Estimation of Optical Trapping Forces with the Eikonal Function

Qin Yu

Supervisor: Prof Bryan Hennelly
Head of Department: Prof Gerard Lacey

A thesis submitted for the degree of Doctor of Philosophy

ELECTRONIC ENGINEERING DEPARTMENT
MAYNOOTH UNIVERSITY

February 2023

Abstract

An optical trap, also known as an optical tweezer, is a scientific instrument that uses a highly focused laser beam to trap and manipulate microscopic particles, such as cells or individual molecules, in three dimensions and has many applications in various fields of science, including biology, physics, and chemistry. The optical trap works by transferring the momentum of photons to the trapped particles, creating a force that can hold the particle in place. Ashkin developed a model to predict the forces in an optical trap in the Mie regime, i.e. for particles larger than the wavelength of the laser. His model is based on geometrical ray optics, whereby the rays from the microscope objective trace straight lines to a single focal point. These rays are refracted by the trapped particle and resultant change in direction of the ray imparts a force due to the conservation of momentum. The model takes no account of the effects of wave optics, which are highly pronounced in the focal region of a lens. The model cannot, therefore, account for the transfer orbital angular momentum from a Laguerre-Gaussian laser in a so-called optical spanner, or the effects of wavefront aberration caused by an imperfect microscope objective. In this thesis, we seek to bridge the gap between wave optics and Ashkin's model using the concept of the Eikonal function, which traces the flux lines through the focus. We call these flux lines non-linear rays, and these are used to replace the rays in Ashkin's model. This augmented model can account for the spin of a particle in an optical spanner and for the deleterious effect of aberration. In the journey towards this final result several significant contributions are made. These include the extension of Ashkin's model to account for absorption within the particle; we see a natural consequence for this is the emergence of a rotational force. Another important contribution are a set of algorithms to sample the three-dimensional diffraction in the focal region of a lens and to trace the flux lines with high accuracy.

Acknowledgements

"It has been a long journey to get to this point, full of challenges and rewards. It would not possible without the support of others. I would like to express my deepest appreciation to my supervisor, Prof. Bryan Hennelly, for his unwavering support, guidance, and patience throughout this research journey. His invaluable feedback, encouragement, and motivation have been instrumental in shaping my academic career. I extend my sincere gratitude to the students, interns, and staff of the Department of Electronic Engineering, who have provided administrative support and resources that have been essential to the success of this thesis project. Finally, I am indebted to my family and friends, who have stood by me through thick and thin, offering their unwavering love, encouragement, and support. Their sacrifices and belief in me have made this achievement possible.

Dedication

I hereby certify that the material contained within this thesis is entirely my own work, does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work

Signed: Qin Yu

ID Number:14183692

Date: February 2023

Contents

List of symbols and abbreviations

λ	wavelength
SLM	Spatial light modulator
OAM	Orbital Angular Momentum
MO	Microscope object
F_s	Scattering force
F_g	Gradient force
F_{mag}	Magnitude of force
r	Iadius
n_1	Iefractive index of surrounding medium
n_2	Refractive index of sphere
P	Input power
R	Reflect coefficient
T	Transmit coefficient
R_s	Reflect coefficient for s polarized
R_p	Reflect coefficient for p polarized
T_s	Transmit coefficient for s polarized
T_p	Transmit coefficient for p polarized
θ	Angle of incidence
r	Angle of refractive
c	Speed of light in vacume
HOT	Holographic optical tweezers
LG	Laguerre Gaussian
P	input power
Z_R	Rayleigh range
w_0	beam waist
l	Azimuthal index
p	Number of radial nodes)
BS	Beam splitter
M	Mirror
∇^2	Laplacian
j	$\sqrt{-1}$
k	$2\pi/\lambda$
δ	Dirac delta function
FT	Fourier Transform
DFT	Discrete Fourier Transform
IDFT	Inverse Discrete Fourier Transform
FFT	Fast Fourier Transform

$\forall$	For all
α	Scaled factor
A	Absorption coefficient
F_t	Rotational force
S/ρ Relative value of distance and radius	
BK7	Schott Borosilicate Crown Glass 7
BG3	Schott Glass BG3
$\Re$	Real part
ASM	Angular Spectrum method
RSM	Rayleigh Sommerfeld methd
ϕ	Phase delay
CCD	Charged coupled device
f	Focal length
NA	Numerical aperture
σ	Standard deviation
w	Weight

Chapter 1

Introduction

1.1 Optical Trapping

Light, as an electromagnetic wave, can have not only thermal effects but also mechanical effects. James Clerk Maxwell developed the theoretical basis for the phenomenon of radiation pressure [1, 2]. He proposed that when electromagnetic radiation, which carries momentum, is incident on a surface, the transfer of momentum must be conserved because the radiation is either reflected or absorbed. As a result of this transfer of momentum, a "pressure" is exerted on the surface in the direction of the incident radiation. The concept of radiation pressure was demonstrated in Nichols's experimental setup using a radiometer in 1901 [3]. However, the radiation pressure of ordinary light sources was too small, and the lack of controllability hindered further research. The invention of lasers in 1960 by Maiman greatly increased the power and controllability of light, making them suitable for a wide range of applications [4]. In 1970, Ashkin and his collaborators successfully used the radiation pressure of two laser beams transmitted relatively to each other to realize the trapping and capturing of latex particles [5]. Then in 1975, Ashkin and Dziedzic [6]. used a laser beam to achieve optical levitation and manipulation of charged or neutral droplets with diameters between 1μm and 40μm. In 1986, Ashkin et al successfully used a single laser beam with a highly uneven intensity distribution [7], which was focused by a microscope objective lens to achieve stable capture of tiny glass spheres. The capture mechanism of the focused laser beam on tiny particles stems from the balance of the two radiation forces generated when the laser beam interacts with the particles. After the laser beam is focused by an objective lens with a large numerical aperture, a micron-scale spot can be obtained, which can form a very high intensity. When this high-intensity and high-gradient beam interacts with particles of matching size, it can generate a so-called gradient force of pico-newton magnitude to trap high-refractive index particles at the maximum light intensity. On the other hand, due to reflection and absorption, scattering forces are also induced, which are usually along the propagation direction of the beam, similar to the Kepler light pressure. When the gradient force overcomes the scattering force, optical tweezers can achieve three-dimensional capture of particles. The theory of optical tweezers technology, in the Mie scattering regime, for particles

larger than the wavelength of the light such as biological cells for example, mainly involves using geometrical ray optics [7] and is based on the simple idea that when the rays of light propagating towards the focal point of the lens, are incident on the particle, they are refracted. The change in direction of these rays results in a transfer of momentum, which in turn leads to trapping forces. This simple theory, however, fails to account for any of the wave optical effects of light. Once simple example is that the orbital angular momentum that is carried by a TEM_{01} laser spatial mode (described using the Laguerre-Gaussian), and which is known to experimentally impart a spin on a trapped particle cannot be accounted for using the current model. Another example is the effect of lens aberration on an optical trap; aberrations are most accurtely described in terms of wave optics. The goal of this thesis is to bridge the gap between the current theory for predicting the forces on an optically trapped particle, and the effects governed by wave optics.

