative projection of constraints in the pupil and focal planes progressively forces the simulation toward a DOE phase profile that generates the targeted beam shape. The new algorithm is applied to the problem of designing a phase-only DOE that transforms a circularly apodized flat-top input beam into a square flat-top irradiance distribution when focused using a 1.4-*NA* objective. In beam shaping, high diffraction efficiency and low uniformity error are known to be mutually exclusive characteristics that must be considered jointly in optimizing DOEs [92]. In this work, we also investigate how κ and δ change as the size of the focused beam profile approaches the diffraction-limited spot size.

5.2 Theory of beam shaping

5.2.1 Vector diffraction integrals

   The optical geometry of the focusing system is depicted in Fig. 5.1. A DOE and an aberration free aplanatic lens which fulfils the sine condition and has focal length f are positioned such that their optical axes are collinear with the z-axis of a Cartesian coordinate system whose origin is located at the Gaussian focus of the lens. The numerical aperture of the lens is $NA = 1.4$ in all calculations. Monochromatic linearly polarized plane waves, with electric field vector parallel to the x-axis, propagate along the z-axis, passing through the DOE and entering the pupil of the lens. The light focuses into a medium of refractive index $n = 1.516$. In the absence of the DOE, this situation is consistent with common applications of high-NA oil-immersion objective lenses.

   The electric field at an arbitrary point $P(x_f, y_f)$ in the focal plane ($z_f = 0$) can be calculated using the vector diffraction integrals [4, 93] given as (see Appendix A for details) :

$$\begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = -i \frac{E_{in} f}{2\pi} \iint\limits_{\sqrt{k_x{}^2 + k_y{}^2} \le k_{max}} T(k_x, k_y) e^{i\Phi(k_x, k_y)} \sqrt{\frac{k_z}{k_t}} \frac{1}{k_z} \begin{bmatrix} \dfrac{k_x{}^2 (k_z/k_t) + k_y{}^2}{k_x{}^2 + k_y{}^2} \\[2ex] \dfrac{k_x k_y (k_z/k_t - 1)}{k_x{}^2 + k_y{}^2} \\[2ex] -\dfrac{k_x}{k_t} \end{bmatrix} \times e^{i(k_x x_f + k_y y_f)} dk_x dk_y \cdot \qquad (5.1)$$

59
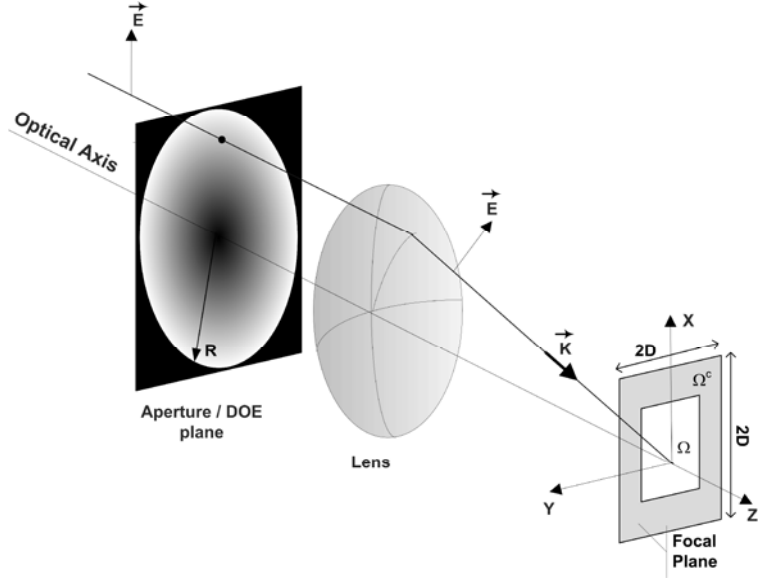
Figure 5.1. Optical setup of the beam shaping problem. The aperture represents the input pupil of the objective lens. The focal plane is divided into two regions. Ω represents the region of interest that contains and bounds the targeted beam shape. Its complement Ω^c represents the remainder of the focal plane.

At any point within the optical system the irradiance is $I = (1/2)nc\varepsilon_0|E|^2$. The reshaped beam is the spatial map of the focused irradiance $I_f(x_f, y_f)$ for all P. The speed of light and electric permittivity in vacuum are c and ε_0, respectively. The wave number of light transmitted through the lens is $k_t = 2\pi/\lambda = [k_x^2 + k_y^2 + k_z^2]^{1/2}$, with k_x, k_y, and k_z being the plane wave components, and λ is the wavelength within the medium. The NA of the lens system sets $k_{max} = k_t NA/n$. The function $T(k_x, k_y)\exp[i\Phi(k_x, k_y)]$ describes the transmission amplitude (T) and phase (Φ) of the DOE. The amplitude of the incident electric field E_{in} is assumed to be spatially constant, so this term was brought outside the integral. The focal plane is divided into a region of interest Ω and its complement Ω^c. The region Ω wholly contains and bounds the targeted beam shape I_t.

60

5.2.2 Normalization

It is helpful to cast the vector diffraction integrals into a form consisting of dimensionless variables by normalizing to the lateral extent of the input beam I_{in} and the targeted beam profile I_t [94]. The Cartesian coordinates of the aperture plane (x_a, y_a) can be related to the x- and y-components of the wave vector by

$$\frac{x_a}{R} = \frac{k_x}{k_{max}} = u \quad \text{and} \quad \frac{y_a}{R} = \frac{k_y}{k_{max}} = v, \tag{5.2}$$

where R is the radius of the entrance pupil, and u and v represent the normalized k_x and k_y components of the wave vector, respectively. Likewise, the focal plane coordinates (x_f, y_f) are normalized by $D = m\lambda$ (see Fig. 5.1), giving the transformed focal plane coordinates (ξ, η):

$$\xi = \frac{x_f}{D} = \frac{x_f}{m\lambda} \quad \text{and} \quad \eta = \frac{y_f}{D} = \frac{y_f}{m\lambda}. \tag{5.3}$$

The size of I_t can be scaled conveniently by m multiples of λ. All free parameters of the system can be combined into the single variable, β, given by [94]:

$$\beta = \frac{2\pi DR}{\lambda f} = \frac{2\pi D(NA)}{\lambda n}. \tag{5.4}$$

Substituting Eqs. (5.3) and (5.4) into Eq. (5.1) and algebraically manipulating gives the following normalized vectorial diffraction integrals for the field at the focal plane:

$$
\begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = -i \frac{E_{in}f}{\lambda} \left(\frac{NA}{n} \right)^2 \iint\limits_{\sqrt{u^2+v^2} \leq 1} T(u,v) e^{i\Phi(u,v)} \frac{1}{\sqrt{q}} \begin{bmatrix} \dfrac{u^2q+v^2}{u^2+v^2} \\[2mm] \dfrac{uv(q-1)}{u^2+v^2} \\[2mm] -\dfrac{NA}{n}u \end{bmatrix} \times e^{i(u\omega_x+v\omega_y)} \, dudv \,, \quad (5.5)
$$

Where $q(u, v) = [1 - (NA/n)^2(u^2 + v^2)]^{1/2}$, and $\omega_x = k_{max}D\xi = \beta\xi$ and $\omega_y = k_{max}D\eta = \beta\eta$ have been introduced so the diffraction integrals in Eq. (5.5) can be evaluated as a Fourier transform.

5.2.3 Chirp-z transform

In this work, the double integrals of Eq. (5.5) were evaluated using a two-dimensional (2D) chirp-z transform (CZT) with 512×512 sampling points in both the aperture and the focal planes, irrespective of the magnitude of R and D. Although 2D fast Fourier transform can be used, CZT is computationally faster and better suited for the present situation because it internalizes zero-padding and allows the spacing of sampling points in the aperture and focal planes to be set independently [95]. This greatly reduces the number of sampling points required when R and D differ substantially in magnitude, as in the present case.

Evaluating Eq. (5.5) effectively propagates the field $E_x(u, v)$ forward into $E_x(\xi, \eta)$, $E_y(\xi, \eta)$, and $E_z(\xi, \eta)$. The fundamental operation of beam shaping involves applying constraints associated with I_t to the field in the focal plane and then calculating a new DOE phase profile that comes closest to generating the reshaped field. A new DOE

transmission profile is obtained by "backward propagating" to the pupil plane through an inverse-CZT applied to the reshaped field.

Because the forward and inverse CZTs are applied to a finite region of the DOE and focal planes, Gibbs artifacts are generated [91]. If these numerical errors are allowed to accumulate, they can degrade the uniformity of I_f or even cause the algorithm to diverge from a solution. Gibbs artifacts were suppressed by applying a Kaiser window to the amplitude of the focal field profiles immediately after they were computed using CZT [91].

5.2.4 Method of generalized projections

For a given input beam I_{in} we seek a DOE phase function Φ that generates a focal plane field distribution such that the sum of the x-, y-, and z-polarized irradiance $I_x + I_y + I_z = I_f$, matches the targeted irradiance I_t for all (ξ, η). An exact match is generally not possible because it is not known *a priori* that an arbitrary I_t is a solution to the wave equation [63]. This is the case for the present example because the targeted square irradiance profile requires a discontinuous drop in the field at the interface of Ω and Ω^c. The problem is further complicated because $\Phi(u, v)$ affects each of E_x, E_y, and E_z, so the field components are not truly independent. Evaluating Eq. (5.5) and integrating I_x, I_y, and I_z over the entire focal plane shows that their fractional power content is 0.74, 0.01 and 0.25, respectively, and these values are independent of $\Phi(u, v)$. So, high-NA beam shaping demands that I_x, I_y, and I_z are optimized collectively. This problem cannot be solved analytically, so iterative numerical techniques must be employed. The MGP is particularly well suited to the current problem because it can find solutions that closely satisfy sets of inconsistent and non-physical constraints [53]. In the MGP the optical

63

field is repeatedly propagated forward and backward between the DOE and focal planes, and constraints associated with both domains are applied in each iteration until a satisfactory solution is found.

5.2.5 Starting conditions

The goal is to design a phase-only DOE, so the initial transmission amplitude T_0 is set to unity. The rate of convergence and quality of the solution can be greatly improved by initiating the vector diffraction algorithm using a well chosen starting DOE phase profile, $\Phi_0(u, v)$. Geometrical optics based methods can be used to identify suitable starting DOEs [82]. Geometrical transformations have been applied successfully to obtain $\Phi_0(u, v)$ analytically when I_{in} and I_t are either separable or axially symmetric [84]; however, the beam shaping example studied here is neither separable nor axially symmetric. To overcome this problem, a procedure described by Aagedal *et al*. [84] was employed to obtain $\Phi_0(u, v)$ as a combination of two separate DOEs that together achieve the required geometric beam transformation. The first element converts the axially symmetric I_{in} into a standard Gaussian beam, which is separable and axially symmetric. The second element converts the Gaussian beam into a square-shaped super-Gaussian beam. The resulting $\Phi_0(u, v)$ does not adequately reshape I_{in} to I_t under high-*NA*, but it provides a good starting point for the vector diffraction algorithm.

5.2.6 Algorithm flow

The analytically calculated starting DOE, $\Phi_0(u, v)$, is substituted into Eq. (5.5). The diffraction integrals are then evaluated using the CZT giving $|E_x|\exp(i\phi_x)$, $|E_y|\exp(i\phi_y)$, and

$|E_z|\exp(i\phi_z)$ in the focal plane, and the corresponding irradiance distribution is I_f. The diffraction efficiency and uniformity error are calculated for the solution as

$$\kappa = \frac{\iint\limits_{\Omega} I_f(\xi,\eta)\,d\xi d\eta}{\iint\limits_{\Omega+\Omega^c} I_f(\xi,\eta)\,d\xi d\eta} \tag{5.6}$$

$$\delta = \frac{\max[I_f(\xi,\eta)] - \min[I_f(\xi,\eta)]}{\max[I_f(\xi,\eta)] + \min[I_f(\xi,\eta)]} \quad \text{for } (\xi,\eta) \in \Omega. \tag{5.7}$$

Under these definitions, a "perfect" solution would have $\kappa = 1$ and $\delta = 0$.

Second, the constraint of the targeted beam shape I_t is applied. For the present example, I_t is defined as

$$I_t = \begin{cases} \alpha & \text{for } (\xi,\eta) \in \Omega \\ I_f & \text{for } (\xi,\eta) \in \Omega^c \end{cases}, \tag{5.8}$$

with the limits of Ω set to $-0.25 \le \xi, \eta \le 0.25$. Within the beam shaping area, total power is conserved and the irradiance is homogenized by setting the latter to its average value:

$$\alpha = \frac{\iint I_f\,d\xi d\eta}{\iint d\xi d\eta} \quad \text{for } (\xi,\eta) \in \Omega. \tag{5.9}$$

The *x*-polarized field within Ω is reshaped as

$$E_x(\xi,\eta) = \sqrt{\frac{2}{nc\varepsilon_0}\left[\gamma I_t(\xi,\eta) - I_y(\xi,\eta) - I_z(\xi,\eta)\right]}\exp[i\phi_x(\xi,\eta)] \quad \text{for } (\xi,\eta) \in \Omega, \tag{5.10}$$

whereas the following are left unchanged: $|E_x(\xi,\eta)|$ outside Ω, $|E_y(\xi,\eta)|$ and $|E_z(\xi,\eta)|$ across all $\Omega + \Omega^c$; and the phases $\phi(\xi,\eta)$ of all field components in $\Omega + \Omega^c$. γ is an adjustable scalar that augments $|E_x(\xi,\eta)|$ in Ω relative to that in Ω. This operation

65

provides a means for slowly pulling energy from Ω^c into Ω [85]. In this work $\gamma = 1.03$ was used for all iterations.

Third, an inverse CZT is applied to the reshaped $E_x(\xi, \eta)$ to generate a complex DOE function $T_{i+1}(u, v)\exp[i\Phi_{i+1}(u, v)]$. Given that a phase-only DOE is required, we apply this constraint by resetting the transmission amplitude to T_0 while retaining the phase. The new complex DOE transmission function becomes $T_0(u, v)\exp[i\Phi_{i+1}(u, v)]$. The electric field is then propagated forward again using the new DOE and the reshaped beam it generates is evaluated based on κ and δ. This process continues until the algorithm converges to a suitable solution or until a fixed number of iterations are completed.

5.3 Results and discussion

Equation (5.10) is intentionally configured so that the beam shaping constraint is only directly applied to the x-component of the field amplitude lying within Ω. This arrangement provides amplitude freedom outside the region of interest and phase freedom across the entire focal plane that help the algorithm converge to a solution [63, 68]. Additionally, it solves the problem of how to reshape three independent field components that are effectively coupled through a common DOE. The intended beam shape is applied repeatedly to the x-polarized field, as it contains the majority of the focused power, and only it is propagated backward to obtain the DOE phase function for the next iteration. The y- and z-polarized components of the focal field are reshaped indirectly when they are calculated by forward propagation through the new DOE. Repeated iterations effectively pull $E_x(\xi, \eta)$, $E_y(\xi, \eta)$, and $E_z(\xi, \eta)$ toward distributions that collectively satisfy $I_x + I_y + I_z \rightarrow I_t$.

We now discuss the results obtained when the vector diffraction algorithm was used to design a phase-only DOE that transforms a circularly apodized flat-top input beam into a focused square irradiance pattern of area $50\lambda \times 50\lambda$ ($D = 100\lambda$). Figure 5.2 shows the normalized focal irradiance distributions of the three polarization components, I_x, I_y, and I_z, and the total focal irradiance distribution I_f generated by the DOE phase profile of Fig. 5.3. This DOE and the associated irradiance distributions were obtained after 600 iterations. The overall beam shape is square as intended with $\delta = 7\%$ and $\kappa = 74.5\%$, indicating that it has good uniformity and power confinement within the region of interest.

In contrast, the irradiance distributions of the constituent polarizations are non-uniform. I_x most resembles the targeted profile, but appears doubly concave, as though squeezed along the x-axis. Although I_x is non-zero across the coordinate axes, I_y and I_z have node(s) at these positions where their field amplitudes drop to zero. I_y is most complex, appearing approximately four-fold symmetric with power concentrated in the corners of Ω. I_z exhibits two-fold symmetry with a single nodal plane lying along the y-axis. The regions of high irradiance in I_y and I_z fill in around the edges of the x-polarized profile making the total irradiance distribution I_f uniform and square. These profiles show that the vector diffraction algorithm successfully generates a DOE for which all polarization components of the field are reshaped concurrently to achieve a targeted irradiance distribution under high-NA focusing.
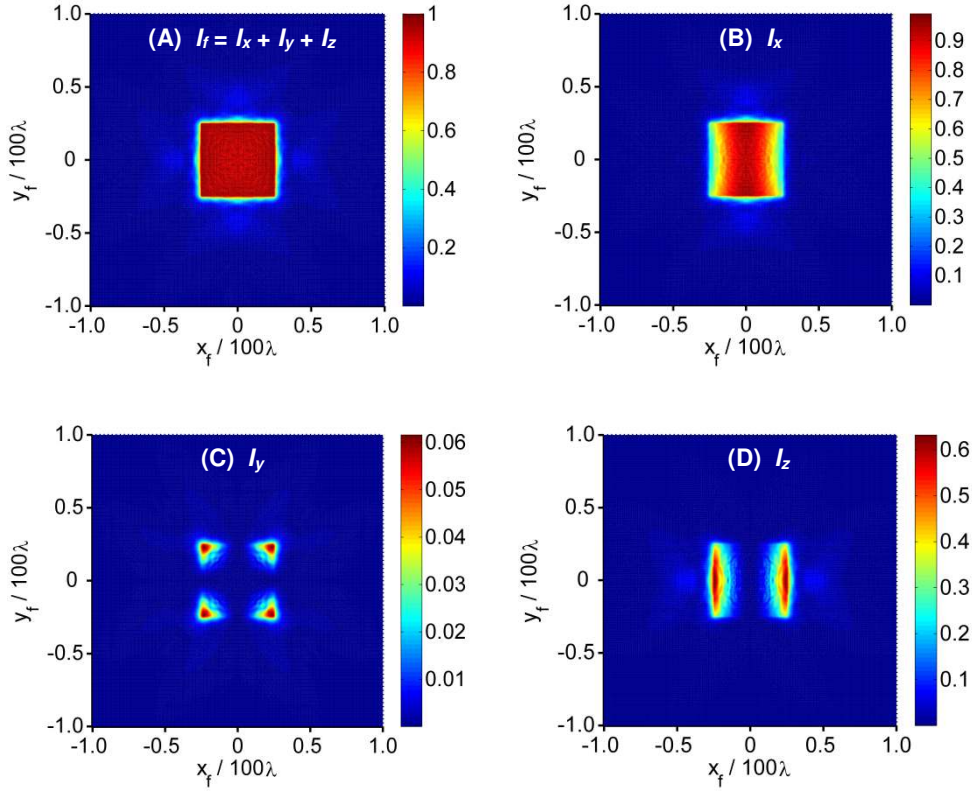
Figure 5.2 (A) Calculated irradiance distribution resulting when a circularly apodized flat-top input beam of radius R is passed through the phase-only DOE shown in Fig. 5.3 and focused using a 1.4-*NA* objective. The DOE was designed to reshape the beam into a flat-top square irradiance pattern of area $50\lambda \times 50\lambda$ ($D = 100\lambda$). (B) - (D) Irradiances of the constituent x-, y-, and z-polarized components of the total field. Each profile is normalized to the peak of I_f.

Figure 5.4 shows how the uniformity error and diffraction efficiency change during the calculation. The diffraction efficiency progressively increases because the parameter $\gamma = 1.03$ causes power to transfer from Ω^c into Ω with each iteration. On the other hand, the uniformity error drops rapidly and reaches an apparent plateau after circa 500 iterations. It is known that high uniformity and high diffraction efficiency are mutually exclusive characteristics in beam shaping [92]. As a result, attempting to improve the diffraction efficiency beyond the level of 74% achieved at approximately 600 iterations
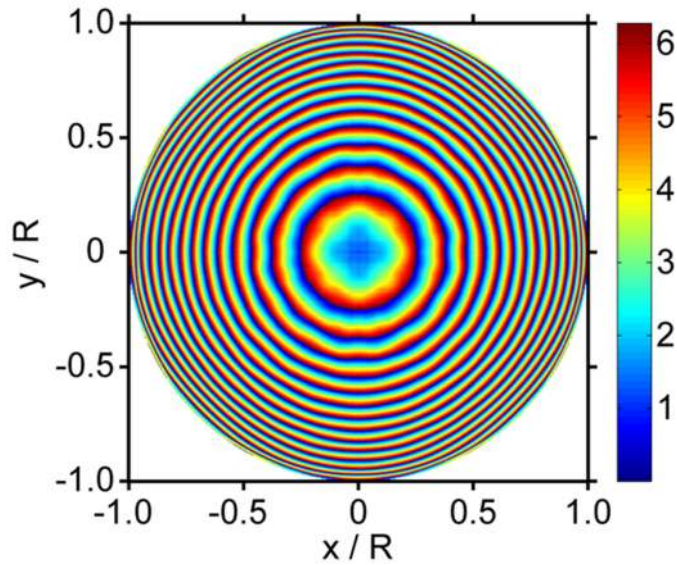
Figure 5.3 DOE phase profile that generates the focal irradiance distributions shown in Fig. 5.2. The phase is plotted in units of radians.

caused the uniformity to erode. Obtaining solutions that are optimized in terms of both δ and κ could be achieved by extending the present vector diffraction algorithm through Tikhonov regularization theory theory [92]. It is noteworthy that the diffraction efficiency and uniformity are very poor for the first iteration. This results because the starting DOE was designed using a geometrical transformation method, which does not account for the vector character of the field. It underscores the importance then of using vector diffraction theory to achieve accurate beam shaping under high-NA conditions, and it demonstrates the improvement that can be achieved in beam shaping using the present vectorial approach.

The problem of shaping a beam whose size approaches the diffraction limit was examined by repeating the calculations described above for $D = 50\lambda$, 25λ, 10λ, and 5λ.

69

The resulting focused irradiance distributions and the corresponding DOE phase profiles are shown in Fig. 5.5. Comparing the irradiance distributions reveals that the intended beam transformation can be achieved, even for targeted beam profiles having an edge-length of $D/2 = 2.5\lambda$ (see Fig. 5.5(G)). The DOEs themselves have approximate four-fold symmetry with respect to rotation about the z-axis, as expected for a square target beam shape (consider also Fig. 5.3). The DOEs are also comprised of many concentric rings of steadily increasing phase reminiscent of a Fresnel lens. These concentric phase rings effectively negate some of the focusing power of the high-NA lens, so understandably their number and radial density decreases as the target beam size is reduced toward the diffraction limit.

As D decreases, the reshaped beam degrades in uniformity and sharpness at the boundary of Ω. The sharpness of the irradiance profiles was characterized empirically by fitting the normalized individual distributions to a super-Gaussian of the form

$$I_f = \exp[-(x_f / a)^{2N} - (y_f / a)^{2N}]. \tag{5.11}$$
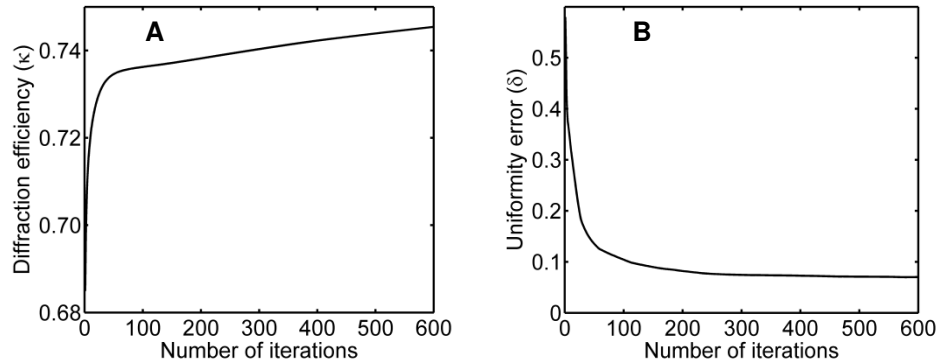
Figure 5.4 (A) Evolution of the beam shaping diffraction efficiency and (B) uniformity error versus iteration number for $D = 100\lambda$.