1.2 Rays and Waves

Ray optics is a branch of optics that deals with the behavior of light as it travels in a straight line through various media, without considering the wave nature of light. Ray optics is also called geometric optics [8, 9] because it uses geometrical principles to explain the behavior of light. The basic assumption of ray optics is that light travels in a straight line until it encounters a boundary between two media, at which point it is either reflected, refracted, or absorbed. In ray optics, light is treated as a collection of rays represented by arrows that indicate the direction of propagation of light. Ray optics explains many phenomena such as reflection, refraction, image formation by mirrors and lenses, and the behavior of optical fibers. The laws of reflection and refraction, also known as Snell's laws [8], are the basic principles of geometrical ray optics. These laws state that the angle of incidence is equal to the angle of reflection for a reflected ray, and that the ratio of the sine of the angle of incidence to the sine of the angle of refraction is constant for a refracted ray. Geometrical ray optics is widely used in many practical applications, including designing lenses, mirrors, and other optical components. It is also used in the study of image formation by optical systems, such as cameras and telescopes. For example, in a camera, the lens focuses light rays onto a film or digital sensor to form an image. In a telescope, a curved mirror or lens collects and focuses light from distant objects to produce an image.

Wave optics, on the other hand, is a branch of optics that deals with the wave nature of light [10]. Wave optics explains the behavior of light using the principles of wave propagation, interference, and diffraction. Unlike ray optics, wave optics takes into account the fact that light travels as a wave and can exhibit wave-like phenomena such as interference and diffraction [11, 12]. Wave optics explains many phenomena such as the interference and the diffraction of light. In wave optics, light is treated as a wave that travels through space and interacts with other waves to produce interference patterns. One of the most important concepts in wave optics is the wave equation, which describes the behavior of waves. The wave equation can be used to solve various prob-

lems in wave optics, including the propagation of light waves, the interference of light waves, and the diffraction of light waves. Another important concept in wave optics is the principle of superposition, which states that when two or more waves meet, their amplitudes add together. This principle is essential in understanding the interference of light waves, which is responsible for the formation of interference patterns in many optical systems. The two branches of optics, ray optics, and wave optics, are complementary to each other. Ray optics provides a simple and intuitive way to understand the behavior of light in many situations, while wave optics provides a more complete and accurate description of light, especially in situations where wave-like phenomena are significant. One area where the wave nature of light must be considered is at the focal point of a lens, precisely the location of an optical trap. Focused light does not in reality form a single infinitely narrow point through which all of the geometrical rays intersect. Instead the light forms a broad diffraction spot (or donut amplitude with spiral phase for the case of a Laguerre-Gaussian beam) at this point. The diffraction pattern is seen to change the shape non-linearly and rapidly as the wavefield approach the focal region of the lens and geomtrical rays cannot account for this effect.

The Eikonal function [8] provides a link between wave and ray optics. The Eikonal function describes the phase of a wavefront at each point in space and time and can be used to determine the path of a light ray through an optical system as a line that travels through paths of constant phase; these flux lines govern the direction in which the power of the wavefield travels. This is done by calculating the gradient of the Eikonal function, which gives the direction of the wavefront normal at each point. The direction of the wavefront normal is then the same as the direction of the light ray at that point. In this way, the Eikonal function provides a way to relate wave optics and ray optics. The Eikonal function is described in more detail in the section 2.5, where we pay particular attention to the relationship between the phase distribution of the wave optical description of light and the geometrical ray. Suffice it to say for now that when the phase of the wavefront varies slowly the geometrical ray and the flux line are one and the same; however, when the phase changes radpidly such as in the focal regions of a focussing Gaussian or Laguerre-Gaussian laser, the flux line follows a very different path to that of the geometrical ray. In Chapters 4 and 5 we develop a numerical metod that can accurately and quickly trace these lines of flux though a focussing laser and then in Chapter 6 we proceed to replace the rays in Ashkins model with the flux lines, enabing concepts like lens aberration and orbital angular momentum to be accounted for using Ahskins model.

1.3 Objective of this thesis

Here, we explicitly state the research question that this thesis seeks to address: ***Can wave optical effects such as wavefront aberration or orbital angular momentum be accounted for using Ashkin's ray-optics model to predict the forces on a particle in an optical trap?*** In order to answer this we were required to ask further questions:

1. Can absorption be included in Ashkin's ray-optics model to predict the forces in

a trap and can this result in a rotational force? This is answered in Chapter 3.

2. Can the three-dimensional wavefield be sampled quickly in three dimensions in the focal region of a lens, and can different laser modes and aberrations be included in this? This is answered in Chapter 3.

3. Can a method be developed that rapidly and accurately traces the flux lines through a focusing laser? This is answered in Chapter 4.

4. Can a computational method be developed to replace the rays with flux lines in Ashkin's model and can this account for the transfer orbital angular momentum from Laguerre-Gaussian lasers? This is answered in Chapter 6.

1.4 Thesis outline and contributions

In this section we provide an overview of the break down of this thesis in the context of the distinct contributions to the field of optics.

In Chapter 2, we provide a review of the background material and fundamental equations that underpin the contributions in the chapters that follow. There are no new contributions in this chapter. We review optical trapping in some detail including the various experimental systems such as holographic optical trapping and mirror trapping. We also provide a through description of the Eikonal function and its relationship to ray optics and wave optical propagation. This description serves as an introduction to the flux-tracing method proposed in Chapter 5.

In Chapter 3, we repeat the derivation of the forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime as first reported by Ashkin; in this case, however, we investigate the impact of partial absorption of light by the sphere on these trapping forces. Our model indicates that the effect of absorption is to reduce the axial and transverse trapping forces as a function of absorption, as well as changing the axial trap position. The trap stiffness is predicted to remain relatively unchanged even for high levels of absorption. The model also predicts that a rotational force is imparted on the sphere as a consequence of absorption. The model reduces exactly to Ashkin's model for the case of zero absorption. *This is the first time that Ashkins model has been shown to predict a rotational force.* The work in Chapter 3 has been prepared as a journal paper: **Qin Yu, Bryan Hennelly, "A new Spin on Ashkin's Laser Trapping Force in the Ray Optics Regime", To be submitted to Optics and Lasers in Engineering**

In Chapter 4, we present several contributions. Chapter 4 and 5 can be described as a two part series to develop a method for non-linear ray tracing, or more accurately,

to trace the lines of flux through a focusing laser. Regardless, there are several distinct contributions in Chapter 4 that are valuable to the optics community without any consideration of the flux.

1. We develop two novel algorithms for the calculation of the diffraction pattern at the focal point of a lens over a 3D grid of sampling points. The first algorithm, which relates to the case of the thin lens approximation, iteratively applies the Angular Spectrum Method (ASM, described in Chapter 2) , where the lens function is first 'pre-aliased' [13, 14] in order improve the computational efficiency. The second algorithm, which relates to the ideal lens involves the use of a single discrete Fourier transform to 'propagate' from the front focal plane of the lens to the back focal plane followed by an iterative application by the ASM.

2. We provide a careful derivation of the three-dimensional sampling requirements for these three-dimensional fields.

3. We extend these algorithms to include different laser spatial modes that are focused by the lens as well as different lens aberrations.

4. We provide for the first time a rigorous proof using fundamental sampling theory, of the superposition-tile method introduced by Kanka [13] and further developed by Kelly, [14] which involves adapting the input to the ASM in order to increase its computational efficiency. This involves breaking the input into smaller 'tiles', adding these together, and using the result as input to the ASM. We also examine the computational efficiency of the approach for the first time, which we find is only improved if the ASM is to be applied iteratively over several distances for the same input.

5. The Rayleigh-Sommerfeld aspheric thin lens function is introduced for the first time, which was found to produce superior focusing when compared to spherical and commercial aspheric lens functions.

6. The computational and memory load associated which each algorithm is quantified.

7. Chapter 4 is a useful reference for those who wish to simulate focused light over three-dimensions with lenses that have numerical aperture < 0.6 and can therefore be modelled using the paraxial approximation. It is particularly useful for those wishing to simulated focused laser beams by lenses with aberrations.

The work in Chapter 4 has been prepared for a submission to a journal with the following title: **Qin Yu, Bryan Hennelly, "Nonlinear Ray Tracing in Focused Fields, Part1: Calculating 3D Complex Wavefields", To be submitted to Applied Optics**.

In Chapter 5, we develop a novel algorithm to trace the flux lines through a the focal regions of a focussing laser. This builds on the contribution in Chapter 4 where we

developed a high-speed calculation of focused three-dimensional complex wavefields in the paraxial approximation for TEM_{00} and TEM_{01} laser modes and in the presence of various lens aberrations. The algorithms developed in the Chapter4 are first used to generate the three-dimensional grid of samples of the complex wavefield in the focal region. In Chapter 5, we focus on tracing a flux through this three-dimensional volume by first calculating the derivative of the phase (normal to the direction of propagation) throughout the three-dimensional volume, which is then used to direct a ray as it 'propagates' in a straight line between two consecutive planes within the volume. The origin of the ray can be chosen arbitrarily at any point and the ray can be then be traced through the volume with appropriate interpolation. We derive an equation for the error and demonstrate the conditions for which the algorithm will have high accuracy. Results are demonstrated for focused wavefields in the presence of aberrations, corresponding to the cases highlighted in the first paper. Some of the most interesting results relate to focused Laguerre-Gaussian beams, for which the rays are found to spiral at different rates of curvature, and for the cases of higher order aberrations such as Trefoil and Quadrafoil. The work in Chapter 4 has been prepared for a submission to a journal with the following title: **Qin Yu, Bryan Hennelly, "Nonlinear Ray Tracing in Focused, Part2: Tracing the flux", To be submitted to Applied Optics**.

In Chapter 6, we augment the extended Ashkin model developed in Chapter 3 to be used with the flux lines in a focusing laser. The core idea is to augment the classical ray-optics approach to calculate the forces acting on the sphere developed by Ashkin (and further developed by us in Chapter 3), such that non-linear rays can replace linear rays in the calculation. Perhaps the greatest advantage of such an approach would be be the capacity to model the orbital angular momentum imparted on the sphere using a Laguerre-Gaussian spatial mode laser. The flux model discussed in Chapter 5 is used to define the non-linear rays that intersect with the surface of the sphere and for each one of these rays, the scattering force, gradient force, and torque can be derived using the Equations defined in Chapter 3. Integration of these forces reveals the total three-dimensional force acting on the sphere as well as total rotational forces which can be decomposed into a 'vertical torque' and 'horizontal torque.' As well as investigating the single beam dielectric trap in the model of Ashkin, we additionally investigated the dual beam trap for all cases, which has the benefit of enhanced trapping forces.

Chapter 2

Background

2.1 Introduction

The aim of this chapter is to introduce and review the background theory and experimental procedures that underpin the contributions that are presented in all of the remaining chapters. In the main, this introductory chapter is broken into two large sections, which relate to the two main concepts that are used at the heart of thesis. The first is the background for optical trapping provided in Section 2.2 and the second is the background theory for the flux provided in Section2.5, which relates to the Eikonal function.

The background section on optical trapping begins with the seminal work by Ashkin and provides a brief overview of the concept and principles that underpin optical trapping. The ray optics description that was proposed by Ashkin is reviewed briefly; only the final equations for the scattering and gradient forces are provided here, since this will be presented in much more detail in Chapter 3 wit ha full derivation of these forces. For completeness these forces are also provided for the case of Rayleigh scattering where the articles are of size less tan the wavelength of the light. This is only covered briefly since the thesis focuses on Mie scattering. The fascinating topic of holographic tweezers is also reviewed, which enables traps to me moved dynamically using a spatial light modulator, which can also generate multiple dynamic traps. When the SLM displays a spiral phase, Laguerre-Gaussian modes can be generated which carry Orbital Angular Momentum, which can be transferred to the trapped particle causing it to spin. Ashkin's model for predicting the forces on an optical trap cannot account for the transfer of OAM to the particle and it precisely this gap in the theory that we attempt to fill in Chapter 6 of this thesis, which builds on the methods developed in Chapter 3-5. Single beam traps require very high numerical aperture microscope objectives. The review of optical trapping also includes the topic of dual beam traps, which require two opposing single beam traps and have significantly stronger trapping forces than single beam traps. Mirror traps use an SLM and a mirror near the image plane of the MO to produce the opposing traps and can move therefore can produce dynamic and stiff traps using low numerical aperture and low magnification objectives which have wide fields of view. Dual beam traps form an important part of our discussion later in

Chapter 6.

The second of the two large sections in this chapter relates to the flux, i.e. the movement of power within the three dimensional diffraction patterns as it moves through space. We hope to replace the use of 'rays' in Ashkin's model with the flux lines or 'non-linear rays'. This section begins with an overview of how the Angular Spectrum, an important concept in optical propagation, can be physically interpreted in terms of plane waves that span all of space that are uniquely defined by a wavevector propagating at some angle through space. This leads to a discussion of the Eikonal and the local wavevector, which naturally precipitates a discussion of the ray and the local spatial frequency. This forms a basis for a computational method to trace the flux lines, described in detail in Section 5.2. Included in this background is the definition of the Fourier transform, which is the basis of the Angular Spectrum, and its disceret counterpart, the discrete Fourier transform, which forms the basis for numerical propagation algorithms that appear in Chapters 4-6. Also included in this background chapter are brief discussion on phase unwrapping and aberration theory.

2.2 Optical Trapping

2.2.1 Background theory for optical trapping

In 1969, Ashkin observed that a focused laser could push particles with a size of several microns. He realized that transparent colloidal particles (diameter $0.6-2.5\mu m$) suspended in water aligned along the optical axis with a focused argon ion laser, could be pushed away from the focus and that particles approaching the beam were also unexpectedly drawn into the beam and pushed away. After repeated experiments with bubbles and droplets, Ashkin determined that the light beam had lateral attraction to particles with a higher refractive index than the surrounding medium but had lateral thrust to particles with a lower refractive index than the surrounding medium. In 1970, Ashkin et al [5]. first proposed the concept of manipulating tiny particles using optical pressure but it was not until 1986 that Ashkin and Dziedzic demonstrated the first stable optical trap [7]. This marked the birth of optical tweezers, whose official name is a single-beam optical gradient force trap. Ashkin won the Nobel Prize in Physics for this technology, which has stimulated more and more research interests internationally and promoted the emergence of new photo-silver technology and its innovative applications. As an example of the importance of optical trapping, one of the authors of the seminal 1986 paper, Steven Chu, would go on to use optical tweezing in his work on cooling and trapping neutral atoms [15], which would later earn Chu the 1997 Nobel Prize in Physics.