The order of the super-Gaussian, N, provides a measure of the sharpness of the beam profile at the boundary of the region of interest. Rather than applying Eq. (5.11) to all of $\Omega + \Omega^c$, the fitting was restricted to a region within which I_f exceeds 0.5. This procedure yields fits that more accurately describe the steepness of the profiles at the interface between Ω and Ω^c because it does not include irradiance fluctuations in Ω^c that are necessarily part of any real solution to the wave equation. The beam shaping parameters κ, δ, and N obtained for each value of D are collated in Table 5.1. The data show that the beam uniformity, diffraction efficiency, and edge sharpness all deteriorate as the size of the reshaped beam is reduced toward the diffraction limit.

The results in Fig. 5.5 suggest that it becomes increasingly difficult to achieve a targeted irradiance distribution when the beam size becomes comparable to the diffraction limit, as has also been observed for scalar beam shaping [96]. This phenomenon can be understood as a manifestation of the uncertainty principle [94].

Table 5.1 Performance parameters for the reshaped beams shown in Figs. 5.2 and 5.5 as function of the focused beam parameter D.

D	Diffraction efficiency (κ)	Uniformity error (δ)	Super-Gaussian order (N)	$\beta^2 \Delta I_{in} \Delta I_t$
100λ	74.5 %	7.0 %	16	6955
50λ	75.0 %	13.8 %	12	1739
25λ	72.4 %	25.9 %	13	435
10λ	65.5 %	28.0 %	6	70
5λ	55.5 %	26.5 %	5	17

The irradiance distribution I_t that can be achieved is inherently limited by the finite spatial bandwidth of the input beam I_{in} and the limited range of wave vectors over which focusing occurs, as quantified by the NA. The limit in the achievable beam shape can be expressed as [94]

$$\beta^2 \Delta I_{in} \Delta I_t \geq 1, \tag{5.12}$$

where

$$\Delta I_{in} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(u^2 + v^2\right) I_{in}(u,v) du dv}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_{in}(u,v) du dv} \quad \text{and} \quad \Delta I_t = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\xi^2 + \eta^2\right) I_t(\xi,\eta) d\xi d\eta}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_t(\xi,\eta) d\xi d\eta}. \tag{5.13}$$

The particular usefulness of β now becomes apparent. Given that β is determined by all free parameters of the system (R, D, λ, and NA), it provides a single measure of the difficulty of the beam shaping problem. Values of $\beta^2 \Delta I_{in} \Delta I_t$ for the beam profiles of Fig. 5.5 are also included in Table 5.1. Because $\beta^2 \Delta I_{in} \Delta I_t$ depends quadratically on D, it drops rapidly within this series and is most comparable to unity at $D = 5\lambda$, for which the reshaped beam quality is poorest. These data and Eq. (5.12) imply then that I_f will differ increasingly from I_t as the targeted beam size is decreased toward the diffraction limited spot size, with all other parameters kept fixed.

5.4 Conclusion

A vector diffraction algorithm was developed for designing phase-only DOEs that reshape beams focused under high-NA conditions. The algorithm accounts for
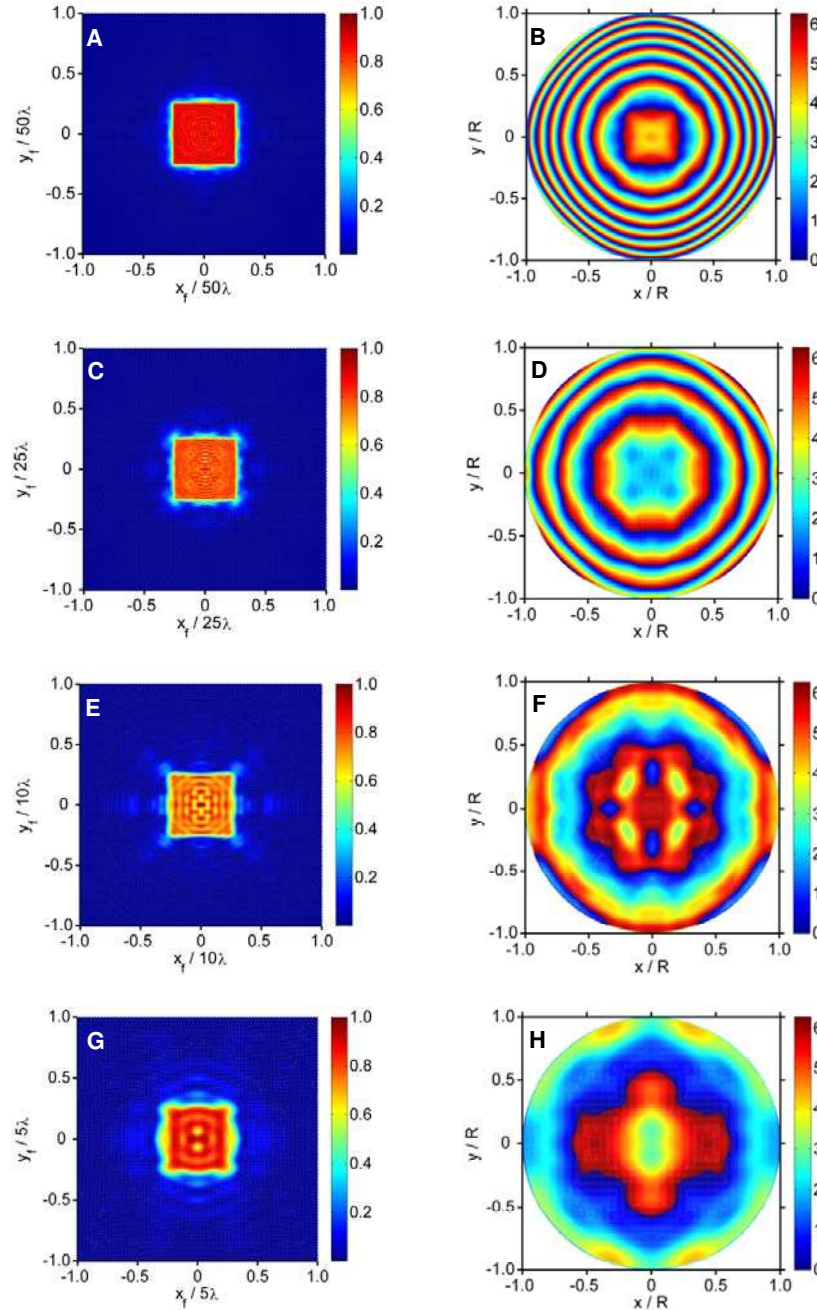
Figure 5.5 (*Left*) Normalized focused irradiance distributions and (*right*) corresponding DOE phase profiles obtained when the vector diffraction algorithm was used to reshape the input beam to a focused square flat-top irradiance profile for which (A, B) $D = 50\lambda$, (C, D) $D = 25\lambda$, (E, F) $D = 10\lambda$, (G, H) and $D = 5\lambda$. The DOE phase is plotted in units of radians.

depolarization effects that occur under high-*NA* focusing by relating the DOE complex transmittance function and the electric field in the focal plane using the vector diffraction integrals. The algorithm was applied in the design of a phase-only DOE that reshapes the focused irradiance distribution of a circularly apodized flat-top input beam into a uniform square profile when focused using a 1.4-*NA* objective lens. We observe that beam uniformity, diffraction efficiency, and edge sharpness all degrade as the size of the targeted flat-top beam is decreased. This suggests that beam shaping becomes increasingly difficult as the area of the targeted irradiance distribution approaches that of the diffraction limited spot size. There are many possibilities for extending the method reported here. A wider range of focused beam shapes could be considered by appropriately modifying the constraints used to define the targeted irradiance distribution. The search for solutions could be made more general by including other free parameters. For example, one could allow for freedom in the size of the homogenized area, so that the targeted size is optimized along with diffraction efficiency and uniformity. Three-dimensional beam shaping could be achieved by applying constraints in multiple planes. This work may also be useful for extending methods employed in the design of phase masks for high-resolution photo-lithography.

CHAPTER 6: THE EFFECT OF EXPERIMENTAL ERRORS ON THE DIFFRACTIVE OPTICAL ELEMENT PERFORMANCE: THEORETICAL STUDY

6.1 Introduction

In chapter 3, the method of generalized projection was used to design a phase pupil filter that superresolves the axial PSF by 29% while holding the side-lobe intensities at below 52% of the peak intensity in the non-paraxial regime. Although the filter's performance is theoretically satisfactory, it can be greatly compromised by imperfections introduced during experimental implementation. Such imperfections include fabrication errors, surface quality variation, and optical misalignment.

The performance parameters G and M are dictated by the structural characteristics of the DOE profile. The profile of a binary DOE is characterized by the structure of its zones. Zone structural parameters include zone width, height and side-wall steepness. Different steps of the fabrication process, such as photo-exposure and etching, can introduce errors to the mentioned structural parameters of the DOE zones. The effect of fabrication errors on the superresolution properties of a phase DOE has been studied based on analytical models developed in the scalar regime [97, 98]. In this chapter, we use the vectorial theory of diffraction to model the effect of structural errors on the performance of the DOE. Further, we study the variation in performance due to errors introduced by the surface roughness of the substrate and due to translating the DOE center off the optical axis. This defines the tolerance allowed in aligning the DOE with the optical system. In this chapter, we utilize the vector diffraction theory to study the

effect of the mentioned imperfection on the performance of the 11-zone DOE designed in chapter 3.

6.2 Theoretical simulation of the experimental errors

Photolithography, e-beam writing, and nano-imprint lithography are among the many fabrication techniques that can be used to fabricate a phase-only binary diffractive element. Irrespective of the fabrication method, a difference between the feature sizes of the fabricated pattern and the theoretical one inevitably results. This is due to the various errors introduced by different steps in the fabrication process, such as photo-exposure and etching. The accuracy in the obtained feature size can differ significantly between the various methods but so does the cost. Pattern feature size accuracy and fabrication cost are mutually exclusive. It is therefore necessary that an error tolerance study is conducted on the binary phase DOE, so that an appropriate fabrication technique is chosen. In what follows, we study the effect of different errors introduced by the fabrication process on the performance of the superresolving DOE, represented through G and M.

6.2.1 Error due to etching line width

Imprecision in the fabrication process can cause the width of the π-zones to differ from the theoretical target. This can have an adverse effect on the performance parameters. To understand how this effect degrades performance and thus obtain a tolerance for fabrication errors, we studied how G and M change as a function of zone width variation. The variable Δw is introduced to specify the difference between the final and intended position of the boundaries of the π-zones. The width of each π-zone is then

changed by $2\Delta w$. The sign of Δw can be positive or negative corresponding to wider and narrower zones, respectively. In this study it is assumed that Δw is equal for all zones, independent of their width.

Figure 6.1 demonstrates the variation in G and M as a function of Δw. G and M were computed using the diffraction integral presented in Sect. 2.2. It can be seen from Fig. 6.1 that for this particular DOE pattern the axial resolution increases (smaller G) with increasing zone width, but this occurs at the expense of more intense side-lobes (larger M). The change in Δw studied was limited to 2 μm, as this resolution is readily achieved using photolithographic fabrication methods. Within the 2 μm change in Δw the performance of the DOE is still acceptable as the variation in G and M is limited to a maximum of 1.42% and 11.16%, respectively.

Another fabrication error, which can be contributed to etching, is sloped side-walls of the binary zones. To simulate the effect of sloped side-walls on the performance parameters we tried an extreme condition. A binary zone of π phase shift is equivalent to a thickness, $d = \dfrac{\lambda_v}{2(n-1)}$, where λ_v is the wavelength in vacuum and n is the index of refraction of the DOE substrate. For $\lambda_v = 800$ nm and $n = 1.5$ (glass substrate), $d \sim 800$ nm. Using dry etching to create the π-zones profile can result in very steep side-walls. This is due to the anisotropic etching property of dry etching. For a zone height of 800
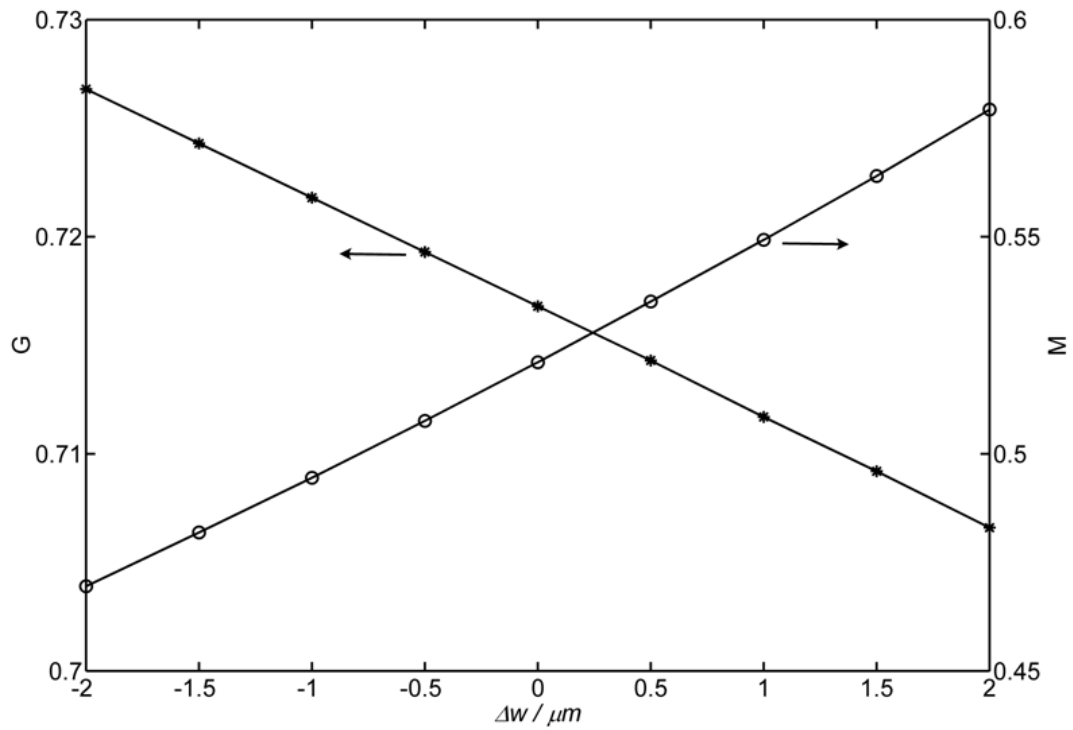
Figure 6.1. *G* (-*-) and *M* (-o-) as a function of variation in π-zone width, $2\varDelta w$.

nm, the side-walls do not slope over a region extending beyond 800 nm, as that would correspond to an inclination angle greater than 45° which is not realistic with dry etching. The sampling resolution across the DOE radius is limited to 0.5 µm, so the side-wall slope was simulated with a roll-off extending over a 1 µm distance. Under this extreme condition in side-wall slope, the performance parameters returned were *G* = 0.7156 and *M* = 0.5290. The side-wall slopes, therefore, have minimal effect on the DOE performance.

6.2.2 Error caused by variation of etching depth

Achieving a π phase difference between consecutive DOE zones requires that the substrate is etched to a depth of 882 nm when $\lambda_v = 800$ nm and $n = 1.45$ for the substrate material. A discrepancy in the etch depth can compromise the performance to the DOE and render the axial intensity distribution asymmetric with respect to the focal plane. The behavior of G and M as a function of the error Δd introduced into the etch depth is shown in Fig. 6.2. This simulation was computed using the diffraction integral of Eq. (2.6).
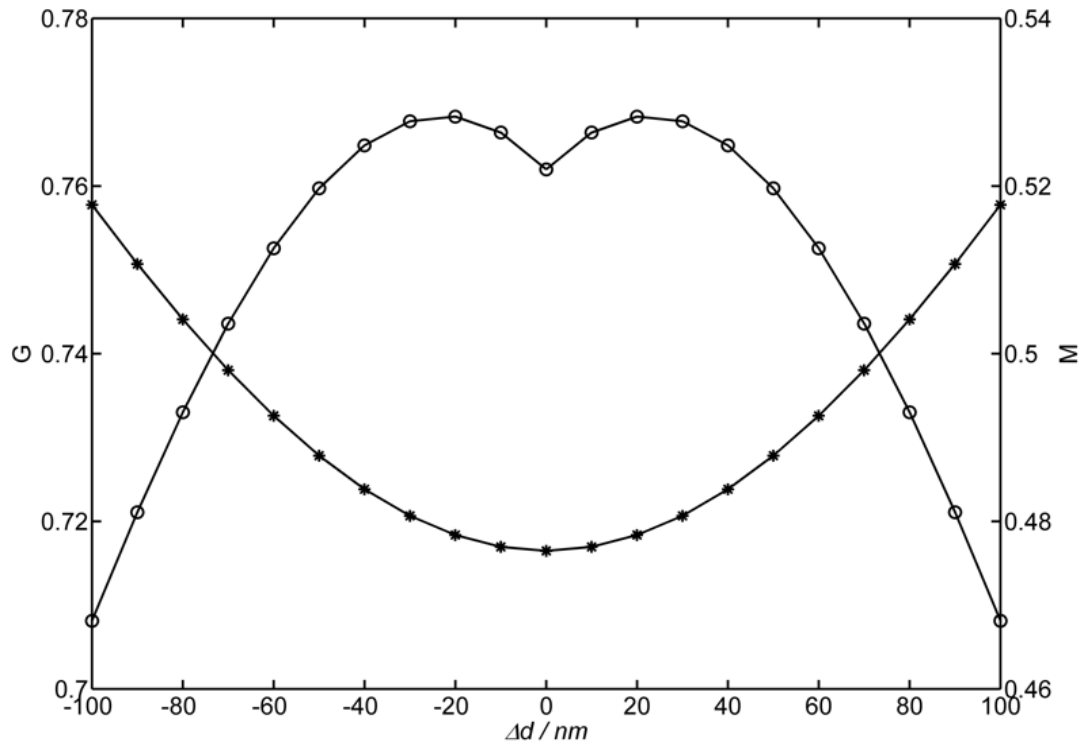
Figure 6.2. G (-*-) and M (-o-) as a function of error caused by a variation in the etch depth, Δd.

It can be seen from Fig. 6.2 that M increases first with reduced resolution in the error range of ± 20 nm and then drops beyond that point. It can be also observed that the

change in *G* and *M* is symmetric with respect to *Δd = 0*, when $|Δd| \leq 100$ nm. For a 100

nm etch depth error the change is the superresolution gain is significant, dropping from

29% down to approximately 24%. To maintain the DOE performance, it is therefore

required that the etching depth error be limited to ± 40 nm.
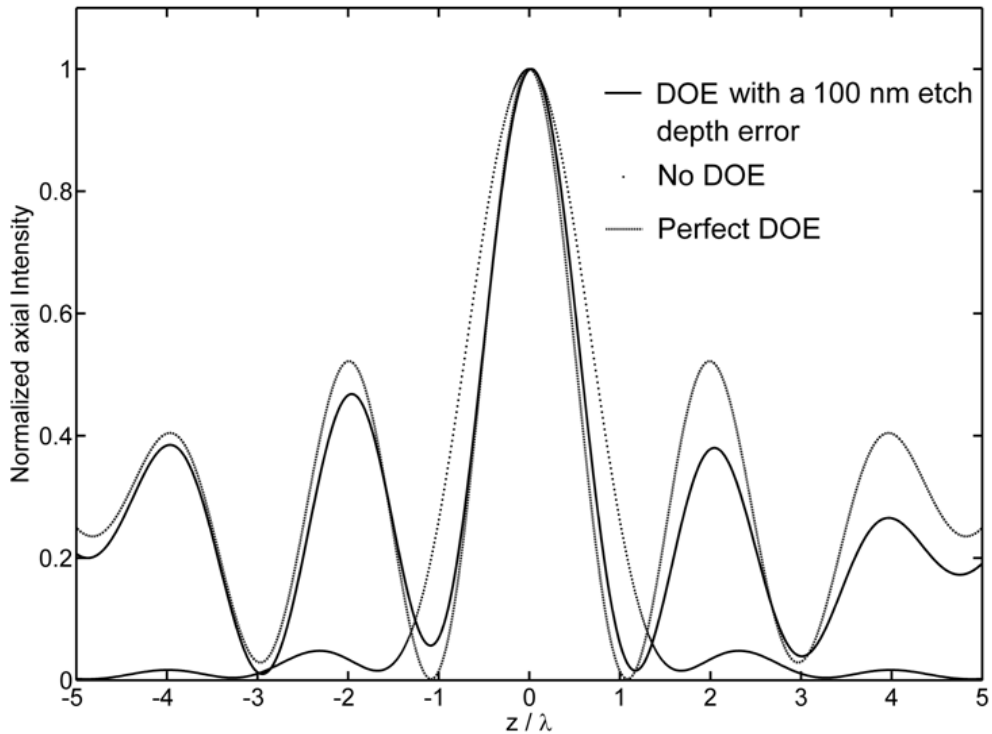
Figure 6.3. A 100 nm etching introduced error can result in an asymmetric axial intensity distribution (solid line). The error can also reduce the resolution gain as can be observed from comparing the intensity due to the DOE without errors (dashed line) and that due to the DOE with errors (solid line). The diffraction-limited axial intensity (dotted line) is provided for comparison reasons.

The variation in etch depth not only compromises the DOE performance but also

breaks the symmetry in the axial intensity distribution. Figure 6.3 illustrates how the

symmetry in the superresolved axial distribution for a 0/π binary DOE is broken by

introducing a 100 nm etch depth error to the DOE profile. Note the axial intensity distribution inverts with respect to the focal plane ($z = 0$) if an error of the opposite sign is introduced.

6.2.3 Error due to surface flatness

The surface flatness of the substrate into which the DOE profile is etched must also be considered during fabrication. Variations in surface height will introduce an additional phase difference between the DOE zones and thereby alter the focused PSF. The surface flatness of commercially available substrates is specified by the quantity λ/Q, which is the maximum variation in surface height across the element as a fraction of the optical wavelength. Typical values of Q are 2, 4, 5, 8, 10 and 20. The surface flatness is typically measured using an emission line from a mercury lamp at $\lambda_{Hg} = 546.1$ nm or the helium-neon laser line at $\lambda_{HeNe} = 632.8$ nm. The simulation described in this section was performed using the *HeNe* wavelength.

Variations in the surface topography could take many different forms. The roughness can have a linear, sinusoidal, zig-zag or random variation across the substrate dimensions. For the purposes of this work, we assume the worst case scenario in which variations in the surface height happen to result in a regular variation of $\pm\lambda_{HeNe}/Q$ between consecutive zones. The new phase of each zone after introducing a maximum height variation Δh is given as,

$$\varphi_N = \pi + (-1)^N \frac{2\pi(n-1)}{\lambda} \cdot \Delta h = \pi + (-1)^N \frac{2\pi(n-1)}{\lambda} \cdot \frac{\lambda_{HeNe}}{2Q}. \qquad (6.1)$$

N is the number of the π-zone, and n is the index of refraction of the substrate used.

Figure 6.4 provides an example showing the variation in the phase profile of DOE fabricated on substrate having a $\lambda_{HeNe}/4$ surface flatness and $n = 1.5$. As Q increases the phase variation between the different zones decreases which results in a phase profile closer to the ideal case.
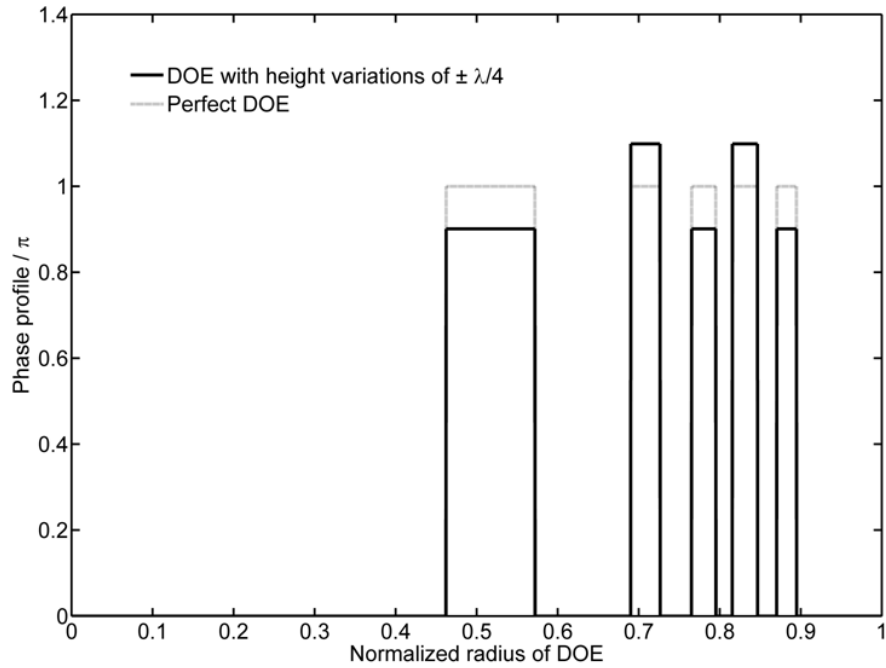
Figure 6.4. Phase profile of a diffractive optical element (DOE) fabricated on a substrate having a surface flatness of $\lambda/4$.