2.2.2 Basic Principles of Optical Trapping

The basic principle of Optical Trapping is to use a high numerical aperture objective lens to focus the laser beam and create an optical potential well (Optical Trap) to confine microscopic particles at the focal spot [7, 16, 17]. This is achieved by using

a focused laser beam that interacts with the particles, leading to two types of forces: scattering force and gradient force. The scattering force arises from photons pushing the object along the direction of beam propagation, while the gradient force is a directional force derived from the interaction between fluctuating dipoles in the object and the focused light field. The gradient force is proportional to the polarizability of the object and the gradient of the light field.

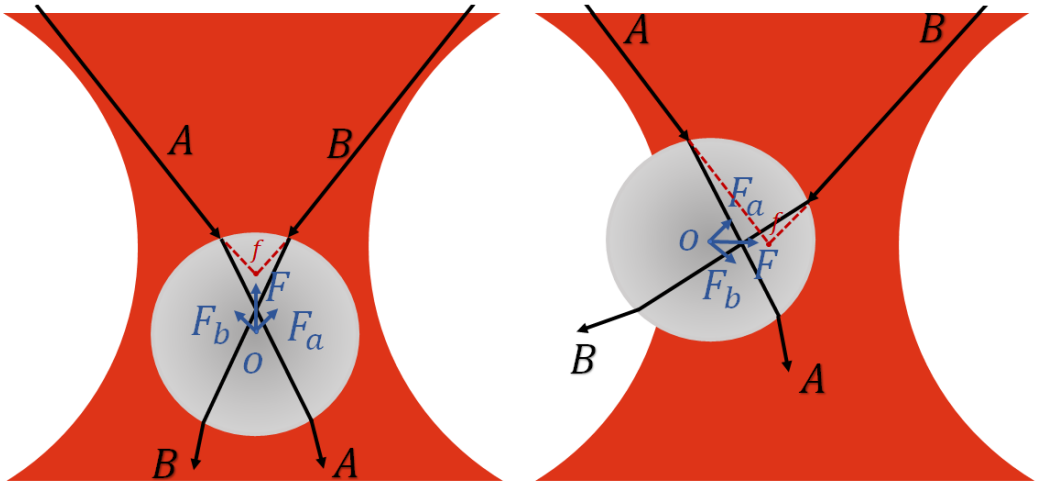

Figure 2.1: Illustration of optical trapping whereby some of the momentum of the light can be transferred to a particle via refraction. This can lead to pushing the particle away from the focus or pulling it towards, depending on angles of incidence and refraction; these angles are highly dependent on the ratio of the refractive index of the particle and the surrounding medium. In the Figure both cases are shown and in each case two rays A, and B, are shown to be refracted within the sphere. The different angles that these particles leave the sphere with respect to their original angles results in the transfer of momentum. The resultant forces F_a and F_b either push or pull the particle.

In optical trapping, the axial gradient force must be strong enough to overcome the scattering force that pushes particles away from the focal area, in order to achieve stable three-dimensional capture. This is why high numerical aperture objectives are used to focus the laser beam in the classical optical trap, as they provide steep light intensity gradients near the focal point, resulting in strong gradient forces that can overcome the scattering force. The use of holographic optical tweezers [18–20] and related technologies has further expanded the capabilities of optical trapping, enabling the manipulation of multiple particles in three dimensions. Holographic tweezers are introduced in Section 2.2.5. In order to remove confusion it is important to define the terms 'scattering force' and 'gradient force' clearly.

- **The Scattering Force,** F_s imparted by a ray acts in the same direction as the incident ray. It will have a component in all three dimensions.
- **The Gradient Force,** F_g imparted by a ray acts perpendicularly to the direction of the incident ray. It will also have a component in all three dimensions.

When the (axial) z-components of all the individual ray scattering and gradient forces added together equals zero, an optical trap is formed in the axial direction. The force

will rises with displacement from the trapping position. Similarly when the lateral x and y-components of all the the individual ray scattering and gradient forces added together equals zero, an optical trap is formed in the lateral direction direction.

2.2.3 The Ray Optics Description

Ray optics, also known as geometric optics, [8, 9] is a branch of optics that studies the behavior of light as it travels in a straight line (or ray) through transparent materials, such as lenses, mirrors, and prisms. It is based on the assumption that light travels in straight lines and interacts with these materials through reflection and refraction. In ray optics, light is represented as a ray, which is an imaginary straight line that represents the path of the light. The path of the ray can be determined using the laws of reflection and refraction. The law of reflection states that the angle of incidence of a ray of light is equal to the angle of reflection, while the law of refraction (Snell's law) states that the ratio of the sine of the angle of incidence to the sine of the angle of refraction is constant for a given pair of media.

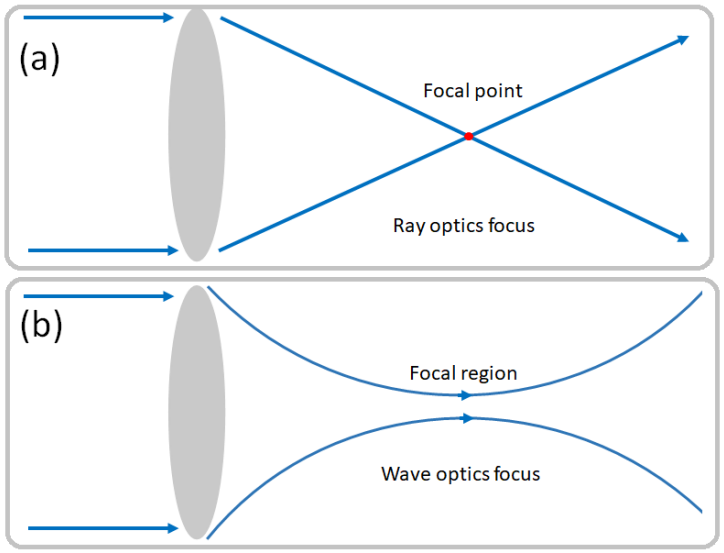

Figure 2.2: Propagation of ray focused by a lens. (a) geometrical interpretation where the rays travel in straight lines. This differs from wave optical description of light propagation which models the light as bending towards the focus (b)

Ray optics is used to predict the path of light through optical systems. For example, it can be used to determine the focal length of a lens, the magnification of an image, and the size and position of an image formed by a lens or mirror. Ray optics is also used in the design and optimization of optical systems in fields such as ophthalmology, microscopy, and telescope design. However, it is important to note that ray optics is an approximation of the behavior of light and only works in situations where the size of the optical elements is much larger than the wavelength of light. For smaller optical elements or in situations where diffraction plays a significant role, wave optics must be used instead. An illustration of the differences between ray optics and wave optics is shown in Fig. 2.2 (a). In the geometrical interpretation of a lens, the rays are