Figure 6.5 illustrates how the axial superresolution is affected by increased surface roughness (decreasing Q). As Q decreases, the superresolution becomes poorer (G increases) whereas the side-lobe intensity decreases. The variation in G is only minor for $Q > 8$, as a result, substrates having a surface flatness of $\lambda_{HeNe}/8$ or better should be more than adequate for fabrication superresolving DOEs.
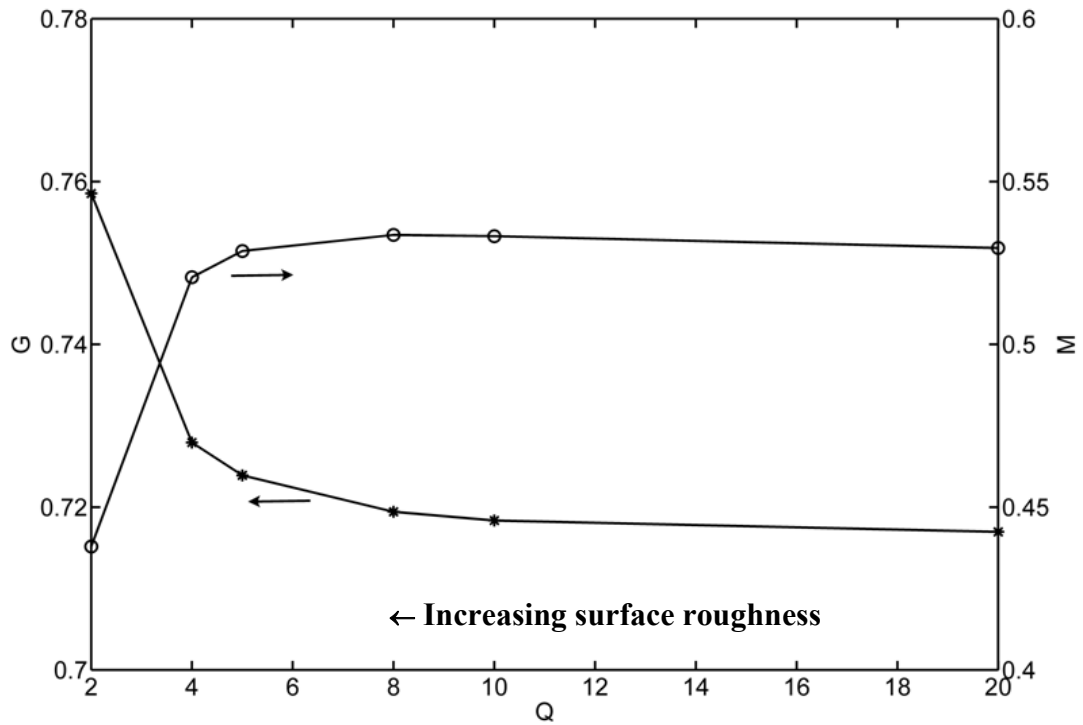
Figure 6.5. *G* (-*-) and *M* (-o-) as a function of increasing surface roughness (decreasing *Q*), modeled by assuming the roughness causes alternate DOE zones to vary in height systematically by $\pm\lambda/Q$.

6.2.4 Errors due to DOE misalignment

Installing the DOE in the optical system can be a challenging task. Any misalignment between the optical axis of the lens system and the mechanical axis of the DOE can compromise the performance of the DOE. It is therefore important to understand how *G* and *M* are affected by shifting the DOE center off the optical axis. Studying this effect provides a tolerance for aligning the DOE with the objective lens. Given the alignment margin, one can choose the appropriate mounting system for the DOE.

Translating a rotationally symmetric DOE off the optical axis will result in a non-rotationally symmetric system. As a result, the diffraction theory provided by Eq. (2.1) cannot be used to calculate the axial intensity distribution. To analyze the properties of the axial intensity distribution it becomes necessary to calculate the full 3D focal intensity distribution. This requires the use of the general vectorial diffraction integrals represented in Eq. (5.1). Once the focal intensity distribution is calculated, the axial intensity can be extracted to study how offsetting the DOE from the optical axis affects the performance parameters G and M.

Figure 6.6 shows the iso-intensity surface plot of the 3D PSF as viewed along the x-axis and generated by a DOE that is shifted of the optical axis by 51.36 μm along a diagonal direction laying at 45° with respect the x and y axes. The iso-surface represents the set of all points at half the peak intensity. As can be seen from the figure, the PSF is tilted with respect to the optical axis. As a consequence, G and M cannot be calculated using the intensity along the optical axis, but instead must be calculated along an axis that passes through the center of the PSF. Calculating the FWHM of the central-lobe in a tilted PSF along the optical axis does not provide the real FWHM of the central-lobe. This actually gives the projection of the real FWHM onto the optical axis. The FWHM of the PSF along the optical axis is narrower than the actual one and thus cannot be compared with the FWHM obtained from a DOE centered on the optical axis. Studying the variation of G and M along the newly defined axis is more relevant to applications where the peak intensity in any focal plane along the z-axis defines the DOE performance. Such is the case in direct laser writing applications.
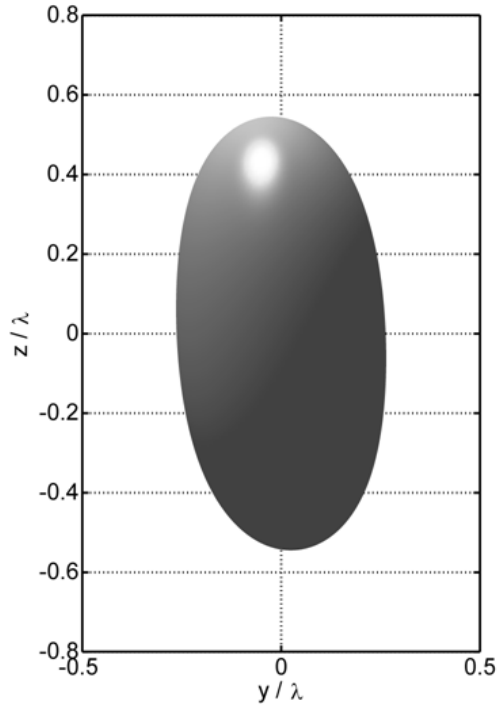
Figure 6.6. Iso-intensity surface of the point spread function (PSF) at half the peak intensity as generated by a DOE shifted of the optical axis by 51.36 μm and as viewed perpendicular to the *yz*-plane.

The 3D PSF and its performance were calculated as the DOE was translated in the aperture plane along the *x*-axis, the *y*-axis, and along a diagonal that bisects the *x*- and *y*-axes. The variable *ΔC* is defined as the distance between the center of the DOE and origin of the aperture plane. The values *G* and *M* as a function of *ΔC* are shown in Fig. 6.7(A), (B), and (C), corresponding to translation along the *x*-, *y*-, and diagonal axes, respectively. The aperture plane was sampled on 512 × 512 square grid over an area of
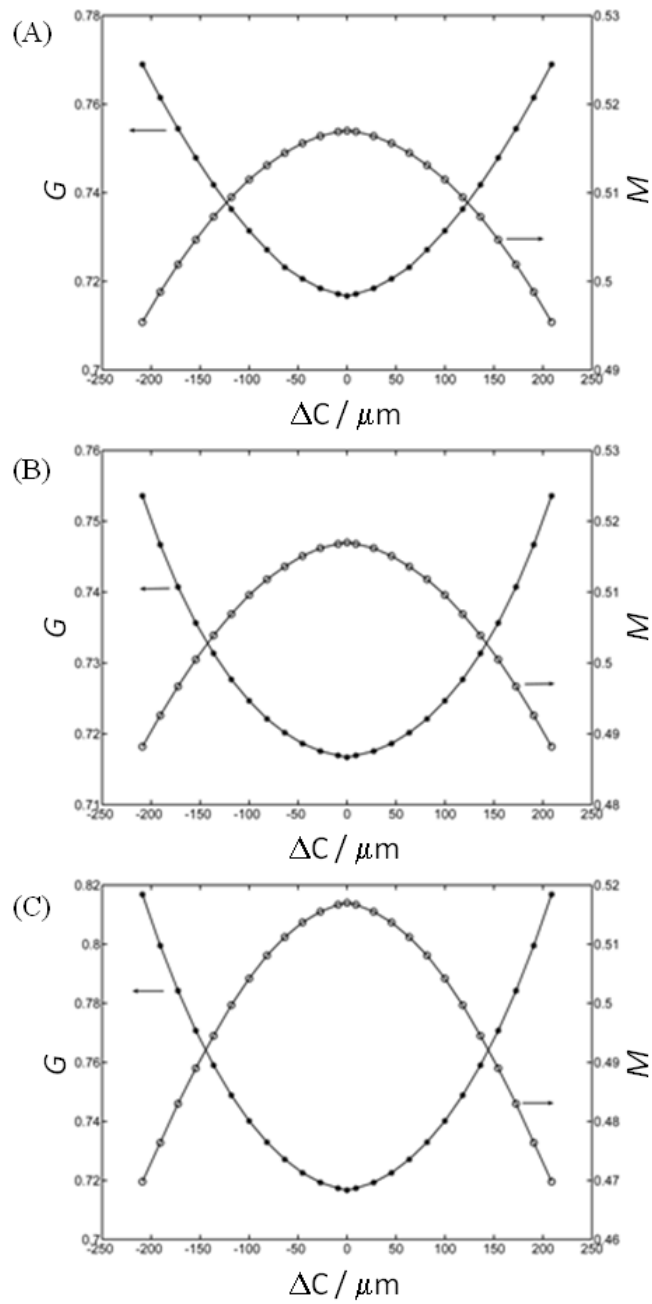
Figure 6.7. *G*(-*-) and *M*(-o-) as a function of DOE translation across the (A) *x*-axis, (B) *y*-axis, and (C) a diagonal that bisects the *x*- and *y*-axes.

9.3 mm × 9.3 mm, set by the diameter of the objective lens entrance pupil. This resulted in a translation step of 18.16 µm along the x- and y-axes, respectively, and 25.68 µm along the diagonal axis. It can be seen from Fig. 6.7 that any off axis translation of the DOE diminishes DOE performance. To optimize performance, the DOE center should fall within a disk surrounding the optical axis having a diameter less than 100 µm.

6.3 Conclusion

Numerical modeling was used to study the effect of fabrication errors and alignment tolerance of an axially superresolving diffractive optical element under high numerical aperture focusing. Fabrications errors studied are variations introduced by zone width etching, sloped side-walls, etching depth, surface flatness and centering the DOE with the optical system. The analyses of superresolution properties G and M as a function of experimental errors provides a mean to create error margins required to properly choose the fabrication technique.   The analysis also provides a theoretical basis for understanding any degradation in performance due to any experimental errors.

CHAPTER 7: EXPERIMENTAL CHARACTERIZATION OF DIFFRACTIVE OPTICAL ELEMENT PERFORMANCE

7.1 Introduction

This chapter provides a description of the experimental tasks required for incorporating a DOE into an optical system which utilizes a high NA objective lens. The goal is to experimentally measure and characterize the PSF obtained with an axially superresovling binary phase-only DOE discussed in previous chapters. This will be accomplished using the 11-zone DOE designed using the MGP algorithm. Although, the DOE designed using the PSO algorithm can yield higher axial superresolution, the former consists of fewer zones and as such it is easier to fabricate while still providing a significant and experimentally useful degree of axial superresolution.

The process of experimentally demonstrating axial superresolution involves three fundamental steps: (1) constructing the DOE, (2) integrating the diffractive element in the optical system and (3) measuring the PSF with and without the DOE. Although accomplishing these goals might seem straight forward, they are in fact challenging, primarily due to aberrations introduced by real optical components. Optical components can be manufactured only up to a certain limit of surface flatness and/or quality. Any manufacturing imperfections will introduce distortion to the phase wavefront or aberrations to an incident laser beam. Depending on the type of optical system, sometimes these aberrations are considered minimal and can be neglected. Unfortunately, experiments involving wide angle optics, such as high NA objective lenses, are exceptionally sensitive to the presence of aberrations. This necessitates a

rigorous characterization of the laser wavefront and the optical quality of the different components used in the system; in addition to, the employment of wavefront cleaning techniques. These additional characterization steps constitute an indispensable complement to the basic tasks required to demonstrate axial superresolution under high *NA* focusing.

7.2 DOE fabrication and characterization

Several fabrication techniques can be utilized for transferring the DOE pattern into a transparent photomask [52]. For binary profiles the method selected should be capable of generating accurate feature sizes, sharp side-walls and exact etch depth. Profile errors resulting from the fabrication process can compromise performance of the DOE (Chapter 6). The minimum zone width in the targeted 11-zone profile is on the order of 100 μm. Such a feature size can be easily produced with contact photolithography. Additionally, when combined with dry etching, photolithography can produce very sharp side-walls with accurately controlled etch depth. As such photolithography and dry etching provide an appropriate and cost effective means for fabricating the DOE.

The DOE fabrication can be divided into three main processes. First, an amplitude photomask having the pattern of the targeted DOE is created via electron beam lithography (EBL). Second, the photomask is used to transfer the binary DOE pattern using optical lithographical into a protective metal layer onto the surface of a fused silica substrate. Last, dry etching is used to bury the pattern into the substrate, thereby creating the DOE phase mask. Each of these steps is described in greater detail below.

The photomask was created on a commercial 4' x 4' quartz substrate that comes coated with a 120 nm layer of chrome topped with a 400 nm thick layer of poly(methylmethacrylate) photoresist (PMMA (Telic co.). A Leica EBPG 5000+ EBL system was used to write the targeted pattern in the PMMA layer. The writing was done at 50 kV accelerating voltage, a current of 100 nA, and an electron fluence of 460 $\mu C/cm^2$. The maximum resolution achieved with these settings is approximately 100 nm. The mask was developed by immersing in methyl isobutyl ketone (MIBK):isopropanol (IPA) (1:3) for 90 seconds to remove the exposed part of the PMMA layer and cleaned by rinsing with IPA for 15 seconds. The exposed chromium was removed etched from the glass surface using a chromium etcher (Air Products, Material no. 64216). The photomask substrate was then rinsed with deionized water and immersed in dichloromethane to strip away remaining photoresist. An acetone/methanol/IPA rinse was used as a final cleaning step to remove any residuals, leaving behind a chromium layer with the targeted pattern inscribed.

The amplitude photomask was then used to transfer the DOE pattern into a fused silica substrate (Dell Optics Inc.) by optical lithography, thereby creating a replica phase mask. The substrate was 12.7 mm in diameter and 3 mm thick with $\lambda/10$ surface flatness. The DOE fabrication procedure is illustrated in Fig. 7.1. The substrate was first sonicated in 1 M KOH(aq) for 30 minutes to form Si-OH surface groups that improve adhesion of photoresist. A 1-μm thick layer of negative-tone photoresist (NR7-1000PY, futurex Inc.) was spin coated onto the substrate and soft-baked for 1 minute at 150 ºC. The DOE profile pattern was then irradiated into the photoresist layer via UV photo-exposure (Karl Suss UV aligner, 12.5 mW cm^{-2}, $\lambda = 365$ nm) for 25 seconds. Following
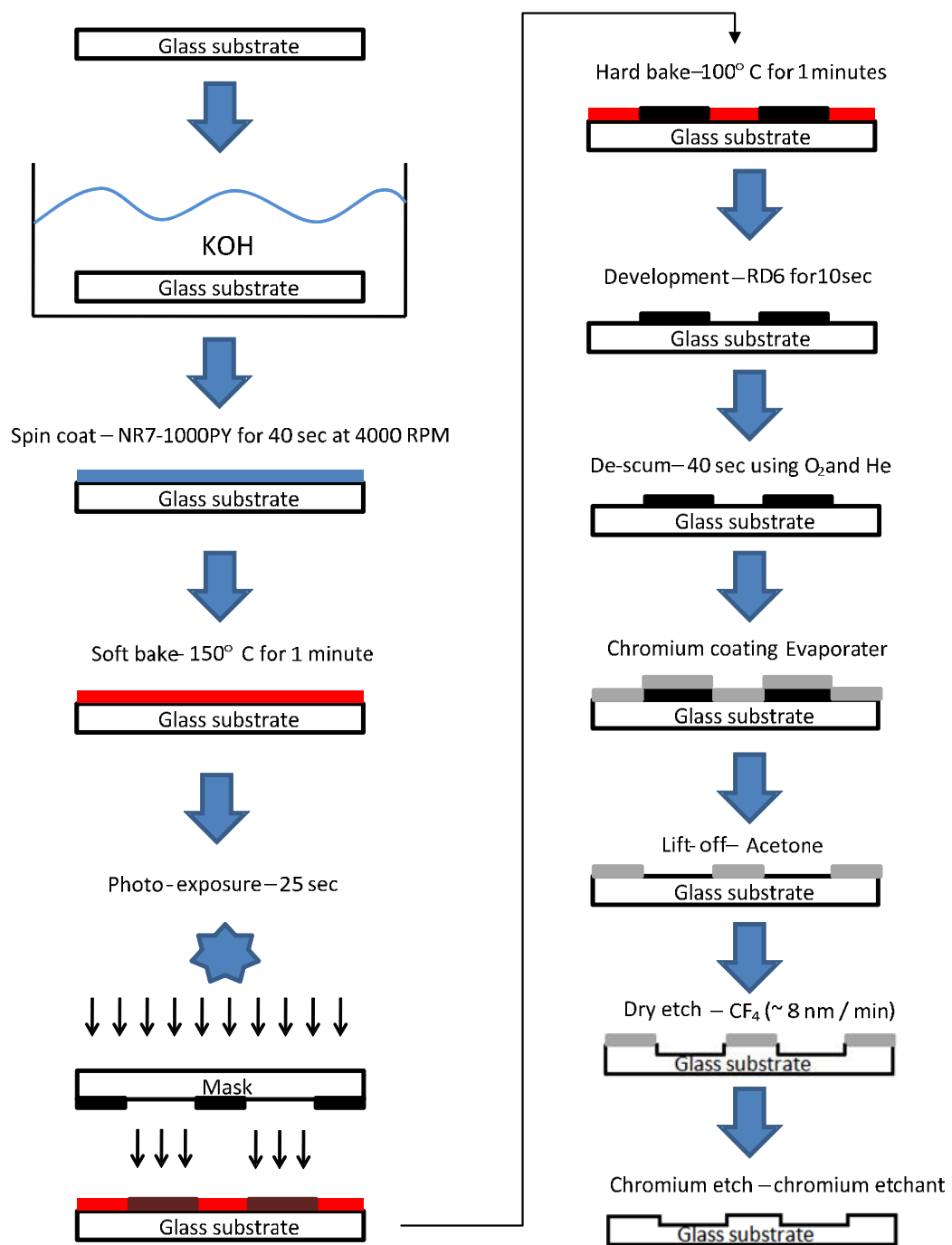
Figure 7.1. Fabrication steps for creating a binary DOE.

photoexposure the substrate was baked for 1 min at 100 ºC and developed with the resist developer (RD6, Futurrex Inc.) for 10 seconds to obtain the DOE pattern in the resist layer. Dry etching was then used to bury the DOE pattern into the fused silica substrate. To accomplish this, the unexposed area of the pattern was cleaned through a "de-scum" step that removes residual photoresist and then coated with metal through thermal evaporation deposition (Edwards, FL 400, Auto 360). The evaporator chamber was evacuated to 10^{-5} torr. For chromium the electric current used to melt the metal is set to 3.4 A. The metal acts as a protective layer that prevents etchant from attacking the area surrounding the π-zones. To clear the zones area for etching, a photoresist lift-off was performed by immersing the substrate in acetone. A plasma etcher (Plasma-Therm, 790 series) using CF_4 gas was then used to dry etch the substrate at a rate of 8 nm/min. Given that the substrate refractive index is $n_s = 1.45332$, an etch depth of 882.38 nm was targeted to create a π-phase shift at a vacuum wavelength $\lambda = 800$ nm. The substrate was then immersed in a chrome etcher (Material no. 64216, Air Products) until all metal was removed from the substrate surface leaving behind the targeted phase-only binary DOE.

The DOE profile was characterized using multiple imaging techniques because the feature sizes vary over several length scales. The height and sharpness of the zone side-walls vary on the micron scale while the zones width can be several millimeters. As such, an optical microscope was used to measure the width of the zones while a scanning electron microscope (SEM) and a profilometer were used to characterize the zone side-wall sharpness and height, respectively. A digital image of the fabricated DOE is shown in Fig. 7.2. The light scattered from some of the circular zones can be seen emanating from the center of the half-inch diameter fused silica substrate. To take a closer look at

the zones and measure their width, Z_w, optical microscope images were collected using 10X and 40X objective lenses. The choice of magnification depended on the width of each zone and the field of view (FOV) of the objective. Although a 40X lens can provide higher imaging resolution compared to the 10X objective, its FOV does not allow capturing the wider zones in a single image without resorting to image stitching. To avoid introducing stitching errors into the zone width measurement, resolution was sacrificed for a wider FOV by using the 10X objective. The collected images of different zones were analyzed to obtain the actual width of each zone. Figure 7.3 illustrates the procedure used to analyze the width of a zone. The image of zones 5, 6 and 7 as collected by the 40X objective lens is shown in Fig. 7.3(A). A line profile data across zone 6, indicated by the yellow stripe in Fig. 7.3(A), is imported into MATLAB for analysis. A plot of the imported data is shown in Fig. 7.3(B). The boundaries of the zone are clearly indicated by the two dips in the profile intensity. The dips result due to the light scattering experienced at the edges of each zone. This measurement was repeated at several positions across the zone. The widths of the inner nine zones are summarized in Table 7.1, along with the minimum zone width measured ($Z_w^{\min}$), maximum zone width measured ($Z_w^{\max}$), the average zone width (Z_w^{avg}), the targeted zone width (Z_w^{theory}) and the difference (Z_w^{Δ}) between Z_w^{theory} and Z_w^{avg}. It can be clearly seen that there is a significant discrepancy between the experimental zone width and the theoretical ones. All the zones share a systematic increase or decrease in width by approximately 7 µm with respect to the targeted value. This is mainly due to errors introduced by the exposure and development steps in the photolithography fabrication procedure.
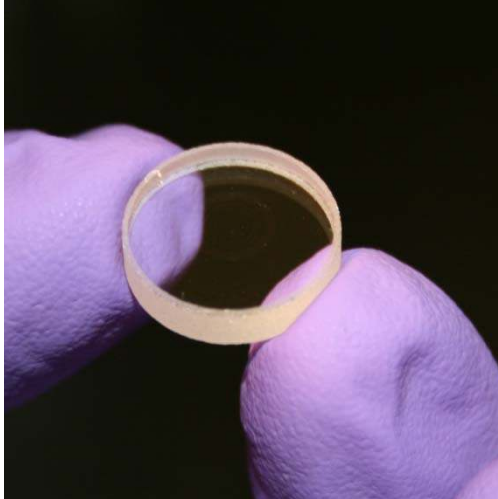
Figure 7.2. Photograph of the DOE zones fabricated on a half-inch fused silica substrate.
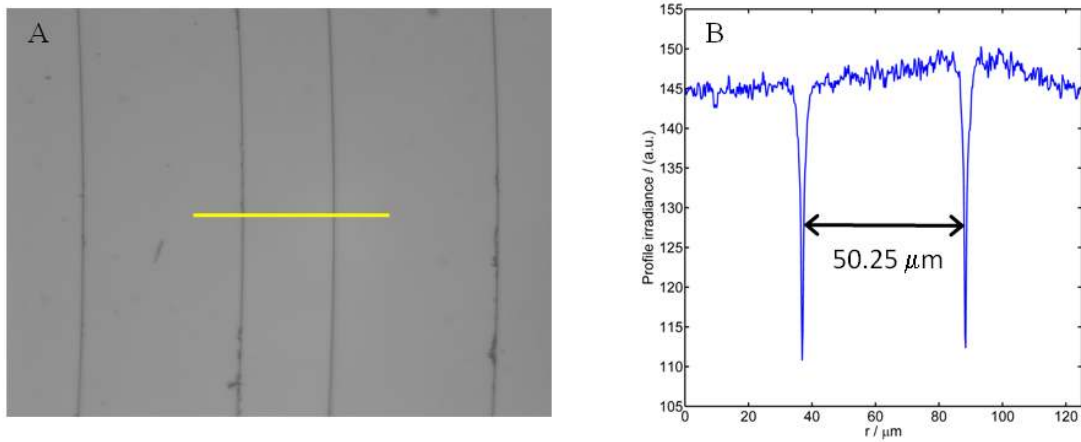
Figure 7.3. DOE zone width measurement procedure. (A) An optical microscope image of zones 5, 6 and 7 as collected by a 40X, 0.6 *NA* objective lens. (B) A line profile across zone 6 corresponding to the yellow stripe in (A).