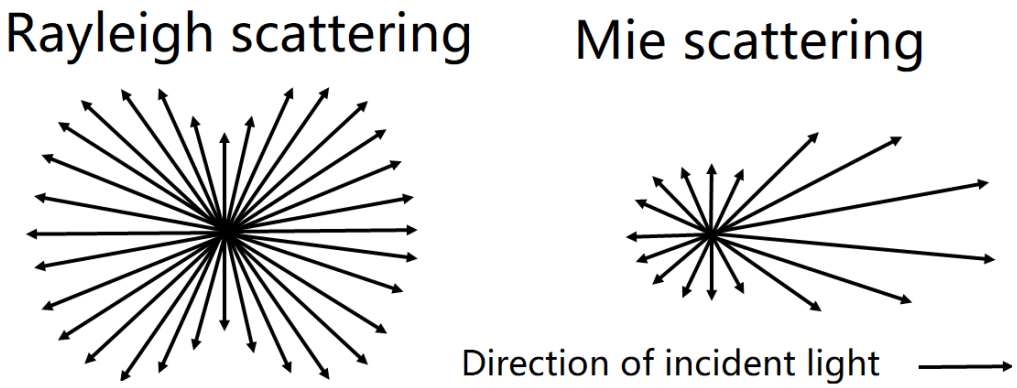

Figure 2.3: Rayleigh scattering and Mie scattering

seen to propagate as straight lines all converging at the focal spot. This model is used by Ashkin to derive the forces acting on larger (Mie Scattering) particles, described in detailed in Chapter 3. Ashkin's method [21] is based on tracing the straight line rays from the lens to the focus and using their angle of incidence on the particle to define the forces imparted by the rays on the trapped particle. In Chapter 6 we try to augment this description using a nonlinear ray model that is based on the bending of light as predicted by wave optics as illustrated in Fig. 2.2 (b). The angles of incidence of these rays on the particle are radically different and the resultant forces are also different.

For the case of Mie Scattering, [22] where the radius of the trapped particle is large ($r > 5\lambda$), the light beam can be regarded as a combination of many rays, each with a certain momentum, which are separately connected to the captured object. Reflection and refraction occur due to interactions, and the change in momentum during these processes can be analyzed to obtain the magnitude of the optical trapping force. This model is called the Geometric Optics Model. Figure 2.1 illustrates the geometric model of optical trapping. A converging laser beam acts on a transparent dielectric sphere, and depending on the relationship between the position of the ball and the focus position, different refraction processes may occur. Refraction of light means that the momentum it carries changes, and according to the law of conservation of momentum, the ball will receive a momentum of equal magnitude but opposite direction. The refractive index of the dielectric sphere is greater than that of the surrounding medium, the propagation of the rays is shown in Figure 2.1. A detailed review of Ashkin's ray-optics model [21] that he used to predict the forces acting on a trapped particle is provided in Chapter 3, which is followed by an extension of his work that accounts for particles that absorb light. For now we simply provide his result for the scattering and gradient forces.

$$F_s = \frac{n_1 P}{c}\left[1 - R\cos(\pi + 2\theta) - \sum_{n=0}^{\infty} T^2 R^n \cos(2(\theta - r) + n(\pi - 2r))\right] \tag{2.1}$$

$$F_g = \frac{n_1 P}{c}\left[0 - R\sin(\pi + 2\theta) - \sum_{n=0}^{\infty} T^2 R^n \sin(2(\theta - r) + n(\pi - 2r))\right] \tag{2.2}$$

where θ and r are the angles of incidence and refraction, R and T are the reflection and

transmission coefficient of the sphere, P is the input power, n_1 is the refractive index of the surrounding medium of the particle. R and T, the reflection and transmission coefficients are defined by the Fresnel equations:

$$\begin{aligned} R_s &= \frac{n_1\cos\theta_1 - n_2\cos\theta_2}{n_1\cos\theta_1 + n_2\cos\theta_2} \\ R_p &= \frac{n_2\cos\theta_1 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1} \\ T_s &= \frac{2n_1\cos\theta_1}{n_1\cos\theta_1 + n_2\cos\theta_2} \\ T_p &= \frac{2n_1\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1} \end{aligned} \tag{2.3}$$

where R_s and R_p are the reflection coefficients for s-polarized and p-polarized light, T_s and T_p are the transmission coefficients, and θ_1 and θ_2 are the angles of incidence and refraction, respectively. It is clear that the scattering and gradient forces are dependent on the angle of incidence θ between the ray and the surface normal. This angle, together with the refractive index of the sphere and the surrounding medium determines the refraction angle of the ray. It is this process of refraction that governs how much momentum is transferred to the particle and by consequence, how much force will act on it. Snell's law describes the relationship between the angle of incidence and the angle of refraction:

$$n_1\sin\theta_1 = n_2\sin\theta 2 \tag{2.4}$$

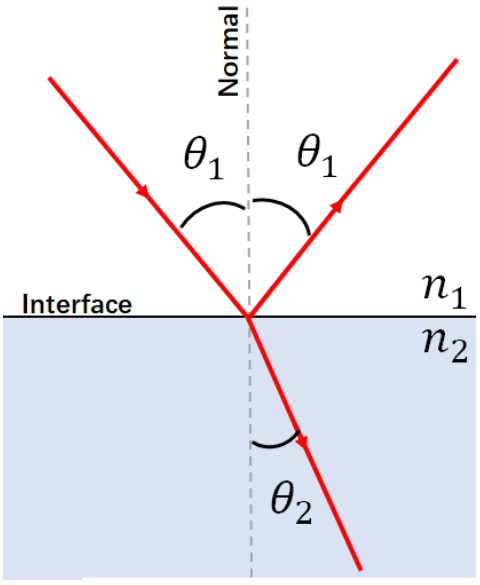

Figure 2.4: Illustration of refraction, which is the basis for the transfer of momentum that occurs in an optical trap.

2.2.4 The Rayleigh scattering description

When the particle size is much smaller than the laser wavelength, it is called a Rayleigh particle. Rayleigh particles satisfy the Rayleigh scattering theory, and they

can be approximated as point dipoles [23]. Under the approximation of radiation theory, the scattering force and the gradient force can be calculated using the following formulas respectively. The scattering power and light intensity are proportional to the particle size. It can be represented as [24],

$$F_s = \frac{I_0 n}{c} \frac{128\pi^5 a^6}{3\lambda^4} \left(\frac{m^2 - 1}{m^2 + 2} \right)^2 \tag{2.5}$$

Among them, I_0 is the light intensity incident on the particle, a is the radius of the particle, c is the speed of light in vacuum, n is the refractive index of the medium around the particle, λ is the wavelength of light, and m is the ratio of the refractive index of the particle to the surrounding medium. The gradient force is proportional to the light intensity gradient which can be expressed as follows:

$$F_g = \frac{2\pi n a^3}{c} \left(\frac{m^2 - 1}{m^2 + 2} \right)^2 \nabla I_0 \tag{2.6}$$

As for the case of the ray-optics description, when the gradient force and the scattering force are balanced, stable capture can be achieved. Rayleigh particles satisfy the Rayleigh scattering theory, and they can be calculated as point dipoles. The scattering force is caused by the absorption and re-radiation of light between dipoles, while the gradient force is caused by the time averaged interaction between the dipole and the light fields. The Rayleigh scattering description of optical trapping is not discussed any further in this thesis, and is included here only for the sake of completeness. All of the particles that will be investigated in this thesis have a large radius relative to the wavelength of the light.