Additional trial and error characterization of the zones' width versus exposure and development time was conducted. It was found out that the error in the zone width can be reduced to less 2 μm with an exposure time of 15 seconds proper exposure and a 10 seconds developing time.

Table 7.1. The width of the inner 9 zones measured using optical microscope images.

Zone number	Z_w^{min} (μm)	Z_w^{max} (μm)	Z_w^{avg} (μm)	Z_w^{theory} (μm)	$Z_w^{theory} - Z_w^{avg}$ (μm)
1	-	-	-	1294.55	-
2	312.95	316.29	314.70	306.61	- 8.1
3	324.02	327.50	326.19	331.53	5.35
4	108.28	108.73	108.46	101.11	- 7.34
5	102.16	102.60	102.44	108.77	6.32
6	91.49	92.43	91.94	84.15	- 7.78
7	49.96	50.46	50.25	57.15	6.79
8	93.27	93.70	93.37	86.30	- 7.07
9	60.55	61.51	60.97	67.22	6.25
10	73.82	74.30	74.13	67.80	- 6.33
11	-	-	-	294.81	-

The depth of the etched zones was measured using an Alpha-Step 200 profilometer (Tencor Instruments). The average depth measured was 878 ± 5 nm.

SEM images were used to characterize the slope of the zone sidewalls. Figure 7.4 shows a front (A) and top (B) view of an arbitrary zone edge. The slope of the sidewall separating two consecutive zones can be seen to drop-off at an approximate rate of 3.5 (~ 778 nm/250 nm). Theoretical simulations in Sect. 6.2.1 show that this rate of sidewall

drop-off has no effect on *G* or *M*. Further, the SEM images indicate a rather smooth rim to the zone boundaries.
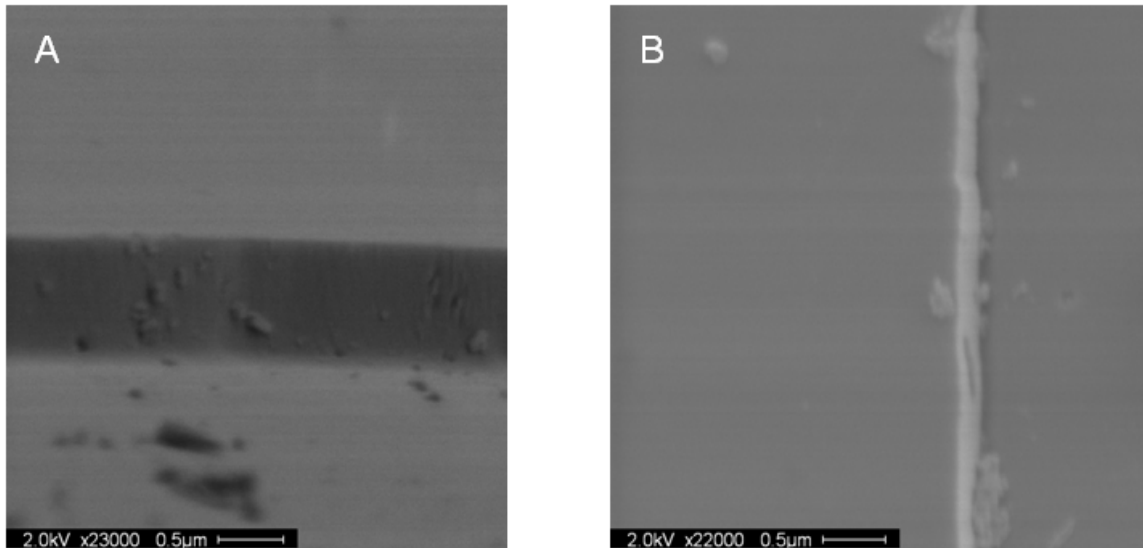
Figure 7.4. SEM images of the a zone sidewall. (A) Front view, (B) Top view.

7.3 Characterizing the point spread function

The most common technique to characterize the performance of an optical lens and/or a diffractive element is to measure the point spread function of the system. Several methods have been developed for examining the optical field near the focus of a high-*NA* objective lens. These include the use of a tapered fiber probe [99], scattering from small nanoparticles [100], optical fiber interferometry [101, 102] and focal point imaging [103]. The later technique is comparatively easy to implement and is appropriate for probing the axial intensity distribution under tight focusing. As such it will be utilized for demonstrating axial superresolution achieved with the DOE.

96

7.3.1 Focal point imaging

Characterizing the PSF by focal point imaging is based on a method used to measure the depth response of a confocal microscope [104]. As illustrated by the optical setup shown in Fig. 7.5, a flat mirror placed in the focal volume reflects the focused laser beam back through the objective lens. A beam splitter then directs the reflected light through a tube lens which images it onto a charge coupled device (CCD). The mirror is translated along the optic axis through the focal volume and a sequence of focal plane images are collected to obtain a three-dimensional intensity map of the PSF. Given a function Φ that describes the aberration of the objective lens, the focal field of the tube lens E can be calculated using the vector diffraction integral as

$$E(x,y,z) = \frac{i}{2\pi} \iint_{k_x,k_y} \frac{E_R(k_x,k_y,k_z)}{k_z} e^{i(\Phi(k_x,k_y) + \Phi(-k_x,-k_y))} e^{i(k_x x + k_y y + k_z(z-2z_d))} dk_x dk_y , \quad (7.1)$$

where E_R denotes the reflected field at the exit pupil of the tube lens, (k_x, k_y, k_z) are the components of the wavevector and z_d represents the displacement of the mirror from the geometric focus. The integration is taken over the spatial frequencies dictated by the *NA* of the lens. Equation (7.1) is identical to that in Eq. (5.1) except that the aberration function appears as a summation of two functions: $\Phi(k_x, k_y) + \Phi(-k_x, -k_y)$. $\Phi(k_x, k_y)$ represents the aberration encountered upon passing forward through the objective and $\Phi(-k_x, -k_y)$ accounts for additional aberration accumulated upon the return pass following reflection by the mirror. If the objective lens is aberration free, then $\Phi(k_x, k_y) = 0$ and so the actual PSF and its image are identical. In the presence of aberrations the collected image does not depict the PSF exactly; effectively the imaged PSF depicts an aberrated

version of the aberrated PSF itself. Although the two are not strictly identical, the imaged PSF still reveals the effect of aberrations on the PSF.

In Sect. 7.6 a method is described for measuring the aberration function $\Phi(k_x, k_y)$ of the objective lens. The validity of this measurement can be checked through PSF imaging. The measured aberration function can be substituted in Eq. (7.1) to calculate a theoretical imaged PSF. A comparison between the imaged PSF and its theoretical equivalence provides a tool to assess the accuracy of the aberration measurement.

Wilson *et al.* [105] show that true confocal imaging is achieved only if the detector diameter d is limited by

$$ d \leq \frac{n \, M \, \lambda}{2\pi \, NA}, \tag{7.2} $$

where M is the system magnification and n is the refractive index of the object plane. The CCD pixels serve as an array of pinholes, and thus the pixel size should be less than or equal to d.

7.3.2 Imaging the diffraction-limited PSF

A CCD camera having 6.45 μm × 6.45 μm pixels (Roper Scientific CoolSnap*ES*, 1392 × 1040 pixels) was used to image the PSF formed by focusing plane waves at $\lambda =$ 800 nm through a 100X, 1.4-*NA* oil immersion objective lens ($n = 1.516$). Substituting these values into Eq. (7.2) gives $d \leq 13.78$ μm, which shows that Wilson's confocal aperture limit is easily satisfied by the CCD used here. Figure 7.6 shows that imaged
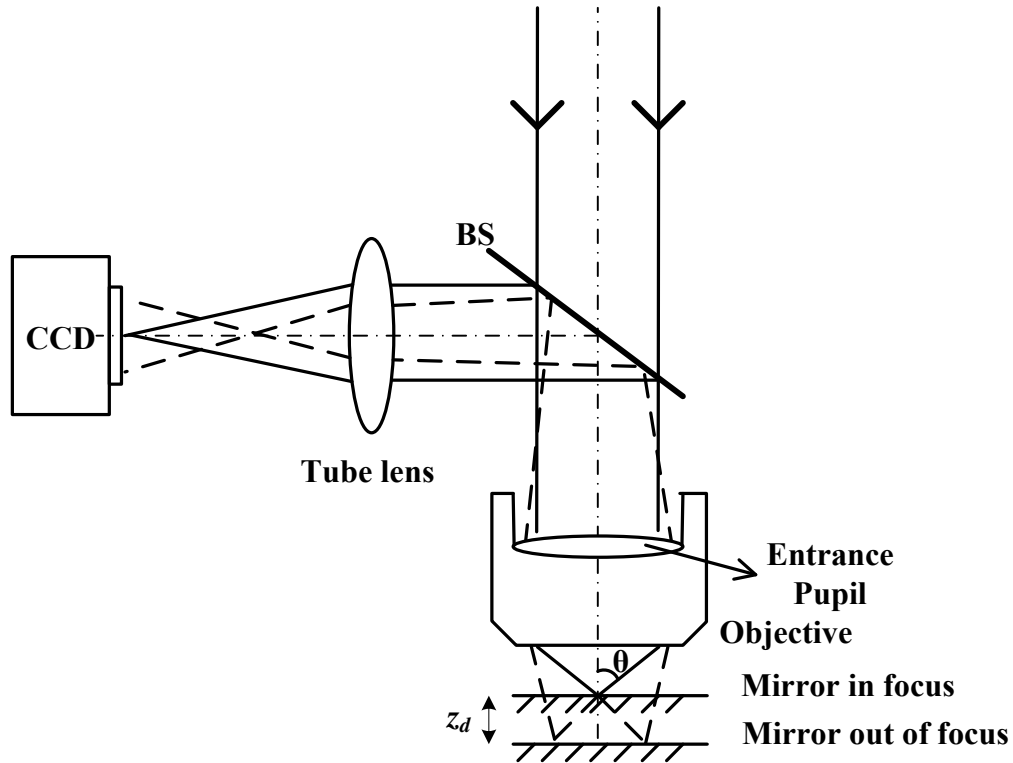
Figure 7.5. The CCD imaging system used to map the point-spread-function of a high *NA* lens.

axial PSF and the theoretical diffraction limited distribution. There is a clear discrepancy between the two profiles. In comparison to the theoretical PSF, the FWHM of the main peak in the experimental plot is twice as wide and the secondary side-lobes are much larger. Further, the experimental PSF is not symmetric with respect to the geometric focal plane ($z = 0$). This asymmetry is diagnostic of spherical aberration in the optical system [106]. The quality of the laser beam entering the objective lens was characterized using a Hartmann sensor ($\lambda/100$ minimum error, 127 μm lateral resolution). The incident beam was found to have a wavefront error less than $\lambda/10$. This implies that the aberrations observed originate in the objective lens itself.
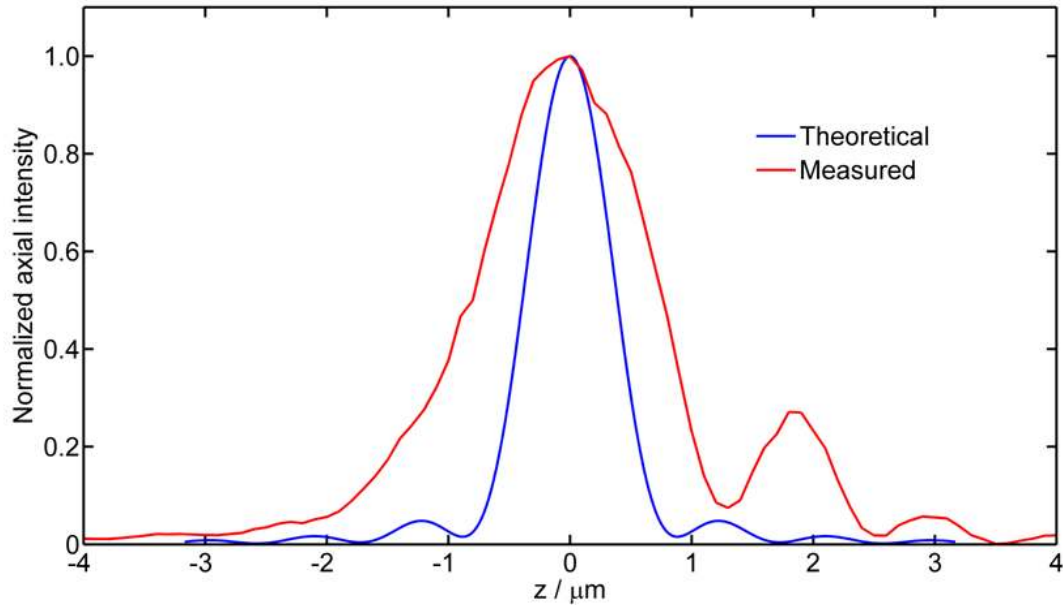
Figure 7.6. Theoretical (blue) and measured (red) axial intensity distributions of 1.4-*NA* immersion oil objective lens at λ = 800 nm.

A study of the irradiance distribution at different focal planes along the optical axis indicates the presence of a different type of aberration in the objective lens. Figure 7.7 depicts the intensity profile in two focal planes at $z = -0.8$ μm (A) and $z = +0.8$ μm (B) with respect to the geometric focus. The elliptically shaped PSF pattern rotates by 90° as the beam propagates through the focal point. This is a clear signature of astigmatism [107]. The fringes surrounding the main lobe in Fig. 7.7(B) confirms the presence of spherical aberrations already verified from the axial intensity profile. The magnitude and type of these aberrations is quantified in Sect. 7.6.

7.3.3 Spherical aberration compensation

High-*NA* objective lenses are very sensitive to spherical aberration. Any variations in the parameters for which the objective was optimized can compromise its PSF. An

objective is "aberration free" only if it is used under conditions for which it was explicitly designed, including immersion oil index, cover slip thickness, and wavelength. Some objective lenses are equipped with a correction collar that can be adjusted to compensate for spherical aberrations. Researchers have also used external variable wavefront modulators to introduce aberrations into the incident beam that offset aberrations inherit

$z = -0.8$ μm $z = 0.8$ μm
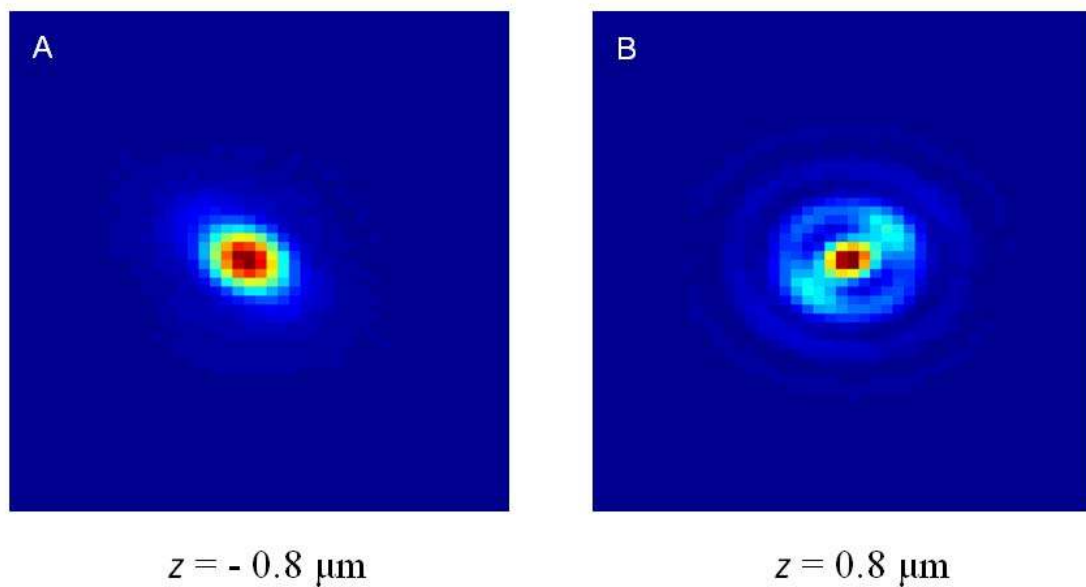
Figure 7.7. Imaged transverse PSF in two planes perpendicular to the optical axis. The two planes are located 0.8 μm (A) before and (B) after the focal plane ($z = 0$). The change in the intensity distribution indicates the presence of astigmatism in the objective lens

to the objective [108]. Such devices are costly and complicate the optical setup. Alternatively, one can compensate for spherical aberration by adjusting the refractive index of the objective immersion liquid. This approach was used to study how changing n affected spherical aberration in and the PSF produced by the 100X/1.4-*NA* objective lens used in this work. The axial distribution of the imaged PSF, the FWHM of its

central lobe, and the symmetry of the distribution (S) with respect to focal plane, were used to gauge the effectiveness of the correction.

The refractive index of the immersion medium was adjusted by blending the immersion oil (IO, type DF, Cargille Laboratories, code 1261) provided by the objective manufacturer with a second immersion liquid (IL, Cargille Laboratories, code 1160) having lower refractive index. The Cauchy formula for the refractive index of each liquid at 23°C is provided by the manufacturer (Cargille Laboratories) as :

$$n_{IO} = 1.497540 + \frac{602657.1}{\lambda^2} + \frac{1.130597 \cdot 10^{11}}{\lambda^4} \qquad (7.3)$$

and

$$n_{IL} = 1.4734957 + \frac{525569}{\lambda^2} + \frac{1.653246 \cdot 10^{12}}{\lambda^4}. \qquad (7.4)$$

Here, the wavelength λ is specified in Angstroms, n_{IO} and n_{IL} indicate the index of refraction of the IO and IL, respectively. Assuming the two liquids are miscible, the refractive index of a mixture can be estimated as

$$n = \frac{n_{IO}V_{IO} + n_{IL}V_{IL}}{V_{IO} + V_{IL}}, \qquad (7.5)$$

where V_{IO} and V_{IL} are the volumes of the IO and IL, respectively. Because these oils are highly viscous, it is difficult to dispense and measure their volumes precisely. Accurate values for volume were obtained by measuring the mass of a targeted volume at ±100 µg

and dividing by the density. Mixtures were prepared in which the volume fraction of IL was increased in steps of approximately 7% by volume. Equations (7.2) – (7.4) were used to calculate n for each mixture. All measurements and calculations described in this section were performed at λ = 8000 Å. An error propagation analysis showed that the uncertainty in calculated n is ±0.0015 (0.1%) and is primarily due to uncertainty in the value of density provided by the manufacturer.

Figure 7.8 illustrates the axial intensity distribution obtained for each mixture. The percent-volume of IL and refractive index calculated for each mixture appears above each plot. It can be clearly observed that a change as small as 0.002 in the refractive index is enough to alter the axial intensity profile significantly. The characterization parameters S and FWHM of the axial profiles are given as a function of refraction index in Fig. 7.9. The symmetry parameter was calculated as follows:

$$S = \frac{1}{N}\sum\left|I_R - I_L\right|^2 , \tag{7.6}$$

where I_R is the part of the axial intensity that lies to the right of $z = 0$ or positive z and I_L is the left-hand component of the distribution. One of the two intensity profiles has to be folded with respect to the $z = 0$ plane before the subtraction in Eq. (7.6) is performed. N is the number of points in I_R or I_L. The range in z over which the subtraction was performed was limited to [-2 μm, 2 μm] as the noise in the low intensity wings of axial profile can introduce a significant error to the calculated values of S.
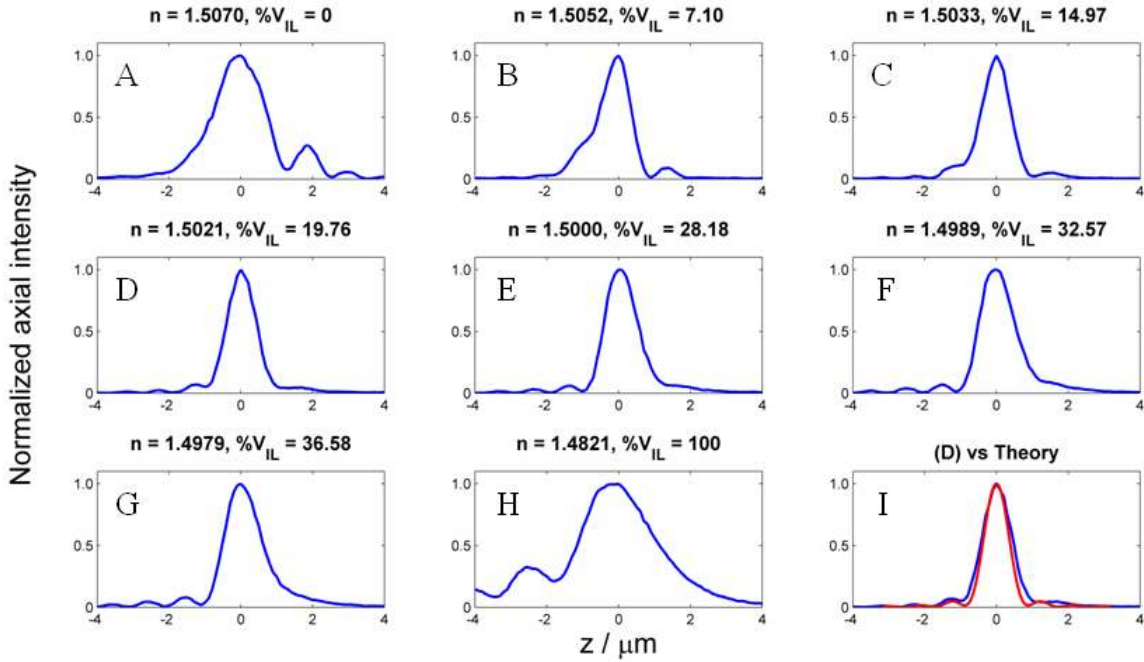
Figure 7.8. Axial intensity distribution measured as a function of the refractive index of the immersion oil mixture.

The values of S, FWHM, and axial intensity distribution all suggest that there is an optimum value for n at 800 nm. The mixture having $n = 1.5021$ yielded the most symmetric profile ($S = 0.012$) and a FWHM $= 1.00$ μm. The corresponding profile is shown in Fig 7.8(D) and compared to the theoretical diffraction limited profile (red) in Fig. 7.8(I). The highest resolution, FWHM $= 0.90$ μm, however, was obtained at $n = 1.5033$, where some symmetry ($S = 0.015$) was lost as illustrated by the axial profile in Fig. 7.8(C). These data suggest that spherical aberration can be minimized with an immersion mixture having refractive index between 1.5021 and 1.5033. Although the performance achieved with an optimized mixture is still below that expected for a diffraction limited focus (FWHM $= 0.79$ μm, $S = 0$), significant improvement in resolution and symmetry were achieved with this approach.
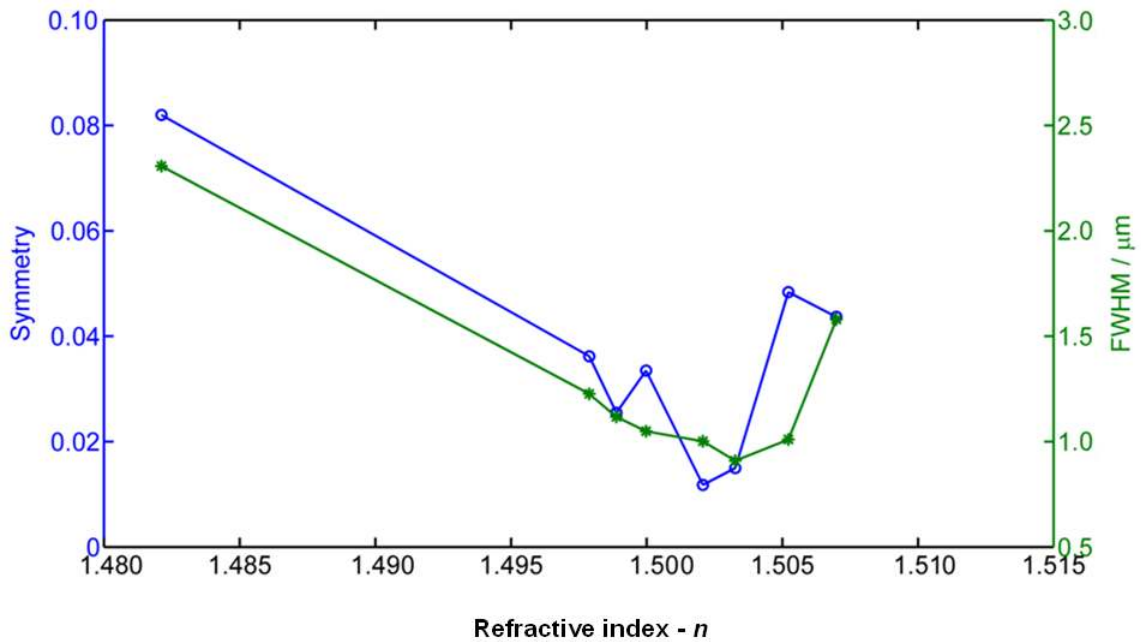
Figure 7.9. Symmetry (−o) and FWHM (−*) of the measured axial intensity distribution as a function of the immersion medium refractive index (*n*).

It is important to note that the calculation of the diffraction limited FWHM was done with the assumption that $NA = 1.4$ at $\lambda = 800$ nm. The *NA* of objectives in practice can be different than that specified by the manufacturer [109]. Furthermore, changing the refractive index of the immersion medium can only correct for spherical aberrations. It has no effect on astigmatism, which is also known to affect the axial intensity profile and resolution [110].

7.4 Incorporating the DOE into the optical setup

The experimental setup for integrating the DOE in the optical system is shown in Fig. 7.10. A mode-locked Ti:sapphire laser operating at $\lambda = 800$ nm is expanded using an

8X Keplerian beam expander. To ensure an aberration free optical wavefront at the output of the telescope, a sub-diffraction 5-μm diameter pinhole is placed at the focal plane of the first lens. The purely spherical phase front exiting the pinhole is collimated into a planewave by an achromatic doublet ($f = 400$ mm, $R_1 = 208.55$ mm, $R_2 = -208$ mm, and $R_3 = -859$ mm) that was optimized to reduce spherical aberration and coma in the near infrared. The expanded beam planar wavefront then passes through the fabricated phase DOE and where it is modulated into a binary profile. The field distribution at the surface of the DOE is then relayed onto the entrance pupil of the objective lens using a 4f system. The beam splitter (BS) and mirror at the focal plane of the objective are part of the PSF measurement system discussed previously and are used to image the PSF generated with the DOE. The pupil plane lens is needed to relay the DOE-modified field onto the pupil plane of the objective lens. The pupil plane relay lens was utilized in combination with the tube lens to image the field at the entrance pupil.

The objective lens is a complicated multi-element optical system, but its intended function is simple. It maps the field distribution at its entrance pupil onto an ideally perfect spherical wavefront converging towards the geometric focus. The DOE is used to modify the input field distribution, so to function properly, the DOE must be positioned immediately before the entrance pupil. For some high-NA objective lenses, such as that used here, the entrance pupil of is located inside the cylindrical metal housing, making it physically inaccessible. To overcome this problem, the 4f relay lens system was used to remotely image a copy of the field at the DOE surface onto the pupil plane.
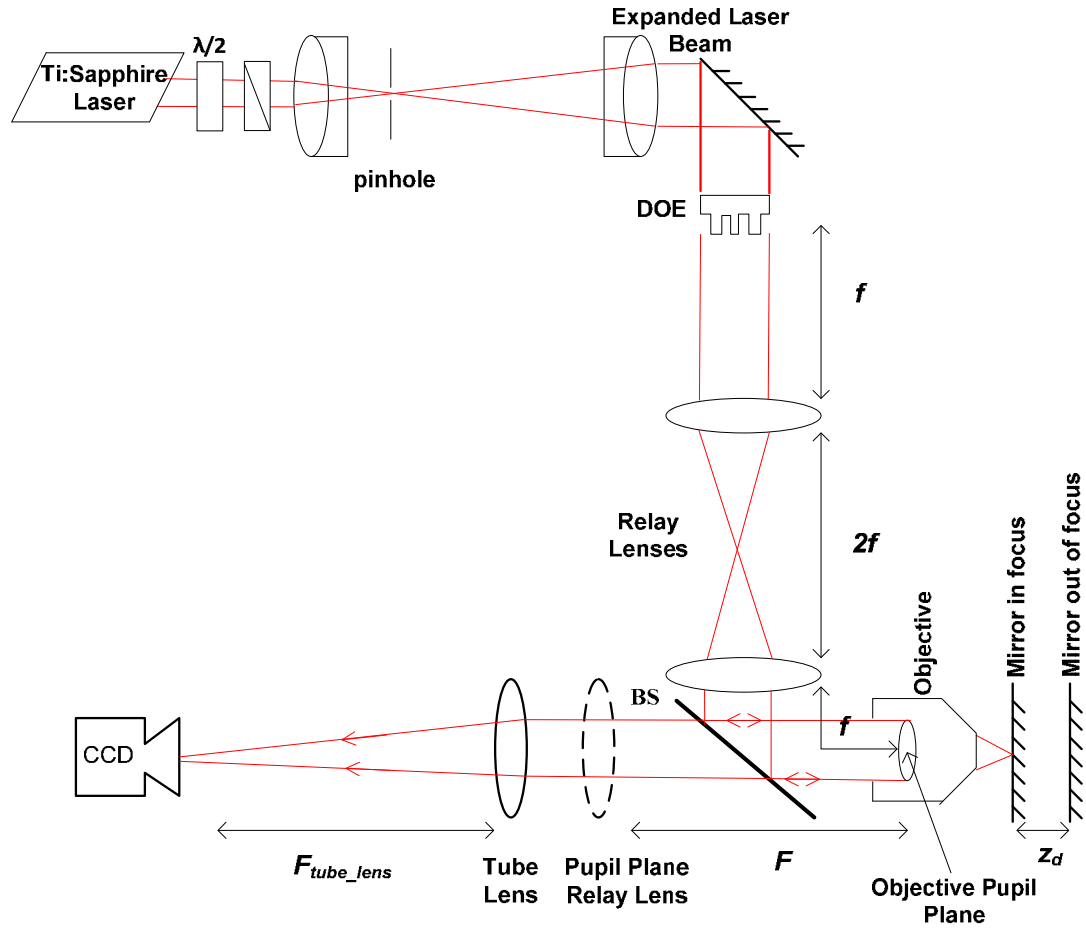
Figure 7.10. Experimental setup used to integrate the DOE into the focusing optical system and image the superresolved axially intensity distribution.

Given that the entrance pupil is located within the objective lens housing, locating it for precise DOE image relaying becomes challenging. This can be accomplished by imaging conjugate planes of the microscope system around which the experimental setup was built. In a properly aligned and focused optical microscope, there exist two sets of conjugate planes in the optical train: four field planes and four aperture planes. Because those planes are parfocal, they can be imaged superimposed on one another. The entrance pupil, also referred to as the back aperture, is one of the four aperture conjugate

planes. This necessitates the imaging of the aperture planes to locate the entrance pupil. The four aperture planes include: (1) the light source filament, (2) the condenser aperture, (3) the entrance pupil and (4) the CCD camera.

To image the aperture planes, a pupil plane relay lens was inserted between the objective and the tube lens (see Fig. 7.10). The tube lens/pupil plane relay lens combination images the entrance pupil field distribution onto the CCD camera. The position of the pupil plane relay lens was scanned until all four conjugate aperture planes came into sharp focus. Figure 7.11(A) shows the image of the source filament imaged using the CCD. By closing the iris of the condenser, both the filament plane and the iris are simultaneously imaged on the third conjugate plane, which is the camera (Fig. 7.11(B)). This implies that the pupil plane relay lens is accurately positioned to image the entrance pupil intensity distribution on the CCD.
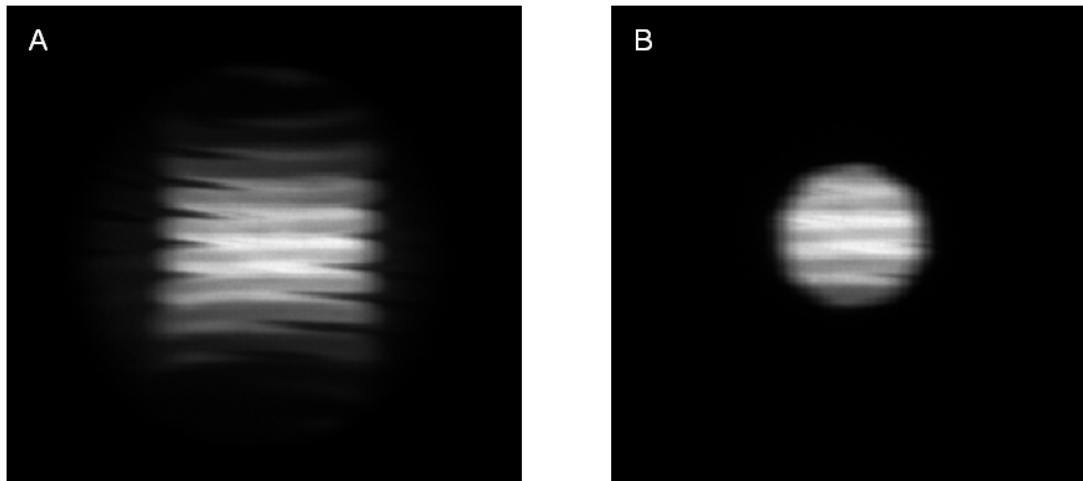
Figure 7.11. CCD image of the aperture conjugate planes. (A) Light source filament. (B) Filament and condenser iris superimposed.

Novotny *et al*. [111] demonstrated that the field distribution at the back aperture reflected by a mirror positioned at the focal plane of the objective is identical to the incident field multiplied by a minus sign. This idea combined with the entrance pupil imaging system described above can be used to accurately relay the DOE surface field onto the entrance pupil. Experimentally the DOE/4f relay lens combination are translated
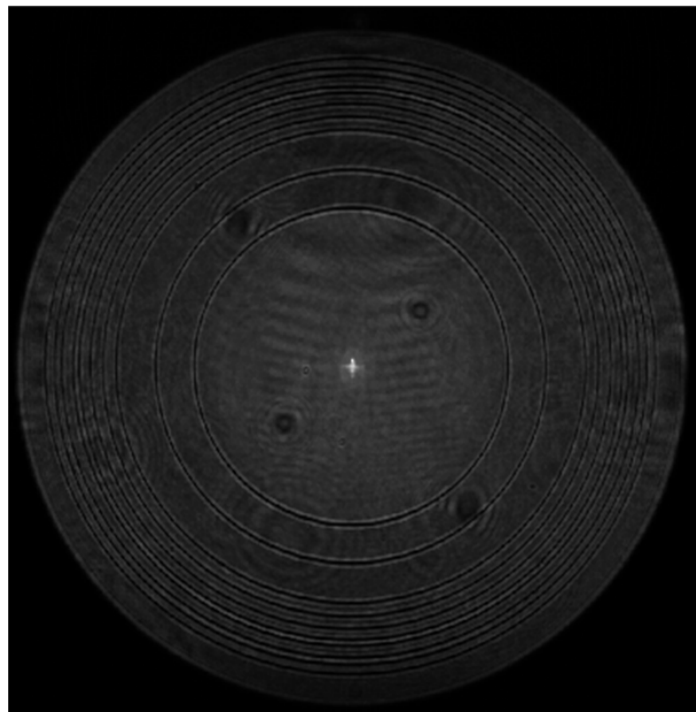
Figure 7.12. CCD image of the DOE at the entrance pupil of the objective lens at λ = 800 nm.

until the DOE profile comes into focus on the CCD camera. The imaged DOE profile at the back aperture is shown in Fig. 7.12. This procedure helps position the relayed DOE profile both longitudinally and transversely. The sharp rings observed in the relayed DOE are attributed to a zero field value at the edge of the zones. This results because the

interface between two consecutive zones lies in a region where two linearly polarized fields with a π phase difference add destructively.

7.5 Imaging the DOE modulated axial PSF

The DOE modified axial PSF was imaged and compared to both the diffraction-limited (DL) pattern (no DOE) and those computed theoretically. The immersion medium used in this experiment was the mixture that minimized spherical aberration (19.76 vol-% IL, $n = 1.5021$, Sect. 7.3.3). The simulated PSF image was calculated with Eq. (2.6) (no account for aberrations in the objective) and using the measured DOE zone widths reported in Sect. 7.2. The axial PSF imaged with and without the DOE is shown in Fig. 7.13. The DL central-lobe has a FWHM = 998 nm where as the central lobe observed with the DOE is apparently superresolved (SR) to FWHM = 760 nm. The experimental superresolution factor is $G = 760$ nm/998 nm $= 0.76$, which corresponds to a 24% improvement in axial resolution. The experimental superresolution is less than the value of 29% predicted theoretically.

Figure 7.14 compares the imaged and theoretically calculated axial DOE-modified PSFs. Even though the theoretical data were generated using experimentally measured DOE zone widths, there remains a clear discrepancy in side-lobe shape and intensity over the region [2 μm < $|z|$ < 6 μm]. This provides further evidence for aberrations in the objective lens. The asymmetry about $z = 0$ in the imaged PSF confirms the presence of spherical aberration, as a binary DOE should give a symmetric pattern for an unaberrated beam. The theoretical and experimental data do agree in the regions where $|z| > 6$ μm. The effect of the difference in zones width of the fabricated DOE from the theoretical one
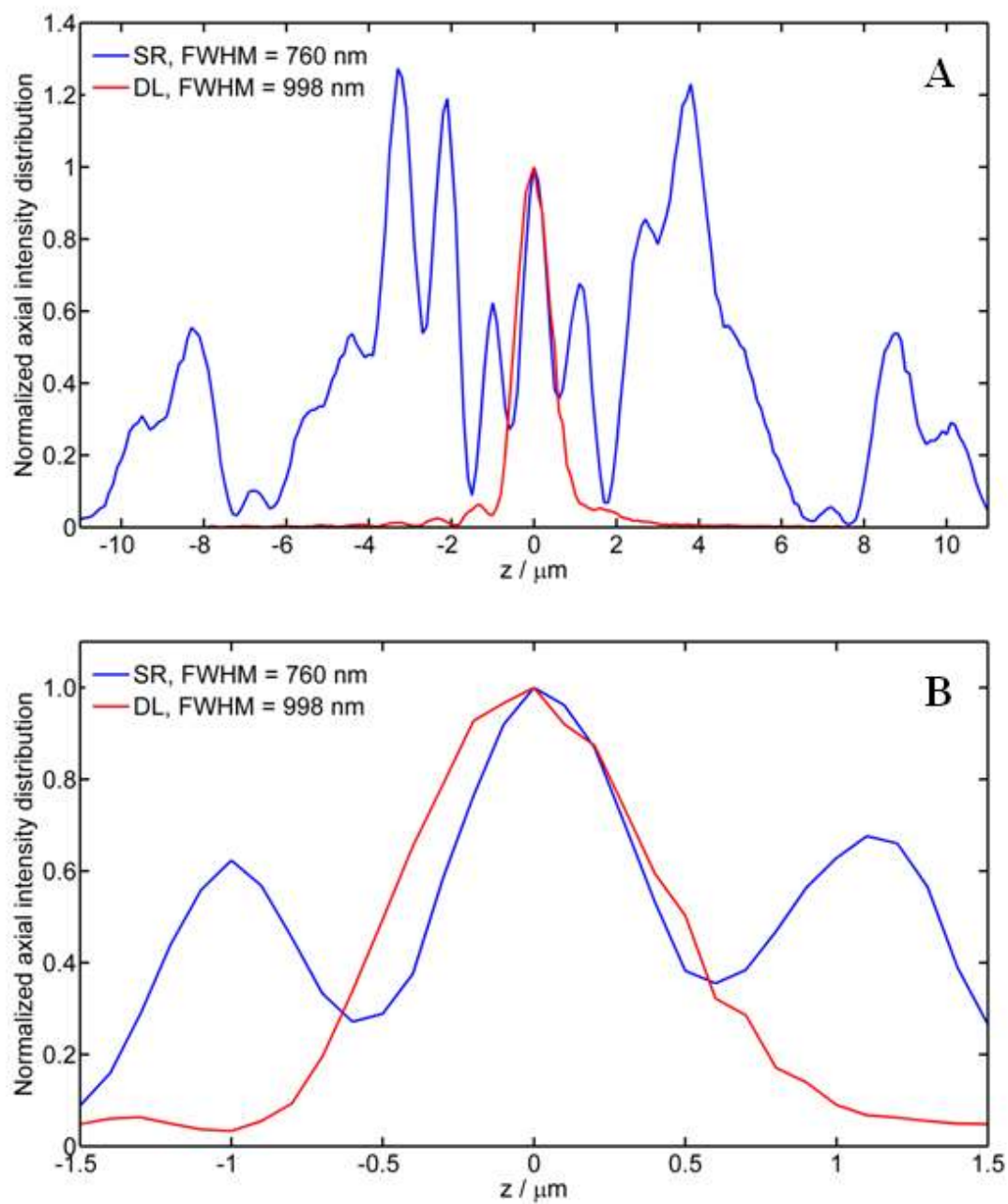
Figure 7.13. Measured axial intensity distribution. DL: Diffraction limit (sans DOE), SR: Superresolved (with DOE). The central-lobe of the axial distribution in (A) is shown in (B).

can be observed in the peak intensity of the first sidelobe. The imaged and calculated PSF yield a peak intensity of approximately 0.70. This is higher than the 0.52 value expected for the designed DOE. Such behavior is expected from the theoretical simulation of Sect. 6.2.1. It was shown that narrower DOE zones can result with a lower G at the expense of higher M.
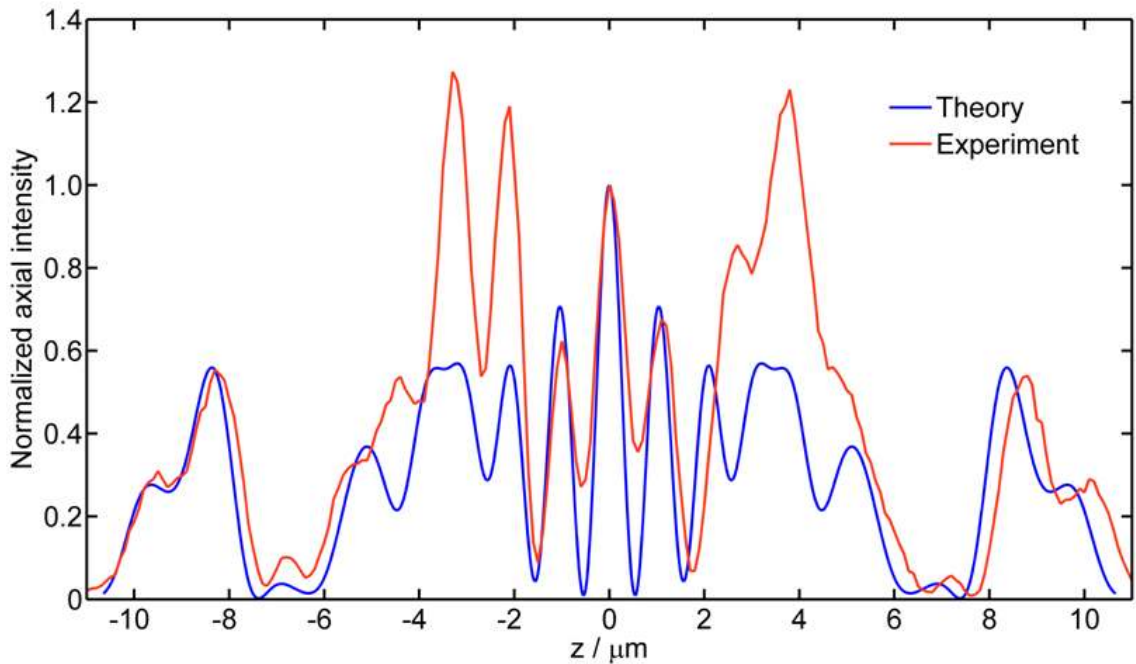
Figure 7.14. The DOE superresolved axial intensity distribution, experiment versus theory.

7.6 Objective lens aberration measurement

7.6.1 Introduction

The discrepancy between the experimental and theoretical axial intensity profile indicates that aberrations are still present in the objective lens despite compensation for

spherical aberration using the optimized immersion oil. As such it was decided to measure the aberrations of the objective quantitatively. Several creative techniques have been developed to characterize objective lenses. These include measuring the complex amplitude PSF (APSF) [101], aberrations at the entrance pupil aperture [112], or both [102]. Although there is an analytical relationship between the pupil field distribution and the APSF, solving for the pupil field distribution based on an APSF measurement is not trivial because of depolarization occurring under high-NA focusing and the problem is mathematically ill-posed [113]. Therefore, aberrations in an objective are best characterized by measuring the phase wavefront at the entrance pupil.

Juskaitis [113] and Charriere $et\ al.$ [114] used a Mach-Zehnder interferometer to obtain the phase distribution at the entrance pupil of an objective lens. Light emitted from a point source was collimated by the objective and interfered with a reference beam to produce an interferogram. The phase wavefront was then extracted from the interferogram. The point source is an essential part of the experimental setup, as it must be small enough to generate spherical waves with a solid angle greater than that of the objective lens. Juskaitis used gold nanoparticle scatterers excited through total internal reflection (TIR) at a prism interface as a point source. To achieve TIR at a glass-immersion oil interface, the refractive index of the prism glass has to exceed that of the oil. The signal-to-noise ratio in this method is generally poor because glass with large refractive index is highly scattering. Charriere $et\ al.$ used a near field scanning optical microscope (NSOM) tip as a point source. Although the NSOM-based method offers higher signal-to-noise ratio (S/N), it remains challenging to implement because the tip

can be easily damaged, particularly when in contact with the immersion oil, and its viscosity introduces drag that can distort the measurement.

In these investigations it was found that many of the limitations cited above can be overcome by using a nano-aperture fabricated in an optically thick metallic film as a point source. To ensure point source operation, the nano-aperture diameter must be limited to less than half the diffraction limited spot size of the objective. High S/N is possible because the nano-aperture can be excited by gently focusing a laser beam onto it in free space. This approach offers both good S/N in the interference pattern and a robust component that can be easily handled.

Phase extraction was accomplished using the two-dimensional Fourier Transform Method (2D-FTM) [115]. Compared to other approaches, such as phase shifting interferometry [116], 2D-FTM offers several advantages. First, it does not require an expensive nanopositioning stage to introduce an accurate phase shift. Second, the phasefront can be extracted from a single interferogram. This is extremely helpful in environments suffering from mechanical vibrations, thermal instability, or air turbulence. Lastly, high accuracy is possible as phase variations as small as $\lambda/100$ can be extracted.