2.2.5 Holographic Optical Tweezers

Holographic optical tweezers (HOTs) [19, 20] are a type of optical trapping technology that uses a computer-generated hologram to create multiple optical traps in three-dimensional space. This technique was first developed in the late 1990s by a group led by Eric Dufresne and David Grier. The basic principle of HOTs is to use a spatial light modulator (SLM) to modulate the phase and amplitude of a laser beam. The SLM acts as a programmable diffraction grating, allowing the creation of complex holographic patterns that can be used to create multiple optical traps. By controlling the hologram displayed on the SLM, it is possible to create an arbitrary number of optical traps in three dimensions. HOTs have several advantages over traditional optical tweezers. Firstly, they can create many traps at once, allowing for the simultaneous manipulation of multiple particles. Secondly, the traps can be moved and repositioned by changing the hologram displayed on the SLM, which allows for precise control over the position and orientation of the trapped particles. Finally, HOTs can be used to create more complex trapping geometries, such as arrays of traps, toroidal traps, and more. There have been many applications of HOTs in the fields of physics, biology, and materials science. For example, HOTs have been used to study the behavior of microorganisms [25–27],

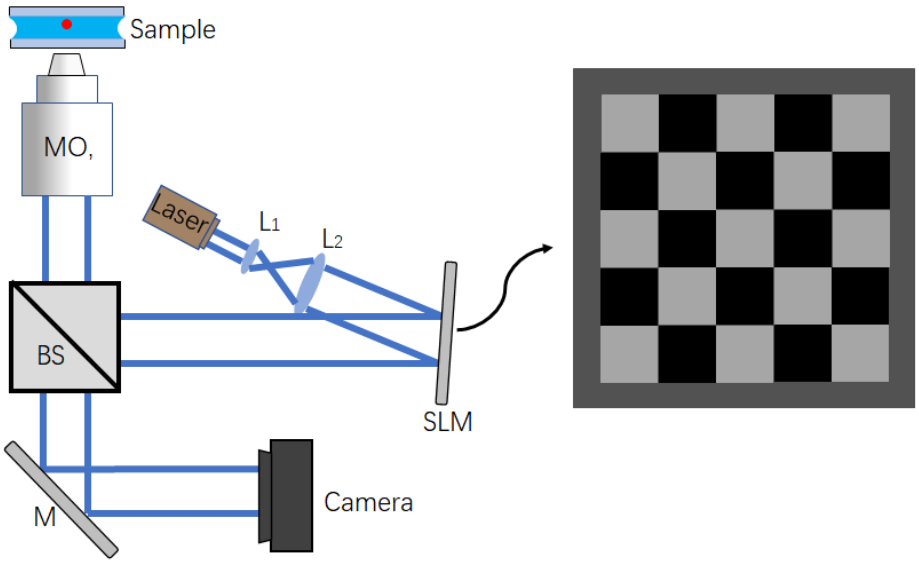

Figure 2.5: Holographic optical tweezers. A spatial light modulator (SLM) can be used to generate multiple dynamic traps in the 3D volume around the focus of the laser.

One of the challenges of HOTs is the need for precise calibration of the optical system, including the SLM and the objective lens. This calibration is necessary to ensure that the holographic traps are accurately positioned and have the correct intensity. In addition, the high intensity of the laser beam can cause heating and damage to the trapped particles, which can limit the use of HOTs for certain applications. Nevertheless, holographic optical tweezers are a powerful tool for the manipulation and study of particles in three dimensions. With their ability to create multiple, complex traps, they have the potential to revolutionize many areas of research and technology.

A HOT set up is illustrated in Fig. 2.5 The SLM is imaged to the back aperture of the microscope objective, such that the relationship between the plane of the SLM and the focal plane of the MO is a Fourier transform (discussed in Section 2.4.2 below). An SLM is an electronic device that can be used to modify the properties of a beam of light, such as its intensity, phase, or polarization, in a spatially dependent manner [28]. In other words, it can manipulate the light in different parts of the beam differently. The SLM typically consists of an array of small pixels as illustrated in Fig. 2.5, which can be controlled independently and at high frame rate to create a pattern that modulates the light passing through the device. This pattern can be used to produce complex optical functions. There are different types of SLMs, including liquid crystal displays (LCDs) and a variant known as Liquid Crystal on Silcon (LCOS) [29], microelectromechanical systems (MEMS) [30], and digital micromirror devices (DMDs) [31]. Each type has its own advantages and limitations, and is suited for different applications but LCOS is the preferred technology for holographic tweezers.

In order to imparts a physical translation of the optical trap by amount Δ_x and Δ_y in the lateral dimensions, a pattern must be displayed on the SLM that imparts the followin linear phase shift/delay:

$$\phi(np_x, mp_y) = \frac{2\pi}{\lambda}(np_x\Delta_x + mp_y\Delta_y) \tag{2.7}$$

where ϕ is the phase delay, n and m are the indices of the pixel array over the two dimensions and λ is the wavelength of the light. To further impart a movement of the trap in the axial direction by an amount Δ_z, this phase delay must have the following form:

$$\phi(np_x, mp_y) = \frac{2\pi}{\lambda\Delta_z}[(np_x - \Delta_x)^2 + (mp_y - \Delta_y)^2] \tag{2.8}$$

To generate multiple traps at arbitrary different positions denoted by index i, the pattern is calculated by adding the complex distribution for each individual trap and taking the phase of the result:

$$\phi(np_x, mp_y) = \angle\left[\sum_i \exp\left[\phi_i(np_x, mp_y)\right]\right] \tag{2.9}$$

In addition, the phase relating to the aberration on the MO can also be removed by including its conjugate of the pattern that is displayed on the SLM. This phase term would be defined using the Zernike polynomials described below and would simply be added to the phase patterns described above.

2.2.6 Optical Spanners and Laguerre-Gaussian Beams

An optical spanner [19, 32–34] is a tool that uses light to apply torque to small objects, such as particles, cells, and biological molecules. It is a type of optical trapping technique that is similar to optical tweezers, but instead of trapping particles in a static position, an optical spanner applies a rotational force to the trapped particle, causing it to spin, via the transfer of orbital angular momentum (OAM) to the particle. An optical spanner typically uses a laser beam with a specific intensity distribution and polarization, known as a Laguerre-Gaussian (LG) beam, to apply torque to the trapped particle. The LG beam creates a structured optical field that exerts the rotational force on the particle, causing it to spin. By controlling the parameters of the LG beam, such as its beam waist and phase structure, one can control the speed and direction of the particle's rotation. To create an optical spanner, we simply display the phase patterns associated with a Laguerre-Gaussian distribution on an SLM. The complex amplitude distribution of the Laguerre-Gaussian mode is uniquely defined by the beam waist width $w(0)$ and the Rayleigh range, z_R, as well as integers l (the azimuthal index) and p (the number of radial nodes) as follows: [8]

$$\begin{aligned} A_{LG_{pl}}(x,y) = &\sqrt{\frac{2p!}{\pi(p+|l|)!}}\frac{1}{w(z)}\left[\frac{\sqrt{2(x^2+y^2)}}{w^2(z)}\right]^{|l|}\exp\left[\frac{-(x^2+y^2)}{w^2(z)}\right] \\ &\times L_p^{|l|}\left[\frac{2(x^2+y^2)}{w^2(z)}\right]\exp\left[jl\tan^{-1}\left(\frac{y}{x}\right)\right]\exp\left[\frac{ikz(x^2+y^2)}{2(z^2+z_R^2)}\right] \\ &\times \exp\left[-j(2p+|l|+1)\tan^{-1}\left(\frac{z}{z_R}\right)\right] \end{aligned} \tag{2.10}$$