7.6.2 Theory of wavefront reconstruction

Extracting the objective entrance pupil phase front from the fringe pattern is accomplished using the two-dimensional Fourier Transform Method (2D-FTM) [115]. 2D-FTM is an extension of a one dimensional approach originally developed by Takeda *et al*. [117]. The principle of 2D-FTM is as follows. Consider the intensity pattern resulting from the interference of a test beam and a reference beam:

$$i(\mathbf{r}) = a(\mathbf{r}) + b(\mathbf{r})\cos[\phi(\mathbf{r})], \tag{7.7}$$

where $\mathbf{r}$ is position vector of an arbitrary point (x,y) in the interferogram. The terms, $a(\mathbf{r})$ and $b(\mathbf{r})$, describe the additive and multiplicative intensity variations, respectively. The phase $\phi(\mathbf{r})$ is composed of a carrier wave $\phi_c(\mathbf{r})$ and the test beam wavefront phase $\phi_T(\mathbf{r})$:

$$\begin{aligned}\phi(\mathbf{r}) &= \phi_C(\mathbf{r}) + \phi_T(\mathbf{r}) \\ &= 2\pi f_C \mathbf{r} + \phi_T(\mathbf{r})\end{aligned} \tag{7.8}$$

The carrier frequency, f_C, is a function of the interference tilt angle. The cosine term in Eq. (7.7) can be expanded using Euler's equality as

$$i(\mathbf{r}) = a(\mathbf{r}) + c(\mathbf{r})\exp[j\phi_C(\mathbf{r})] + c*(\mathbf{r})\exp[-j\phi_C(\mathbf{r})], \tag{7.9}$$

where

$$c(\mathbf{r}) = (1/2)b(\mathbf{r})\exp[j\phi_T(\mathbf{r})]. \tag{7.10}$$

The asterisk denotes the complex conjugate. In order to calculate $\phi_T(\mathbf{r})$, the term $c(\mathbf{r})$ must be filtered out of the interference pattern. This can be achieved in the frequency domain. Taking the Fourier transform of $i(\mathbf{r})$ gives

$$I(\mathbf{f}) = A(\mathbf{f}) + C(\mathbf{f} - f_C) + C*(\mathbf{f} + f_C), \tag{7.11}$$

where uppercase letters are used to indicate the Fourier transform of the corresponding function in the spatial domain. The vector $\mathbf{f}$ gives a position in frequency space. Assuming that $a(\mathbf{r})$, $b(\mathbf{r})$, and $\phi_T(\mathbf{r})$ vary slowly with respect to f_C, the C term can be separated from the others by multiplication with an appropriate frequency filter.

115

Translating the C component by $-f_C$ to the origin and calculating the inverse Fourier transform yields $c(\boldsymbol{r})$. The wrapped phase wavefront $\phi_T(\boldsymbol{r})$ is then given by

$$\phi_T(\boldsymbol{r}) = \text{Im}\{[\ln[c(\boldsymbol{r})]\} , \qquad (7.12)$$

where "Im" denotes the imaginary part of a complex number. An unwrapping algorithm can then be applied to obtain the continuous phase wavefront profile.

7.6.3 Experimental setup

The experimental setup utilized to measure the pupil aberrations of the objective lens is shown in Fig. 5.15. The layout is based on a Mach-Zehnder interferometer. A 532 nm frequency doubled Nd:YVO$_4$ laser was split between a reference arm and a test arm, the latter of which contains the objective lens being characterized. Laser light in the test arm was delivered to the microscope stage through a single-mode fiber. The light at the fiber exit was collected, collimated and focused onto a nano-aperture.

The nano-aperture was fabricated in a 225-nm thick aluminum (Al) film deposited by vacuum thermal evaporation onto a 1-mm thick borosilicate microscope slide. A 30 keV Ga$^+$ focused ion beam (FIB, FEI 200 TEM) optimized at 10 pA current and 10 μs dwell time was used to mill a series of circular apertures with diameters ranging between 50 nm and 500 nm in the Al film. An SEM image of the 150-nm diameter pinhole used in this experiment is shown in Fig. 7.16. Aluminum was chosen because at 532 nm it offers high optical extinction ($\kappa \cong 6.5$) and high reflectivity in the bulk at normal incidence ($R = 92\%$) [118]. These characteristics ensure that a nano-aperture created in a sub-micron thick film of Al will be optically opaque around the hole, which ensures a large signal-to-noise ratio.
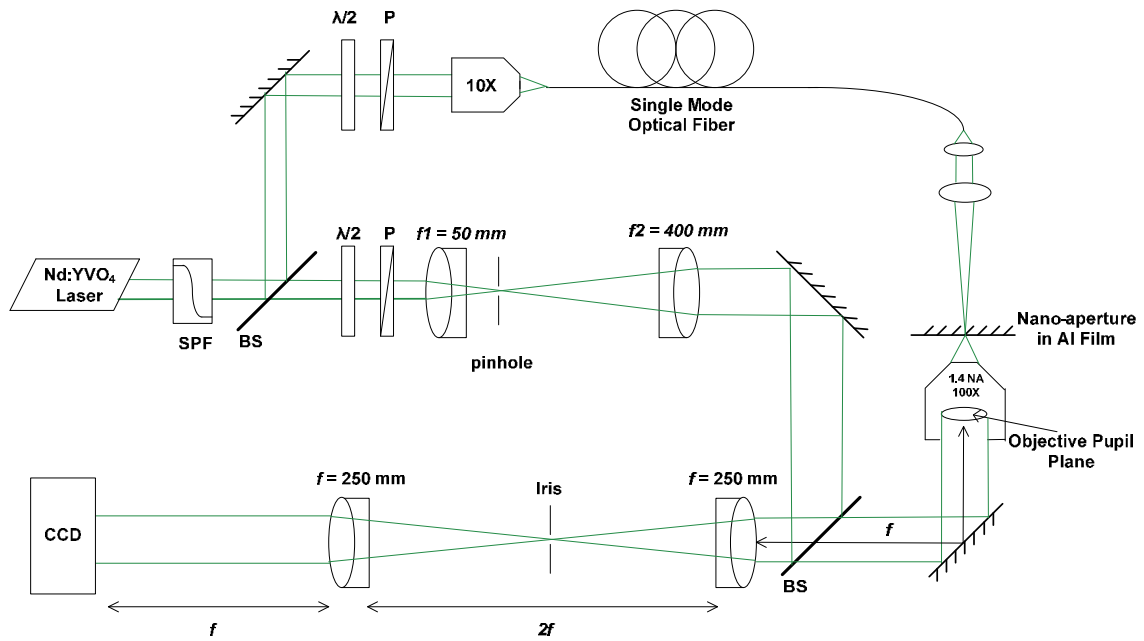
Figure 7.15. Experimental setup used to characterize the aberrations of a high *NA* objective lens.

The sub-wavelength aperture functions as a point source, so radiation emanates from it as an aberration free spherical wavefront. The objective lens collects the spherical wave and collimates it into a nearly planar test wave at the exit pupil. The collimated test wavefront is only perfectly planar if the objective lens is aberration free, so phase deviations that are present are a direct result of aberrations in the objective lens.

To observe aberration introduced by the objective, the test wave was interfered with light in reference arm. A beam expander combined with sub-diffraction spatial filter pinhole in the reference arm generated a reference wave with less than $\lambda/10$ wavefront distortion. The test wave and reference wave were interfered at the image plane of a high resolution CCD camera (6.45 µm pixel size, 1392 × 1040 pixels). The test wave was
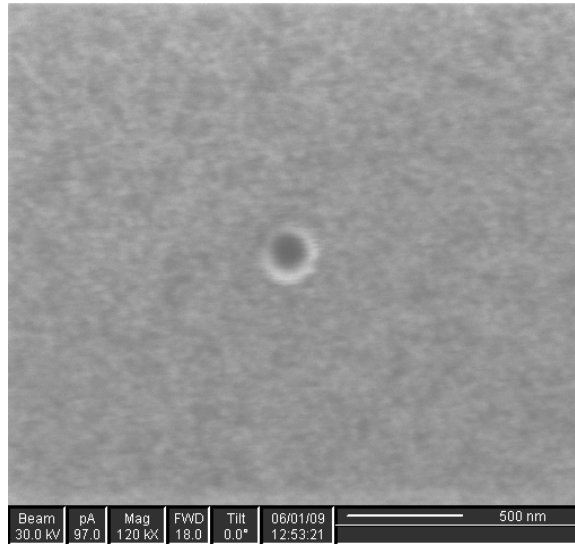
Figure 7.16. Scanning electron microscope image of a 150 nm pinhole in an Al film.

relayed onto the CCD array from the objective entrance pupil using a 4f relay system. This is necessary to measure the wavefront of the field at the entrance pupil. The test and reference waves were interfered in an off-axis configuration to introduce a tilt angle that sets the carrier frequency of the interference pattern. The tilt angle was optimized to introduce as many fringes across the field of view as possible while maintaining the pattern resolution at no less than 20 pixels per fringe. The intensity in each arm was adjusted with halfwave plate-polarizer combinations to maximize the contrast in the interference pattern. A short pass filter (SPF) blocked stray 809-nm radiation coming from the diode pump of the Nd:YVO$_4$ laser. The iris at the focal plane of the relay lenses was adjusted to block high spatial frequencies in the pupil irradiance without introducing artifacts into the interference pattern.

As in any interferometric measurement there are several sources of noise that can degrade the accuracy of the wavefront measurement. These sources can take the form of stray reflections, quantization errors, detector nonlinearity, laser instability, thermal instability, mechanical vibrations, and air turbulence. The last two sources are the most prominent cause of phase measurement errors. To minimize their effect, efforts were taken to stabilize the optical setup and minimize the time required to record an interferogram. All optomechanics were bolted to the vibration damping optical table. Cooling fans and rotating motors were switched off or removed from the table. The entire optical setup was isolated from air turbulence in the surrounding environment using plastic curtains. The fiber optic was attached to a vibration damping post at several locations along its length. The distance separating any two consecutive affixing points was limited to 10 cm. The laser light exiting the fiber was focused onto the nano-aperture to increase its brightness sufficiently that the interferogram recording times could be reduced to ~10 ms. Experimental data were collected at night when the laboratory was unoccupied.

7.6.4 Error in wavefront retrieval

2D-FTM, like other wavefront reconstruction algorithms, is susceptible to both experimental and computational errors. An ideal interferogram would consist of a uniform intensity envelope ($a = b = $ constant in Eq. (7.7)), high signal-to-noise, high modulation contrast, and an infinite spatial extent [119]. These conditions can only be achieved approximately in a real experiment, yet deviations from the ideal introduce error in the phase wavefront retrieval. Experimental errors can also arise due to detector sensitivity and resolution, optical misalignment of the interferometer, mechanical

vibration, or imperfections in optical components. Computational inaccuracies, on the other hand, are associated with theoretical approximations and calculations used in reconstructing the phase wavefront. These errors include, but are not limit to, aliasing, energy leakage between consecutive frequency components in the Fourier domain, and incorrect spectrum filtering [120]. From here forward, it is assumed that experimental errors were minimized by employing the measures detailed in Sect. 7.6.3. A discussion of data quantization and random noise error estimation can be found elsewhere [121]. In what follows, the implementation of the different theoretical steps will be illustrated, in addition to the several refinement techniques utilized to improve the wavefront retrieval accuracy.

7.6.5 Wavefront reconstruction procedure and results

Figure 7.17 shows the interferogram recorded by the CCD camera. The size of the interference pattern window is governed by the diameter of the objective entrance pupil. For an objective having a magnification $M = 100X$, $NA = 1.4$, and a tube-lens focal length $F_{tube} = 200$ mm, the entrance pupil diameter is $2F_{tube}NA/M = 5.6$ mm. This corresponds to approximately 868 pixels or sampling points across the interferogram.

Two important features should be noted in the recorded interference pattern. First, the interferogram intensity envelope decreases with distance from the center of the data window. This results because by theory the test wave radiates from the nano-aperture with a squared-cosine intensity distribution [122]. Referring now to Eq. (7.11), variations in the fringe intensity broaden the envelope of A, C and C^* in the frequency spectrum, making it more difficult to isolate lobe C (or C^*), which contain the phase wavefront
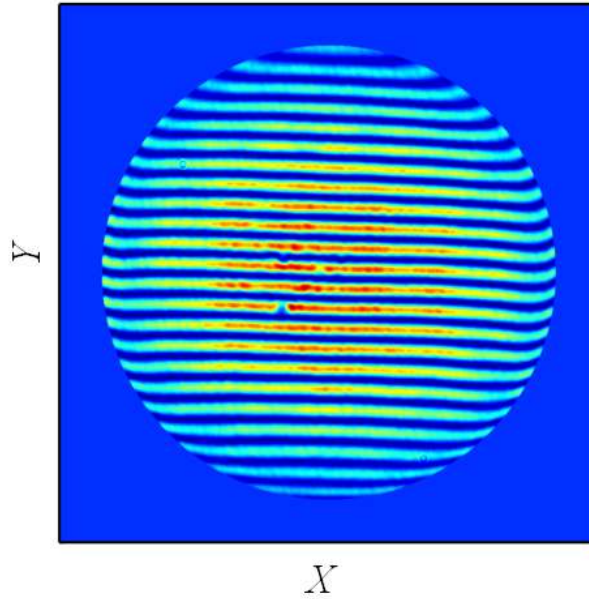
Figure 7.17. Interference of the reference and test beams recorded by the CCD camera.

information. Second, the interferogram is truncated at the periphery of the data window. The Fourier transform of a truncated signal extends to infinity and thus adds noise to wavefront information in the frequency space. Fringe extrapolation [119, 123] and applying a Hamming window to the data [117] can be used to minimize the effects of fringe truncation. Fringe extrapolation extends the interferogram beyond the window border, thus eliminating the discontinuity. A Hamming window, on the other hand, slowly attenuates the fringe amplitude toward the data window boundary.

Several methods have been developed for fringe extrapolations [119, 120, 123]. Among those, the iterative algorithm developed by Roddier *et al*. [119] and modified later by Massig *et al*. seems to give the best results [124]. The iterative approach, applied to the data in Fig. 7.17, begins by computing the 2D Fourier transform of the interferogram. The resulting spectrum, Fig. 7.18(A), is then multiplied by a frequency

filter consisting of narrow passbands centered around *A, C* and *C**. The frequency filter used, shown in Fig. 7.18(B), was a super-Gaussian of order ten. The smooth edges of the filter profile reduce the Gibb's effect error. An inverse Fourier transform is then computed to reconstruct the interferogram. The fringes inside the data window are replaced with the original pattern while data outside the window is left unchanged. These steps are repeated over many cycles. The extrapolated interferogram obtained after 1000 iterations is shown in Fig. 7.18(C). It can be seen that this procedure also extends the intensity envelope of fringe pattern because the frequency components of the *A* term are retained in the extrapolation. The width of each frequency filter was optimized through trial and error so that it was wide enough to retain high frequency features in the interferogram but sufficiently narrow to eliminate components associated with truncation of the data window. A key interferogram feature that was monitored during this process was the curves present in the fringes at the periphery of the data window.

A final step in the preconditioning of the interference pattern before applying the 2D-FTM is the multiplication by the Hamming window. The extrapolated interferogram multiplied by a Hamming window is shown in Fig. 7.18(D). The width of the Hamming window was chosen so that the first zeros of the cosine function fell at the edge of the extrapolated window.

The preconditioned interferogram was imported into the 2D-FTM algorithm to extract the phase wavefront. The algorithm computationally implements the steps described in Sect. 7.6.2. The spatial frequency filter used to isolate the *C* term was identical to that used during interferogram extrapolation, except that the filters for *A* and *C** were set to zero. After translating the *C* term by $-f_C$ to the origin and taking an inverse 2D FFT, the
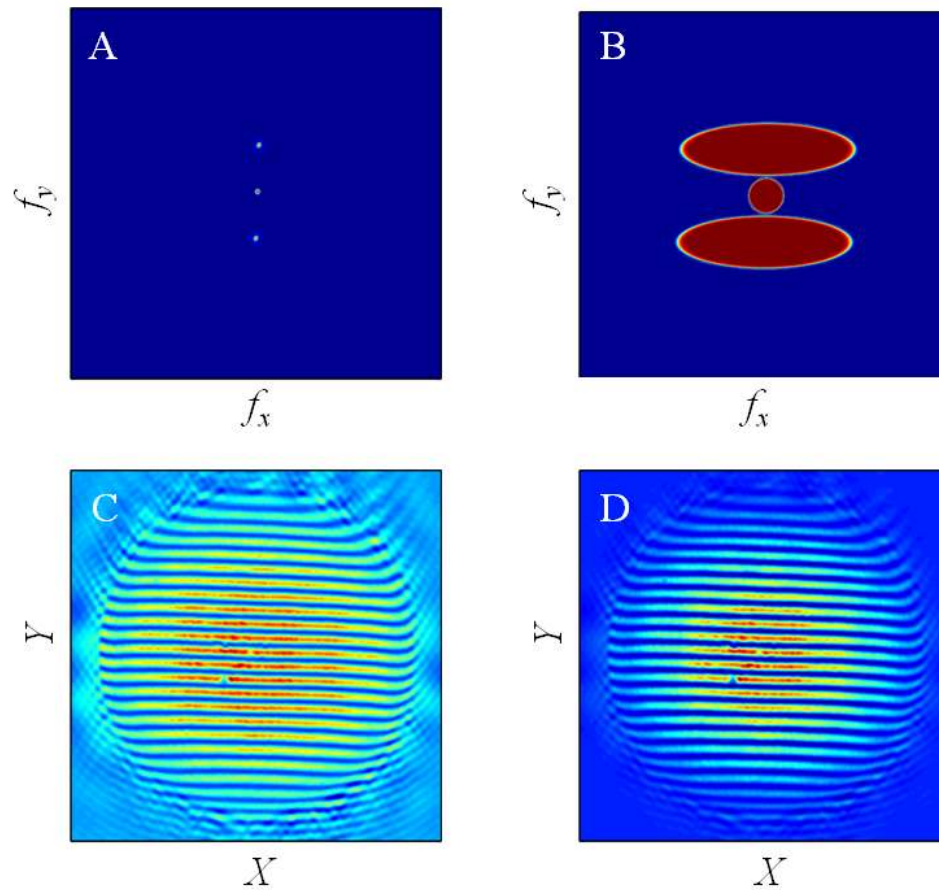
Figure 7.18. Interferogram treatment for wavefront reconstruction. (A) Fourier transform of the interference pattern shown in Fig. 5.17. (B) Frequency filter used in the fringe extrapolation algorithm. (C) Extrapolated inteferogram. (D) Extrapolated interferogram in (C) multiplied by a hamming window.

phase wavefront was calculated using Eq. (7.12). Two dimensional phase unwrapping was then applied to recover the full phase distribution using the procedure described by Macy [115]. The resulting phase wavefront inside the data window is shown in Fig. 5.19. The maximum deviation obtained across the phase profile is $\lambda/1.7$.

7.6.6 Zernike decomposition of the reconstructed wavefront

A Zernike polynomial decomposition of the reconstructed wavefront can be used to quantify each type of aberration present and identify those resulting from the objective alone. The maximum phase variation of the reconstructed wavefront shown in Fig. 7.19 is larger than can be attributed to the objective lens alone because the experimental and
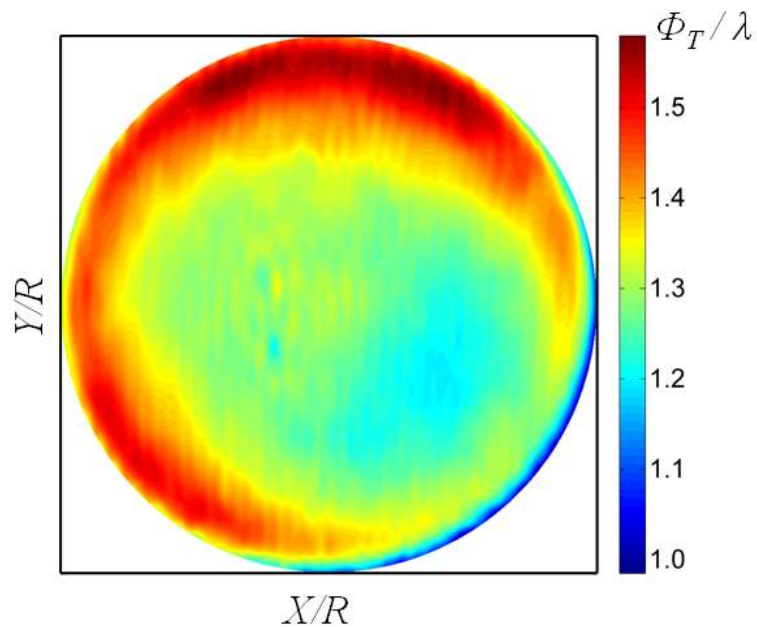
Figure 7.19. Reconstructed wavefront obtained using the 2D Fourier transform method.

data processing methods can introduce additional phase distortions that include piston, tilt, and defocus. Defocus results from mispositioning of the nano-aperture outside the focal plane. Residual tilt, on the other hand, appears when the C component is not shifted to exactly the origin in the frequency space. This will always be in error by at least the frequency space discretization T^{-1}, where T is the field window size.

Normalizing the radial extent of the data window to unity permits the reconstructed phase ϕ_T to be described as a sum of normalized Zernike polynomials $Z_j(x,y)$ [125] as

$$\phi_T = \sum_{j=1}^{\infty} a_j Z_j(x,y),$$ (7.13)

where a_j are coefficients corresponding to the jth Zernike polynomial. Each Zernike polynomial describes a specific aberration function and the index j identifies each aberration term. The Zernike polynomials are orthogonal over a unit circle and thus the coefficients a_j can be easily calculated as [126]:

$$a_j = \iint\limits_{x^2+y^2 \le 1} \phi_T(x,y) Z_j(x,y)\, dxdy .$$ (7.14)

The wavefront shown in Fig. 7.19 was decomposed into a summation over the first 45 Zernike polynomials. The number of polynomials used was chosen so the residual error between the original and decomposed wavefront was less than 4%. Figure 7.20(A) shows the reconstructed wavefront obtained after removing the distortions introduced by piston, defocus and tilt. The remaining phase error across the entrance pupil is $\lambda/2.8$. The coefficients corresponding to the first 25 aberrations are given in Fig. 7.20(B). It can be seen that the dominant wavefront distortions are astigmatism and spherical aberrations with some secondary coma. This is consistent with the qualitative observation discussed in Sect. 7.3.2.

To validate the wavefront measurement, Eq. (7.1) was used to calculate the imaged PSF of the objective with aberrations included as $\Phi_T(x,y) + \Phi_T(-x,-y)$ (see Sect. 7.3.1), and the result was compared to the experimentally measured distribution. The calculated and experimental imaged axial PSFs are shown in Fig. 7.21. The theoretical and

125

experimental data are in good agreement, which confirms that this approach is suitable

for characterizing pupil plane aberrations of a high-*NA* oil immersion objective lens.

Others researchers have suggested that the actual *NA* of a high performance objective
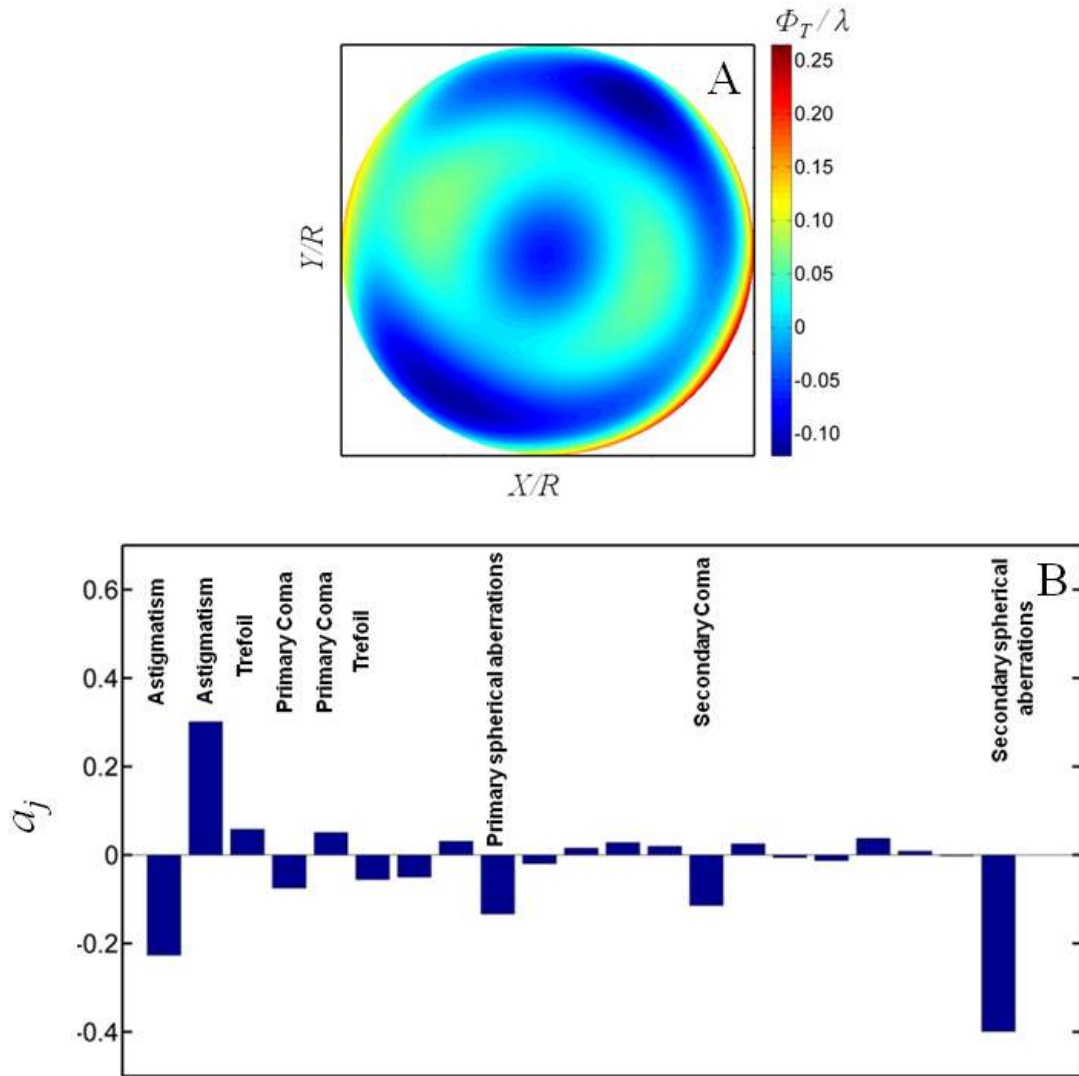
Figure 7.20. (A) Reconstructed wavefront minus piston, tilt and defocus. (B) The coefficients of the first 25 Zernike polynomials used to reconstruct wavefront in A. Those corresponding to piston, tilt, and defocus are not shown.

lens can differ slightly from that stated by the manufacturer [109], and such variations can explain discrepancies between calculated and experimentally measured PSFs. This possibility was considered by recalculating the imaged axial PSF with *NA* as a free parameter. As can be seen in Fig. 7.22, a better fit to the experiment data was obtained for *NA* = 1.42.
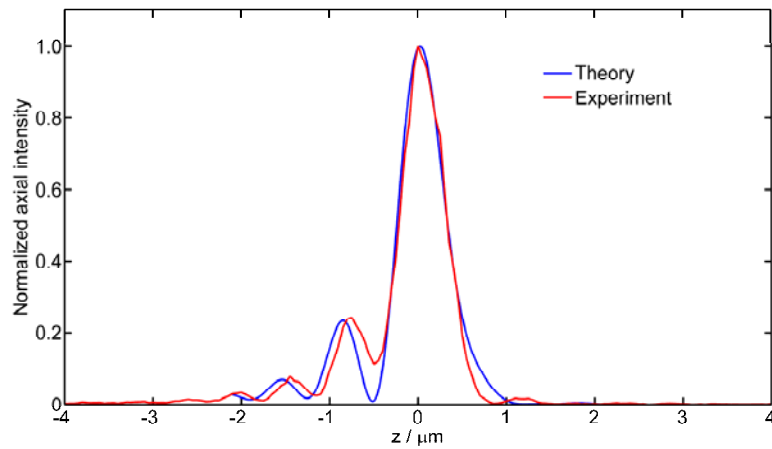
Figure 7.21. Comparison of the experimentally imaged axial PSF and that calculated with *NA* = 1.4.
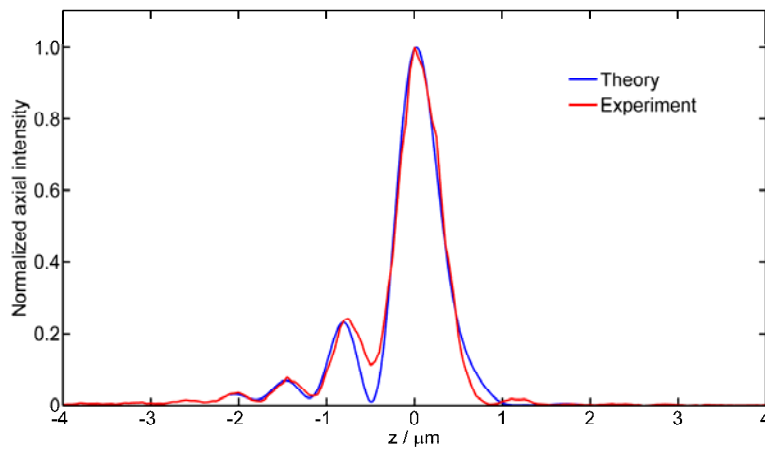
Figure 7.22. Comparison of the experimentally imaged axial PSF and that calculated with *NA* = 1.42.

Decomposition of the aberrated pupil wavefront into a set of Zernike polynomials permits a study of how each type of aberration affects the axial PSF. Using the methods described above, a set of axial intensity distributions was calculated that corresponds to the input field being distorted in phase by one or more of the experimentally determined quantities of astigmatism, coma, and primary and secondary spherical aberrations. Figure 7.23 shows the calculated axial PSFs obtained when all the aberrations are present and that due to primary and secondary spherical aberrations only. It appears that the dominant type of aberration in the particular objective characterized is secondary spherical aberration. The profile of the secondary spherical aberration along the radius of the entrance pupil is shown in Fig. 7.24. Apparently, the wavefront aberrations are concentrated around the center and toward the periphery of the pupil.
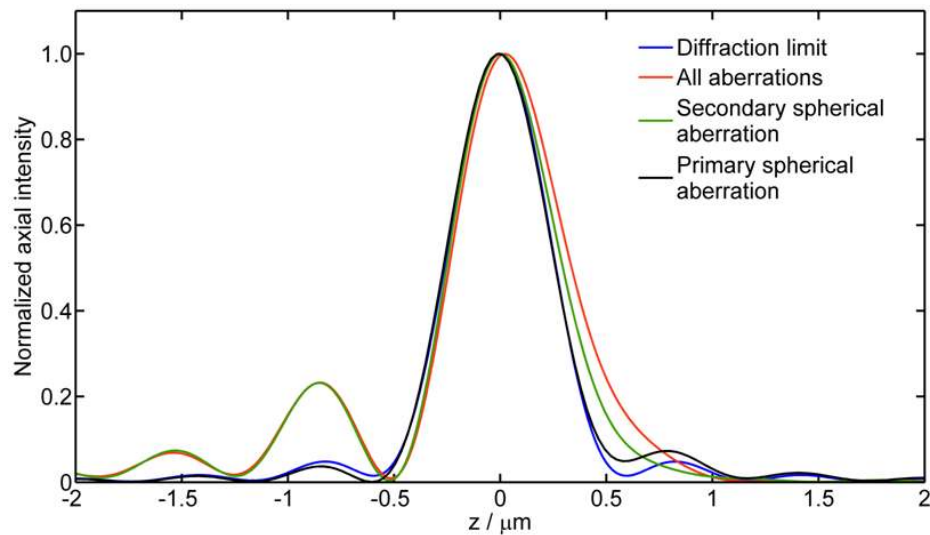
Figure 7.23. Comparison of the calculated axial intensity distributions obtained when the input field is aberrated by the experimentally measured amount of primary or secondary spherical aberration. The diffraction limited profile is superimposed for comparison.
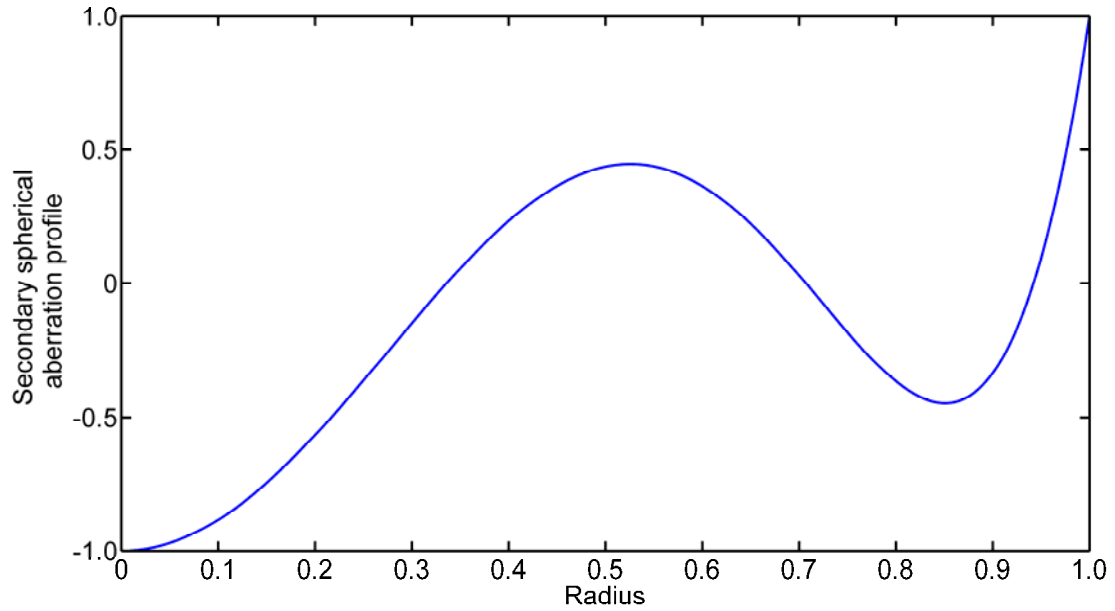
Figure 7.24. The radial phase profile of Zernike secondary spherical aberration.

The diffraction integral of Eq. (7.1) was used to model the effect of the measured aberrations on the DOE performance. The phase profile used in the calculation is $\Phi_T(k_x,k_y) + \Phi_T(-k_x,-k_y) + \Phi_{DOE}(k_x,k_y)$, where Φ_T and Φ_{DOE} describe the objective aberration and DOE phase profile, respectively. $\Phi_T(-k_x,-k_y)$ was introduced to account for the field passing through the objective twice. Figure 7.25 shows the axial intensity distribution as it would be imaged on the CCD camera. It should be noted that the axial plot in Fig. 7.25 was simulated for the conditions under which the pupil aberrations were characterized. These include using a wavelength $\lambda = 532$ nm and Type-DF immersion oil (no spherical aberration compensation). It can be clearly observed that the axial intensity shown in Fig. 7.25 shares a similar behavior to the experimental profile in Fig. 7.14 for negative z. Both profiles suffer from high side-lobe peak intensity corresponding to $M \approx$

1.2. The discrepancy between the two profiles for positive z is attributed to the difference in the immersion medium used or, in other words, the level of spherical aberration present in the optical system.
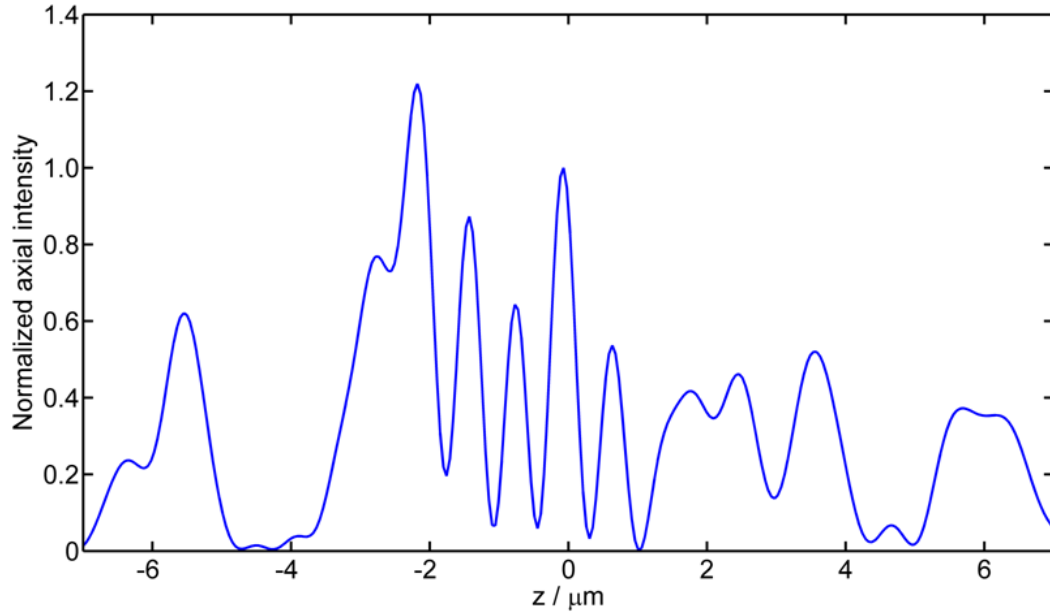
Figure 7.25. Axial intensity distribution demonstrating the effect of measured aberrations on the DOE performance. The phase profile used is $\Phi_T(k_x,k_y) + \Phi_T(-k_x,-k_y) + \Phi_{DOE}(k_x,k_y)$. The simulation was done for $\lambda = 532$ nm and without spherical aberration compensation.

7.7 Conclusion

The experimental work required to incorporate an axially superresolving DOE into a high NA optical system and to characterize its performance is described. The major steps of the experimental work included: 1) fabricating the DOE, 2) integrating the DOE into the optical system, 3) imaging the PSF and 4) measuring the pupil plane aberrations of the objective lens. The PSF images obtained with and without the DOE indicate the

presence of aberrations in the objective lens. These aberrations are detrimental to the performance parameters G and M of the DOE. The measured $G = 0.76$ is larger than the theoretically predicted value of 0.71. The largest disagreement between theory and experiment lies in the side-lobe regions. The aberrations can be partly compensated by adjusting the refractive index of the immersion oil used with the objective lens. An experiment was performed to study how the value of the refractive index affects the focal symmetry and FWHM of the axial PSF. It was found out by reducing n from 1.5070 to 1.5021, the asymmetry of the PSF increases from $S = 0.01$ to $S = 0.04$. The FWHM also improves significantly and decreases from 1.8 μm to 1 μm. To quantitatively measure and identify the aberrations, a Mach-Zehnder interferometer based experiment was used to measure pupil aberrations of the objective lens. The maximum wavefront distortion across the entrance pupil was found to be $\lambda/2.8$. The effect of the different aberrations present in the objective on the axial PSF was studied. It was found that secondary spherical aberration had the strongest impact on the axial PSF.

CHAPTER 8: OUTLOOK

The work described in this dissertation can be extended both theoretically and experimentally. The MGP algorithm utilized in chapter 5 to reshape the focal plane irradiance distribution can be developed to design DOEs capable of reshaping the full (3D) PSF of high-NA systems. The MGP is an iterative process that requires forward and backward propagation between the DOE and the 3D PSF. The most computationally efficient method to compute the 3D PSF is to represent the electromagnetic diffraction integrals in a 2D Fourier transform form and calculate the PSF at different planes in the focal region. Even if the 3D PSF is obtained, propagating it backward to obtain the field distribution at the pupil plane is not a trivial task. An alternative technique would be to propagate iteratively between the pupil plane and a single focal plane while choosing a different focal position in each iteration. Consider dividing the focal volume into a set of planes at z_1, z_2,.., z_N, where N is the total number of focal planes. The steps of the algorithm are described as follows: (1) propagate forward to obtain the field at z_1, (2) apply the constraints at z_1, (3) propagate backward to obtain the pupil field, (4) apply the pupil plane constraints. In the second iteration, the same steps are applied but now plane z_2 replaces z_1, Once all N z-planes are covered, the entire process is repeated over several cycles until the targeted 3D PSF distribution is achieved. This type of algorithm is referred to as block iterative. The constraints are partially applied in each iteration or in "blocks" [53].

The extension to the experimental work described in chapter 7 lies in using a spatial light modulator (SLM) to replace the fabricated DOE. An SLM consists of an array of

pixels that can be individually controlled to modulate the phase and/or amplitude distribution of a laser beam. The SLM can be used not only to impose the DOE pattern on the pupil plane but also to add the appropriate aberrations to negate those of the objective lens. The phase modulation needed to correct for the objective aberrations can be inferred from a pupil plane phase front measurement, as discussed in chapter 7, preformed at the operational wavelength.

The SLM, focal point imaging technique, and the PSO algorithm can be combined to perform beam shaping experimentally, thereby skipping the theoretical modeling step. In analogy to the theoretical beam shaping problem described in chapter 5, the focal irradiance distribution is obtained using the focal imaging technique and the DOE phase profile corresponds to the modulation applied at the SLM. This eliminates the need to propagate forward and backward between the DOE plane and the focal plane. In the experimental implementation, however, no constraints can be applied to the focal irradiance. As such, the MGP method would need to be replaced with the PSO algorithm. The PSO changes the design variables based on performance parameters describing irradiance distribution and thus it is well suited for this application. The uniformity error and diffraction efficiency can be used as parameters to the gauge the performance of the algorithm. The design variables can take the form of Zernike coefficients, which can be used to reconstruct the phase profile imposed by the SLM. The DOE phase profile designed using the proposed algorithm has several advantages over the MGP method. First, all the experimental imperfection such as objective aberrations and alignment errors are internally accounted for. Second, the PSO does not suffer from local traps and thus can achieve a more optimal solution.

133

APPENDIX A: DERIVATION OF THE VECTOR DIFFRACTION INTEGRALS

Consider the optical system shown in Fig. A.1. A linearly polarized electric field (E_0), having an arbitrary angle with respect to the meridional plane, is refracted by the lens towards the Gaussian focus O. The meridional plane is defined as the plane containing the optical axis and a typical ray traversing the optical system. The angle
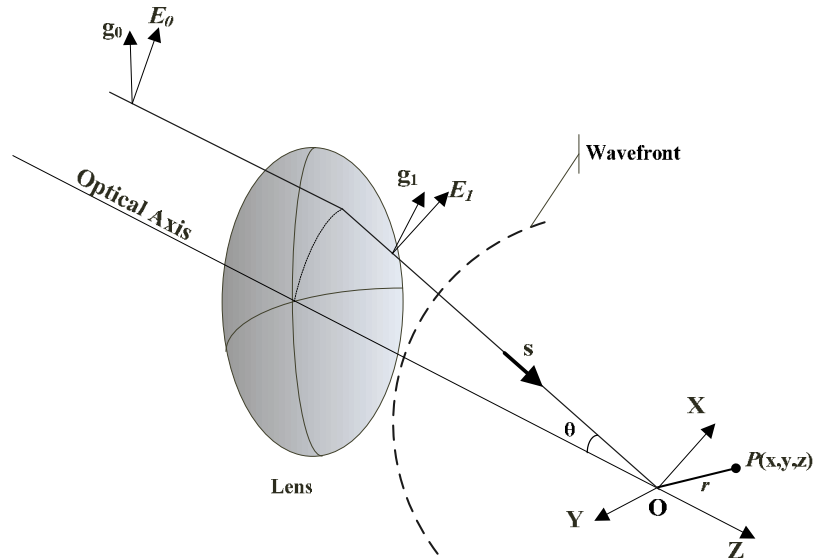
Figure A.1. Optical configuration utilized to derive the electromagnetic diffraction integrals. The meridional plane contains the optical axis and a typical ray traversing the system.

subtended between E_0 and the meridional plane is maintained by the field E_1 upon refraction. The space dependent electric field E at an arbitrary point P in the focal region is given by the electromagnetic diffraction integral derived by Wolf [4] as:

$$E(P) = -\frac{ik}{2\pi} \iint\limits_{\Omega} \frac{\mathbf{a}(s_x, s_y)}{s_z} \exp\{ik[\Phi(s_x, s_y) + \mathbf{s} \cdot \mathbf{r}(P)]\} ds_x ds_y , \qquad \text{(A.1)}$$

where $\mathbf{r}(P)$ is a radius vector connecting the point P to the origin O of the Cartesian coordinate system. The axis OX is oriented in the direction of the electric vector E_0 in the object space. The vector $\mathbf{s} = s_x \mathbf{i} + s_y \mathbf{j} + s_z \mathbf{k}$ defines the direction of a typical ray in the image space, where $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ are unit vectors in the direction of the coordinate axes. The function Φ denotes the aberrations in the optical system, and k is the wavenumber. The integration is taken over the solid angle Ω subtended between the optical axis and the periphery of the exit pupil. The unperturbed electric field strength vector $\mathbf{a}(s_x, s_y)$ is derived by Richards and Wolf [5], to be:

$$\mathbf{a} = fl_0\left(\cos^{1/2}\theta\right)\left[(\mathbf{g_0} \cdot \mathbf{i})\mathbf{g_1} + (\mathbf{g_0} \cdot \mathbf{j})(\mathbf{g_1} \times \mathbf{s})\right], \qquad \text{(A.2)}$$

where $\mathbf{g_0}$ and $\mathbf{g_1}$ are two unit vectors, lying in the meridional plane, perpendicular to the ray in the object space and the image space, respectively. The parameters f and l_0 are the focal radius of the Gaussian sphere and amplitude factor in the image space, respectively. As described by Kant [93], $\mathbf{g_0}$ and $\mathbf{g_1}$ can be evaluated as:

$$\mathbf{g0} = \frac{(\mathbf{s} \times \mathbf{k}) \times \mathbf{k}}{|\mathbf{s} \times \mathbf{k} \times \mathbf{k}|} = \frac{-s_x \mathbf{i} - s_y \mathbf{j}}{\left(s_x^2 + s_y^2\right)^{1/2}}, \qquad \text{(A.3)}$$

and

$$\mathbf{g_1} = \frac{(\mathbf{s} \times \mathbf{k}) \times \mathbf{s}}{|\mathbf{s} \times \mathbf{k} \times \mathbf{s}|} = \frac{-s_x s_z \mathbf{i} - s_y s_z \mathbf{j} + \left(s_x^2 + s_y^2\right)\mathbf{k}}{\left(s_x^2 + s_y^2\right)^{1/2}}. \tag{A.4}$$

Substituting Eqs. (A.4) and (A.3) in Eq. (A.2) we obtain:

$$\mathbf{a}(s_x, s_y) = fl_0 \sqrt{s_z} \left[\frac{\left(s_x^2 s_z + s_y^2\right)\mathbf{i} - s_x s_y (s_z - 1)\mathbf{j} + s_x\left(s_x^2 + s_y^2\right)\mathbf{k}}{\left(s_x^2 + s_y^2\right)} \right]. \tag{A.5}$$

The components s_x, s_y, s_z of the unit vector $\mathbf{s}$ are related to the components (k_x, k_y, k_z) of the wavenumber (k) as:

$$s_x = \frac{k_x}{k}, \quad s_y = \frac{k_y}{k}, \quad s_z = \frac{k_z}{k}. \tag{A.6}$$

Utilizing Eqs. (A.5) and (A.6), and expanding $\mathbf{s} \cdot \mathbf{r}(P) = s_x x_f + s_y y_f + s_z z_f$, where x_f, y_f and z_f are the coordinates of P, Eq. (A.1) can be rewritten following some algebraic manipulations and variables regrouping as:

$$\mathbf{E} = \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = -i\frac{l_0 f}{2\pi} \iint\limits_{\sqrt{k_x^2 + k_y^2} \leq k_{max}} e^{i\Phi(k_x,k_y)} \sqrt{\frac{k_z}{k}} \frac{1}{k_z} \begin{bmatrix} \dfrac{k_x^2(k_z/k) + k_y^2}{k_x^2 + k_y^2} \\ \dfrac{k_x k_y (k_z/k - 1)}{k_x^2 + k_y^2} \\ -\dfrac{k_x}{k} \end{bmatrix} \times e^{i(k_x x_f + k_y y_f + k_z z_f)} dk_x dk_y. \tag{A.7}$$

136

APPENDIX B: SPECIAL FUNCTIONS

B.1 Gegenbauer polynomials

The Gegenbauer polynomials $C_n^\alpha(x)$ are solutions to the differential equation [127] :

$$\left(1-x^2\right)\frac{d''y}{dx^2} - \left(2\alpha+1\right)x\frac{d'y}{dx} + n(n+2\alpha)y = 0. \tag{B.1}$$

For α < -0.5, the polynomials can be evaluated recursively as follows [127]:

$$
\begin{aligned}
&C_0^\alpha(x) = 1, \\
&C_1^\alpha(x) = 2\alpha\,x, \\
&C_n^\alpha(x) = \frac{(2n-2+2\alpha)x\,C_{n-1}^\alpha(x) + (-n+2-2\alpha)C_{n-2}^\alpha(x)}{n}.
\end{aligned} \tag{B.2}
$$

B.2 Spherical Bessel functions of the first kind

The spherical Bessel functions of the first kind, $j_n(z)$, are particular solutions to the differential equation [128]:

$$z^2\frac{d''w}{dz^2} + 2z\frac{d'w}{dz} + \left[z^2 - n(n+1)\right]w = 0; \; (n = 0,\pm1,\pm2,...). \tag{B.3}$$

$j_n(z)$ is related to the ordinary Bessel functions $J_v(z)$ as [128]:

$$j_n(z) = \sqrt{\frac{\pi}{2z}}J_{n+1/2}(z). \tag{B.4}$$

137

$J_v(z)$ is defined as [128]:

$$J_v(z) = \left(\frac{z}{2}\right)^v \sum_{k=0}^{\infty} \frac{\left(-\frac{z^2}{4}\right)^k}{k!\Gamma(v+k+1)},$$

(B.5)

where the gamma function Γ is given as [128]:[128]

$$\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt .$$

(B.6)

LIST OF REFERENCES

1. S. F. Pereira and A. S. van de Nes, "Superresolution by means of polarization, phase and amplitude pupil masks," Opt. Commun. 234, 119-124 (2004).