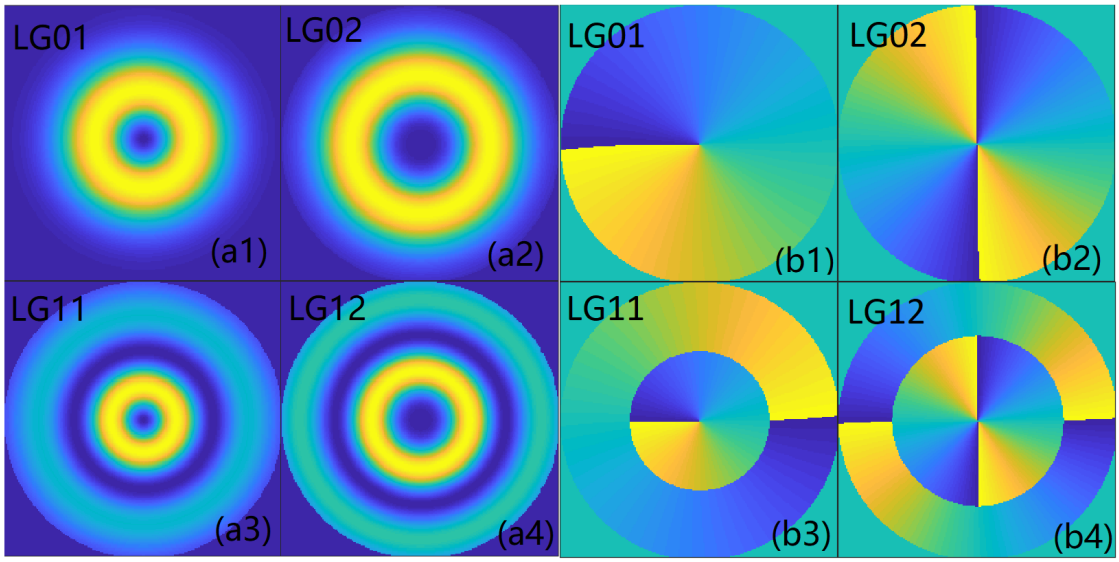

Figure 2.6: Amplitude (left) and phase (right) of LG01, LG02, LG11 and LG12

where the $1/e$ width of the Gaussian term is given by:

$$w(z) = w(0)\sqrt{\frac{z^2 + z_R^2}{z_R^2}} \tag{2.11}$$

The argument of the last exponential term in Equation 2.10 is the Guoy phase. The term $L_p^{|l|}$ is the associated Laguerre-Gaussian polynomial obtained from Laguerre polynomials:

$$L_p^{|l|}[a] = (-1)^{|l|}\frac{d^{|l|}}{da^{|l|}}L_{p+|l|}[a] \tag{2.12}$$

We note that for $l = p = 0$ and for the case of negligible beam expansion $z_R \to \infty$ the above expressions reduces to the form of the Gaussian mode defined in Equation 2.11 where $w(0) = 2\sqrt{2}\sigma$. The amplitude and phase distributions associated with several different L-G modes are shown in Fig. 2.6.

Optical spanners have a wide range of applications in fields such as biophysics, materials science, and cell biology, where they can be used to study the mechanical properties of small objects and to manipulate biological structures. For example, optical spanners have been used to study the mechanical properties of individual cells and to investigate the rotation of cilia and flagella in biological organisms. A nice review of the applications of optical spanners can be found in Refs 22,35. Overall, optical spanners provide a non-invasive way to apply rotational forces to small objects, and they are a powerful tool for investigating the mechanical properties of biological and materials systems. It should also be mentioned that the use of light beams carrying orbital angular momentum can also enable free-space information transfer with high bandwidth and security [36] but this is outside the scope of optical trapping, which is covered in this thesis. Optical spanners are of particular interest in this thesis and we attempt to model the spinning forces associated with these spanners using Ashkin's ray-optics model.

2.3 Dual Beam and Mirror Traps

A dual-beam optical trap, [16, 22] which is is illustrated in Fig. 2.7 also known as a two-beam optical trap or counter-propagating optical trap, is a type of optical trap used to manipulate microscopic objects. The trap consists of two opposing laser beams, each focused by a microscope objective lens. The laser beams are typically of equal intensity and frequency, and they propagate in opposite directions, intersecting at the focal point between the two objective lenses. The light exerts a force on a microscopic object in the focal point of the trap, and the opposing beams generate a stable trap that can hold the object in place. Like the single beam optical trap, the dual-beam optical trap can be used to manipulate microscopic objects in various ways, including trapping, positioning, and rotating.

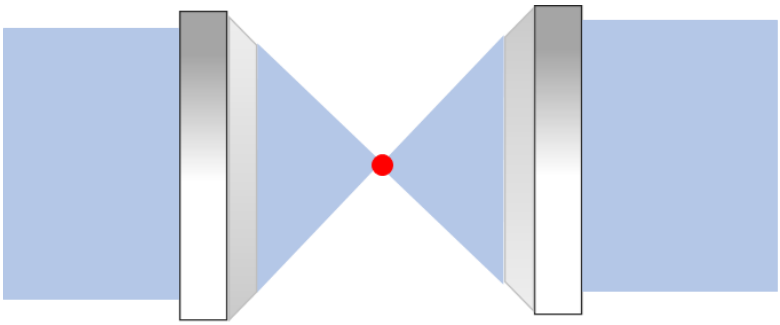

Figure 2.7: Illustration of dual beam laser trap. Two opposing single beam traps are strengthened due to the cancellation of their respective scattering forces.

Dual-beam optical traps offer several advantages over single-beam optical traps for manipulating microscopic objects. Firstly, they have improved trapping stability and can generate a more stable trap than single-beam traps, which helps to hold the trapped object more securely in place. This is because the opposing beams effectively cancel each others' scattering forces. The result is that dual-beam optical traps can generate a stronger trapping force than single-beam traps, which allows for the trapping and manipulation of larger or heavier particles. An additional benefit of dual-beam optical traps is that they offer more precise control over the position and movement of the trapped object than single-beam traps. This is because the two beams can be independently adjusted to control the direction and strength of the trapping force, allowing for more precise positioning and manipulation of the object. There is also the advantage of reduced heating effects due to the lower intensity of the laser beam, which can cause thermal damage to the trapped object. Overall, dual-beam optical traps offer several advantages over single-beam traps for manipulating microscopic objects, including improved trapping stability, enhanced trapping strength, improved manipulation precision, and reduced heating effects. These advantages make dual-beam traps a useful tool for a wide range of applications in science and engineering, including in biological research to study the mechanical properties of cells and to manipulate microorganisms and other biological samples [37, 38]. They have also been used in materials science and nanotechnology to manipulate and assemble nanopar-

ticles and other microscale structures [39].

One of the the most famous examples of multiple beam traps is the application to capture and cool single atoms together [15], which won the 1997 Nobel Prize in Physics for Chu, who was also one of the co-authors of Ashkins seminal paper in 1986 [7].