2. J. Turunen and F. Wyrowski, *Diffractive Optics for Industrial and Commercial Applications* (Wiley, Jena, 1998).

3. M. Martinez-Corral, R. Martinez-Cuenca, I. Escobar, and G. Saavedra, "Reduction of focus size in tightly focused linearly polarized beams," Appl. Phys. Lett. 85, 4319-4321 (2004).

4. E. Wolf, "Electromagnetic diffraction in optical systems I. An integral representation of the image field," Proc. Roy. Soc. A 253, 349-357 (1959).

5. B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems II. Structure of the image field in an aplanatic system," Proc. Roy. Soc. A 253, 358-379 (1959).

6. P. Adamson, "High-aperture focusing systems: control of light concentration in focal region by pupil filtering," J. Mod. Optics 51, 65-74 (2004).

7. C. M. Blanca and S. W. Hell, "Axial superresolution with ultrahigh aperture lenses," Opt. Express 10, 893-898 (2002).

8. M. P. Cagigal, J. E. Oti, V. F. Canales, and P. J. Valle, "Analytical design of superresolving phase filters," Opt. Commun. 241, 249-253 (2004).

9. V. F. Canales and M. P. Cagigal, "Pupil filter design by using a Bessel functions basis at the image plane," Opt. Express 14, 10393-10402 (2006).

10. V. F. Canales, J. E. Oti, and M. P. Cagigal, "Three-dimensional control of the focal light intensity distribution by analytically designed phase masks," Opt. Commun. 247, 11-18 (2005).

11. H. Y. Chen, N. Mayhew, E. G. S. Paige, and G. G. Yang, "Design of the point spread function of a lens, binary phase filter combination and its application to photolithography," Opt. Commun. 119, 381-389 (1995).

12. S. Ching-Cherng and L. Chin-Ku, "Ultrasmall focusing spot with a long depth of focus based on polarization and phase modulation," Opt. Lett. 28, 99-101 (2003).

13. D. M. de Juana, V. F. Canales, P. J. Valle, and M. P. Cagigal, "Focusing properties of annular binary phase filters," Opt. Commun. 229, 71-77 (2004).

14. D. M. de Juana, J. E. Oti, V. F. Canales, and M. P. Cagigal, "Design of superresolving continous phase filters," Opt. Lett. 28, 607-609 (2003).

15. H. Ding, Q. Li, and W. Zou, "Design and comparison of amplitude-type and phase-only transverse super-resolving pupil filters," Opt. Commun. 229, 117-122 (2004).

16. C. Dorrer and J. D. Zuegel, "Design and analysis of binary beam shapers using error diffusion," J. Opt. Soc. Am. B 24, 1268-1275 (2007).

17. J. Jia, C. Zhou, X. Sun, and L. Liu, "Superresolution laser beam shaping," Appl. Opt. 43, 2112-2117 (2004).

18. S. Ledesma, J. Campos, J. C. Escalera, and M. J. Yzuel, "Symmetry porperties with pupil phase-filters," Opt. Express 12, 2548-2559 (2004).

19. S. Ledesma, J. C. Escalera, J. Campos, and M. J. Yzuel, "Evolution of the transverse response of an optical system with complex filters," Opt. Commun. 249, 183-192 (2005).

20. L. Lin, X. Wang, Z. Wang, and G. Mu, "Binary phase-only pupil filters application to high-aperture confocal microscopy," Optik 114, 329-332 (2003).

21. H. Liu, Y. Yan, and G. Jin, "Design theories and performance limits of diffractive superresolution elements with the highest sidelobe suppressed," J. Opt. Soc. Am. A 22, 828-838 (2005).

22. H. Liu, Y. Yan, D. Yi, and G. Jin, "Theories for the design of a hybrid refractive-diffractive superresolution lens with high numerical aperture," J. Opt. Soc. Am. A 20, 913-924 (2003).

23. H. Liu, Y. Yan, D. Yi, and G. Jin, "Design of three-dimensional superresolution filters and limits of axial optical superresolution," Appl. Opt. 42, 1463-1476 (2003).

24. M. Martínez-Corral, P. Andrés, J. Ojeda-Castañeda, and G. Saavedra, "Tunable axial superresolution by annular binary filters. Application to confocal microscopy," Opt. Commun. 119, 491-498 (1995).

25. M. Martínez-Corral, P. Andrés, C. J. Zapata-Rodríguez, and M. Kowalczyk, "Three-dimensional superresolution by annular binary filters," Opt. Commun. 165, 267-278 (1999).

26. M. Martinez-Corral, M. T. Caballero, E. H. K. Stelzer, and J. Swoger, "Tailoring the axial shape of the point spread function using the Toraldo concept," Opt. Express 10, 98-103 (2002).

27. M. Martinez-Corral, C. Ibáñez-López, G. Saavedra, and M. T. Caballero, "Axial gain resolution in optical sectioning fluorescence microscopy by shaded-ring filters," Opt. Express 11, 1740-1745 (2003).

28. J. McOrist, M. D. Sharma, C. J. R. Sheppard, E. West, and K. Matsuda, "Hyperresolving phase-only filters with an optically addressable liquid crystal spatial light modulator," Micron 34, 327-332 (2003).

29. D. Mugnai, A. Ranfagni, and R. Ruggeri, "Pupils with super-resolution," Phys. Lett. A 311, 77-81 (2003).

30. S. Quabis, R. Dorn, M. Eberler, O. Glockl, and G. Leuchs, "Focusing light to a tighter spot," Opt. Commun. 179, 1-7 (2000).

31. A. Ranfagni, D. Mugnai, and R. Ruggeri, "Beyond the diffraction limit: Super-resolving pupils," Appl. Phys. 95, 2217-2222 (2004).

32. T. R. M. Sales and G. M. Morris, "Diffractive superresolution elements," J. Opt. Soc. Am. A 14, 1637-1646 (1997).

33. T. R. M. Sales and G. M. Morris, "Axial superresolution with phase-only pupil filters," Opt. Commun. 156, 227-230 (1998).

34. C. J. R. Sheppard and A. Choudhury, "Annular pupils, radial polarization, and superresolution," Appl. Opt. 43, 4322-4327 (2004).

35. C. J. R. Sheppard and Z. S. Hegedus, "Axial behavior of pupil-plane filters," J. Opt. Soc. Am. A 5, 643-647 (1988).

36. H. Wang, Z. Chen, and F. Gan, "Phase-shifting apodizer for next-generation digital versatile disk," Opt. Eng. 40, 991-994 (2001).

37. H. Wang and F. Gan, "High focal depth with pure-phase apodizer," Appl. Opt. 40, 5658-5662 (2001).

38. A. I. Whiting, A. F. Abouraddy, B. E. A. Saleh, M. C. Teich, and J. T. Fourkas, "Polarization-assisted transverse and axial optical superresolution," Opt. Express 11, 1714-1723 (2003).

39. F.-R. Xiao, J.-H. Yuan, G.-Y. Wang, and Z.-Z. Xu, "Factors in Design of Optical Super-Resolving Filters," Chin. Phys. Lett. 21, 70-72 (2003).

40. Y. Zhang, C. Zheng, and Y. Zou, "Focal-field distribution of the solid immersion lens system with an annular filter," Optik 115, 277-280 (2004).

41. X.-F. Zhao, C.-F. LI, and H. Ruan, "Improvement of Three-Dimensional Resolution in Optical Data Storage by Combination of Two Annular Binary Phase Filters," Chin. Phys. Lett. 21, 1515-1517 (2004).

42. S. Zhou and C. Zhou, "Discrete continuous-phase superresolving filters," Opt. Lett. 29, 2746-2748 (2004).

43. T. G. Jabbour and S. M. Kuebler, "Vector diffraction analysis of high numerical aperture focused beams modified by two- and three-zone annular multi-phase plates," Opt. Express 14, 1033-1043 (2006).

44. T. G. Jabbour and S. M. Kuebler, "Axial field shaping under high-numerical-aperture focusing," Opt. Lett. 32, 527-529 (2007).

45. T. G. Jabbour and S. M. Kuebler, "Design of axially super-resolving phase filters using the method of generalized projections," Opt. Commun. 281, 2002 - 2011 (2008).

46. T. G. Jabbour and S. M. Kuebler, "Particle-swarm optimization of axially superresolving binary-phase diffractive optical elements," Opt. Lett. 33, 1533 - 1535 (2008).

47. T. G. Jabbour and S. M. Kuebler, "Vectorial beam shaping," Opt. Express 16, 7203 - 7213 (2008).

48. A. Diaspro, *Confocal and two-photon microscopy : foundations, applications, and advances* (Wiley-Liss, New York, 2002).

49. G. Yang, "An optical pickup using a diffractive optical element for a high-density optical disc," Opt. Commun. 159, 19-22 (1999).

50. C. Ibáñez-López, G. Saavedra, G. Boyer, and M. Martínez-Corral, "Quasi-isotropic 3-D resolution in two-photon scanning microscopy," Opt. Express 13, 6168-6174 (2005).

51. C. Ibáñez-López, G. Saavedra, K. Plamann, G. Boyer, and M. Martínez-Corral, "Quasi-spherical focal spot in two-photon scanning microscopy by three-ring apodization," Microsc. Res. Tech. 67, 22-26 (2005).

52. B. Kress and P. Meyrueis, *Digital Diffractive Optics* (John Wiley & Sons, LTD, New York, 2000).

53. R. Piestun and J. Shamir, "Synthesis of three-dimensional light fields and applications," Proc. of the IEEE 90, 222-244 (2002).

54. J. Rosen, "Synthesis of nondiffracting beams in free space," Opt. Lett. 19, 369-371 (1994).

55. J. Rosen and A. Yariv, "Synthesis of an arbitrary axial field profile by computer-generated holograms," Opt. Lett. 19, 843-845 (1994).

56. M. Yun, L. Liu, J. Sun, and D. a. Liu, "Three-dimensional superresolution by three-zone complex pupil filters," J. Opt. Soc. Am. A 22, 272-277 (2005).

57. Z. S. Hegedus and V. Sarafis, "Superresolving filters in confocally scanned imaging systems," J. Opt. Soc. Am. A 3, 1892 - 1896 (1986).

58. J. J. Stamnes, *Waves in Focal Regions: Propagation, Diffraction and Focusing of Light, Sound and Water Waves*, The Adam Hilger Series on Optics and Optoelectronics (Adam Hilger, Bristol, 1986).

59. S. M. Kuebler and M. Rumi, "Nonlinear optics -- applications: three-dimensional microfabrication," in *Encyclopedia of Modern Optics*, R. D. Guenther, D. G. Steel, and L. Bayvel, eds. (Elsevier, Oxford, 2004), pp. 189-206.

60. S. M. Kuebler, M. Rumi, T. Watanabe, K. Braun, B. H. Cumpston, A. A. Heikal, L. L. Erskine, S. Thayumanavan, S. Barlow, S. R. Marder, and J. W. Perry, "Optimizing two-photon initiators and exposure conditions for three-dimensional lithographic microfabrication," J. Photopolym. Sci. Technol. 14, 657-668 (2001).

61. H. B. Sun, K. Takada, M. S. Kim, K. S. Lee, and S. Kawata, "Scaling laws of voxels in two-photon photopolymerization nanofabrication," Appl. Phys. Lett. 83, 1104-1106 (2003).

62. A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," J. Opt. Soc. Am. A 1, 932-943 (1984).

63. F. Wyrowski, "Diffractive optical elements: iterative calculations of quantized, blazed phase structures," J. Opt. Soc. Am. A 7, 961-969 (1990).

64. A. D. Polyanin and A. V. Manzhirov, *Handbook of Integral Equations* (CRC Press, Boca Raton, 1998).

65. R. Kant, "Superresolution and increased depth of focus: an inverse problem of vector diffraction," J. Mod. Optics 47, 905-916 (2000).

66. W. C. Catino, J. L. LoCicero, and H. Stark, "Design of continuous and quantized amplitude holograms by generalized projections," J. Opt. Soc. Am. A 15, 68-76 (1998).

67. N. K. Bao, Z. Chen, and J. Zhang, "Phase-only diffractive optical element for improving the beam quality of a vertical-cavity surface-emitting laser array," Optik 115, 23-27 (2004).

68. F. Wyrowski, "Diffraction efficiency of analog and quantized digital amplitude holograms: analysis and manipulation," J. Opt. Soc. Am. A 7, 383-393 (1990).

69. P. D. Higdon, P. Török, and T. Wilson, "Imaging properties of high aperture multiphoton fluorescence scanning optical microscopes," J. Microsc. 193, 127-141 (1999).

70. I. Akduman, U. Brand, J. Grochmalicki, G. Hester, R. Pike, and M. Bertero, "Superresolving masks for incoherent high-numerical-aperture scanning microscopy in three dimensions," J. Opt. Soc. Am. A 15, 2275-2287 (1998).

71. P. Venkataraman, *Applied Optimization with MATLAB Programming* (Wiley, New York, 2002).

72. J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *International Conference on Neural Networks*, (IEEE, 1995), 1942-1948

73. J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *International Conference on Systems, Man, and Cybernetics*, (IEEE, 1997), 4104 - 4108.

74. X. Hu and R. C. Eberhart, "Multiobjective optimization using dynamic neighborhood particle swarm optimization," in *Congr. Evolutionary Computation (CEC'2002)*, (IEEE, 2002), 1677-1681.

75. S. Lee, H. Park, and M. Jeon, "Binary particle swarm optimization with bit change mutation," IEICE Trans. Fundamentals E90-A, 2253 - 2256 (2007).

76. L. Yuan and Z.-D. Zhao, "A modified binary particle swarm optimization algorithm for permutation flow shop problem," in *International Conference on Machine Learning and Cybernetics* (IEEE, 2007), 19 - 22.

77. Y. Shi and R. C. Eberhart, "Parameter selection in particle swarm optimization," in *Evolutionary Programming VII: Proceedings of the Seventh Annual Conference on Evolutionary Programming*, (Springer-Verlag, 1998), 591 - 600.

78. D. A. Pommet, M. G. Moharam, and E. B. Grann, "Limits of scalar diffraction theory for diffractive phase elements," JOSA A 11, 1827 - 1834 (1994).

79. F. M. Dickey, S. C. Holswade, and D. L. Shealy, *Laser Beam Shaping Applications* (Taylor & Francis Group, Boca Raton, 2006).

80. V. Ramírez-Sánchez and G. Piquero, "Global beam shaping with nonuniformily polarized beams: a proposal," Appl. Opt. 45, 8902-8906 (2006).

81. L. A. Romero and F. M. Dickey, "Lossless laser beam shaping," J. Opt. Soc. Am. A 13, 751-760 (1995).

82. O. Bryngdhal, "Geometrical transformations in optics," J. Opt. Soc. Am. 64, 1092-1099 (1974).

83. D. L. Shealy and S. H. Chao, "Geometric optics-based design of laser beam shapers," Opt. Eng. 42, 3123-3138 (2003).

84. H. Aagedal, M. Schmid, S. Egner, J. Müller-Quade, T. Beth, and F. Wyrowski, "Analytical beam shaping with application to laser-diode arrays," J. Opt. Soc. Am. A 14, 1549-1553 (1997).

85. M. Johansson and J. Bengtsson, "Robust design method for highly efficient beam-shaping diffractive optical elements using iterative-Fourier-transform algorithm with soft operations," J. Mod. Optics 47, 1385-1398 (2000).

86. V. V. Kotlyar, P. G. Seraphimovich, and V. A. Soifer, "An iterative algorithm for designing diffractive optical elements with regularization," Opt. Laser Eng. 29, 261-268 (1998).

87. J. S. Liu and M. R. Taghizadeh, "Iterative algorithm for the design of diffractive phase elements for laser beam shaping," Opt. Lett. 27, 1463-1465 (2002).

88. G. Zhou, X. Yuan, P. Dowd, Y. L. Lam, and Y. C. Chan, "Design of diffractive phase elements for beam shaping: hybrid approach," J. Opt. Soc. Am. A 18, 791-800 (2001).

89. M. Kuittinen, P. Vahimaa, M. Honkanen, and J. Turunen, "Beam shaping in the nonparaxial domain of diffractive optics," Appl. Opt. 36, 2034-2041 (1997).

90. Y. Zhao, Y.-P. Li, and Q.-G. Zhou, "Vector iterative algorithm for the design of diffractive optical elements applied to uniform illumination," Opt. Lett. 29, 664-666 (2004).

91. J. L. Bakx, "Efficient computation of optical disk readout by use of the chirp z transform," Appl. Opt. 41, 4897-4903 (2002).

92. H. Kim, B. Yang, and B. Lee, "Iterative Fourier transform algorithm with regularization for the optimal design of diffractive optical elements," J. Opt. Soc. Am. A 21, 2353-2365 (2004).

93. R. Kant, "An analytical solution of vector diffraction for focusing optical systems with Seidal aberrations I. Spherical aberration, curvature of field, and distortion," J. Mod. Optics 40, 2293-2310 (1993).

94. F. M. Dickey and D. L. Shealy, *Laser Beam Shaping*, Optical Engineering (Marcel Dekker, Inc., New York, 2000).

95. M. Leutenegger, R. Rao, R. A. Leitgeh, and T. Lasser, "Fast focus field calculations," Opt. Express 14, 11277-11291 (2006).

96. J. S. Liu, A. J. Caley, and M. R. Taghizadeh, "Symmetrical iterative Fourier-transform algorithm using both phase and amplitude freedoms," Opt. Commun. 267, 347-355 (2006).

97. L. Qiu, W. Zhao, and X. Ding, "Effect of fabrication errors on axial super-resolution property of a three-zone pupil filter," Optik 117, 563-568 (2006).

98. W. Zhao, L. Qiu, and Z. Feng, "Effect of fabrication errors on superresolution property of a pupil filter," Opt. Express 14, 7024-7036 (2006).

99. S. K. Rhodes, K. A. Nugent, and A. Roberts, "Precision measurement of the electromagnetic fields in the focal region of a high-numerical-aperture lens using a tapered fiber probe," J. Opt. Soc. Am. A 19, 1689 - 1693 (2002).

100. T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," PNAS 97, 8206 - 8210 (2000).

101. R. Juškaitis and T. Wilson, "The measurement of the amplitude point spread function of microscope objective lenses," J. Microsc. 189, 8 - 11 (1997).

102. A. Marian, F. Charrière, T. Colomb, F. Montfort, J. Kühn, P. Marquet, and C. Depeursinge, "On the complex three-dimensional amplitude point spread function of lenses and microscope objectives: theoretical aspects, simulations and measurements by digital holography," J. Microsc. 225, 156 -169 (2007).

103. E. J. Botcherby, R. Juškaitis, M. J. Booth, and T. Wilson, "An optical technique for remote focusing in micorscopy," Opt. Commun. 281, 880 - 887 (2008).

104. T. R. Corle, C.-H. Chou, and G. S. Kino, "Depth response of confocal optical microscopes," Opt. Lett. 11, 770 - 772 (1986).

105. T. Wilson and A. R. Carlini, "Size of the detector in confocal imaging systems," Opt. Lett. 12, 227 - 229 (1987).

106. M. Gu and G. Zhou, "Spherical aberration and its compensation for high numerical aperture objectives," in *3D Laser Microfabrication*, H. Misawa and S. Juodkazis, eds. (Wiley-VCH, Weinheim, 2006), pp. 37 - 56.

107. G. M. Gallatin, "Analytic evaluation of the intensity point spread function," J. Vac. Sci. Technol. B 18, 3023 - 3028 (2000).

108. M. R. Beversluis, L. Novotny, and S. J. Stranick, "Programmable vector point-spread function engineering," Opt. Express 14, 2650 - 2656 (2006).

109. L. Dai, I. Gregor, I. v. d. H. Hocht, T. Ruckstuhl, and J. Enderlein, "Measuring large numerical apertures by imaging the angular distribution of radiation of fluorescing molecules," Opt. Express 13, 9409 - 9414 (2005).

110. T. Wilson and A. R. Carlini, "The effect of aberrations on the axial response of confocal imaging systems," J. Microsc. 154, 243 - 256 (1989).

111. L. Novotny, R. D. Grober, and K. Karrai, "Reflected image of a strongly focused spot," Opt. Lett. 26, 789 - 791 (2001).

112. J. L. Beverage, R. V. Shack, and M. R. Descour, "Measurement of the three-dimensional microscope point spread function using a Shack-Hartmann wavefront sensor," J. Microsc. 205, 61 - 75 (2002).

113. R. Juškaitis, "Characterizing high numerical aperture microscope objective lenses," in *Optical Imaging and Microscopy*, 2nd ed., P. Torok and F.-J. Kao, eds. (Springer, Berlin Heidelberg, 2007), pp. 21 - 43.

114. F. Charrière, A. Marian, T. Colomb, P. Marquet, and C. Depeursinge, "Amplitude point-spread function measurement of high-NA microscope objectives by digital holographic microscopy," Opt. Lett. 32, 2456 - 2458 (2007).

115. W. W. Macy, Jr., "Two-dimensional fringe-pattern analysis," Appl. Opt. 22, 3898 - 3901 (1983).

116. H. Schreiber and J. H. Bruning, "Phase shifting interferometry," in *Optical Shop Testing*, Third ed., D. Malacara, ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).

117. M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," J. Opt. Soc. Am. 72, 156 - 160 (1982).

118. M. J. Weber, *Handbook of Optical Materials*, Laser and Optical Science and Technology Series (CRC Press, Boca Raton London New York Washington D.C., 2003).

119. C. Roddier and F. Roddier, "Interferogram analysis using Fourier transform techniques," Appl. Opt. 26 1668 - 1673 (1987).

120. M. Kujawinska and J. Wójciak, "High accuracy Fourier transform finge pattern analysis," Opt. Laser Eng. 14, 325 - 339 (1991).

121. G. Frankowski, I. Stobbe, W. TIscher, and F. Schillke, "Investigation of surface shapes using carrier frequency based analysing systems," in *Interferometry*, (Proceedings of the SPIE, 1989), 89 - 100.

122. H. A. Bethe, "Theory of diffraction by small holes," Phys. Rev. 66, 163 - 182 (1944).

123. M. Dai and Y. Wang, "Fringe extrapolation technique based on Fourier transform for interferogram analysis," Opt. Lett. 34, 956 - 958 (2009).

124. J. H. Massig and J. Heppner, "Fringe-pattern analysis with high accuracy by use of Fourier-transform method: theory and experimental tests," Appl. Opt. 40, 2081 - 2088 (2001).

125. V. N. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils," Appl. Opt. 33, 8121 - 8124 (1994).

126. L. Kay, A. Podoleanu, M. Seeger, and C. J. Solomon, "A new approach to the measurement and analysis of impact craters," Int. J. Impact Engng 19, 739 - 753 (1997).

127. G. Szego, *Orthogonal Polynomials,* 4th ed., Colloquium Publications (American Mathematical Society, Providence, 1981).

128. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Selected Government Publications (Wiley, New York, 1972).