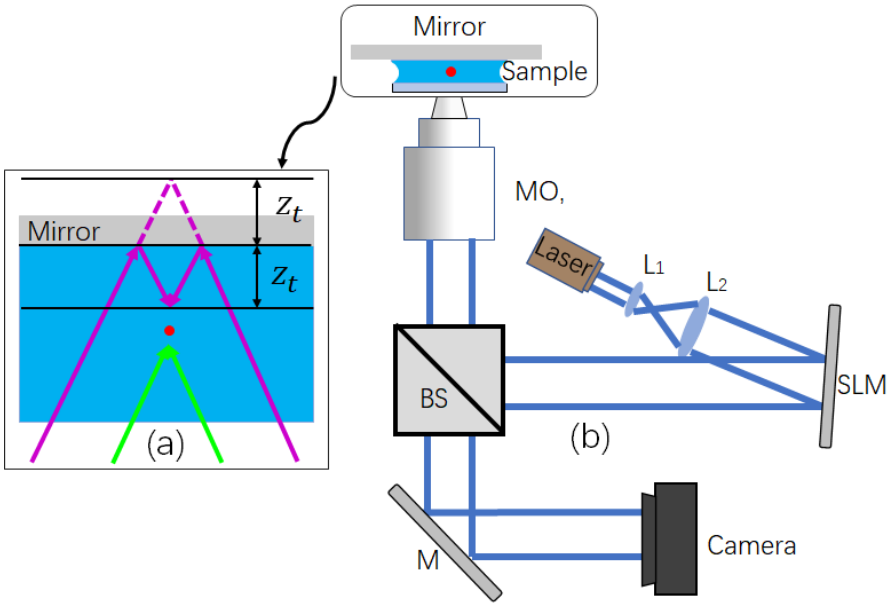

Figure 2.8: Set up for the dual beam mirror trap whereby a single beam is incident on an SLM, which produced two independent beams, which come to focus a slightly different distances. A mirror reflects one of these back such that both beams focus to the same point and are counter propagating. These set-up have sometimes been named macro-tweezers owing to the wide field of view that is afforded by the lower numerical aperture objective that can be used to achieve trapping.

While the use of an SLM to control the trap position is difficult for the case of a dual beam trap (since this would require two independent and synchronously controlled SLMs), a more recent evolution of the dual-beam approach has been proposed that uses only a single MO. The so called mirror trap, [40–42] is a variation of the optical tweezers technology that uses an SLM pattern to split a laser beam into two trapping beams. The basic setup of such a trap is illustrated in Fig. 2.8 and consists of a laser beam that is expanded onto an SLM; a pattern is displayed on the SLM to split the single laser beam into two beams with controlled phase and intensity profiles. One of these beams is unchanged with respect to the original laser, while the second is expanded slightly such that it will come to focus several hundred micrometres after the first. A mirror is placed behind the sample plane, which reflects the second beam back towards the sample. Both beams will come to focus at the same point in space but both are counterpropagating, which creates the effect of the dual-beam trap but only using a single MO. By carefully controlling the phase and intensity of the two beams, a stable trap can be created that holds the sample in place. One of the main advantages of using an SLM, is that dymaic holographic tweezers can be generated as described above. By manipulating the phase and intensity profiles of the two beams, it is possible to create a wide range of trapping patterns, including multiple traps, optical spanners,

and more complex 3D trapping patterns. This makes it possible to manipulate more complex structures or to perform more precise manipulations of smaller structures. Another advantage of using an SLM-based mirror optical trap is the ability to dynamically adjust the trap configuration. By changing the phase and intensity profiles of the beams using the SLM, the trap can be repositioned or modified as needed, allowing for more precise control over the sample. Dual traps and mirror traps are of particular interest in Chapter 6, where we apply our newly proposed flux-model to predict the forces acting on the trapped particle.

2.3.1 Applications of laser trapping

Optical trapping, of one variant or another, has a wide range of applications in fields such as biophysics, cell biology, materials science, and optics, due to its ability to trap and manipulate objects with high precision and without causing any damage. Although we do not research any specific applications of optical trapping in this thesis, it is important to provide an overview of the potential applications as an overarching justification for the work in this thesis. We have attempted to review the applications under a number of classification in an effort to simplify the discussion. In some cases there is overlap in the citations.

- **Force measurement**: This is perhaps the most obvious application. Optical tweezers can measure the force exerted by a single molecule, allowing researchers to study the mechanical properties of biological molecules such as DNA and proteins. For example, optical tweezers have been used to study the elasticity of single DNA molecules and the mechanical properties of proteins such as myosin, kinesin, and dynein [22, 43, 44].

- **Microrheology**: Optical tweezers can be used to measure the viscoelastic properties of materials, such as biological fluids and gels, at the microscale. By measuring the motion of a trapped particle in response to an external force, researchers can determine the material's viscoelastic moduli. Optical tweezers have been used to study the viscoelastic properties of biological materials such as cytoskeletons and extracellular matrices [45, 46]

- **Cell manipulation**: Optical tweezers can be used to manipulate living cells and study their behavior. Researchers can use optical tweezers to hold and move cells, measure their mechanical properties, and study their response to external stimuli. Optical tweezers have been used to study cell division, migration, and signaling [37, 47, 48].

- **Single-molecule manipulation, in particular for studying DNA and proteins** This is arguably the most important application and likely the most cited. Optical tweezers can be used to manipulate individual molecules and study their interactions. By attaching a single molecule to a trapped particle, researchers can measure the molecule's mechanical properties and study its interactions with other molecules. Optical tweezers have been used to study the interactions of

DNA-binding proteins, RNA polymerases, and ribosomes with DNA and RNA molecules [44, 49–52].

- **Optogenetics**: Optogenetics is a technique that uses light to control the activity of cells genetically modified to express light-sensitive proteins. Optical tweezers can be used to stimulate or inhibit the activity of individual cells by manipulating the position. Optical tweezers have been used to control the activity of neurons, muscle cells, and other cells expressing light-sensitive proteins [53].

- **Cell sorting**: Optical tweezers can be used to sort cells based on their physical properties, such as size, shape, and mechanical properties. By trapping cells in a laser beam and moving them to different locations, researchers can separate cells based on their properties. Optical tweezers have been used to sort cells based on their mechanical properties, such as stiffness, which can be an indicator of disease or other cellular abnormalities [54, 55].

- **Biological imaging** Optical tweezers can be used to manipulate cells and other biological samples during imaging experiments. By holding cells in place or moving them to different locations, researchers can image specific regions of cells or study the dynamics of cellular processes. Optical tweezers have been used in combination with fluorescence microscopy, confocal microscopy, and super-resolution microscopy to image cells and cellular components [7, 56].

- **Optical stretching**: Optical stretching is a technique that uses optical tweezers to stretch cells and measure their mechanical properties. By trapping a cell with two laser beams and moving them apart, researchers can stretch the cell and measure its response to the applied force. Optical stretching has been used to measure the stiffness of cells and study the effects of drugs and other compounds on cell mechanics [57].

- **Neuroscience**: Optical tweezers have been used in neuroscience to study the mechanical properties of neurons, the mechanics of neuronal growth and differentiation, and to measure the force exerted by neuronal growth cones [58, 59].

- **Chemical and material science**: Optical tweezers have been used in chemical and material science to study the properties of polymers, colloids, and other soft materials [60].

- **Cooling atoms** The cooling of atoms using optical traps is an important area of research in physics and related fields, because it allows researchers to study the behavior of atoms at very low temperatures, where quantum mechanical effects become more pronounced. Cooling is achieved using very strong optical traps which prevent motion and thereby cool. Laser-cooled atoms can be used for a wide range of applications, such as high-precision measurements, quantum computing, and the study of fundamental physics [15].

  Overall, optical trapping is a versatile and powerful technique that has a wide range of applications in many different fields. Its ability to trap and manipulate

small objects with high precision and without causing any damage has made it an important tool for a variety of scientific and technological applications.

2.4 Background theory for optical/numerical propagation

Numerical propagation plays an important role in this thesis, particularly in Chapters 4 and 5. Here we review some of the most important concepts that are particularly relevant in those chapters. These concepts will be expanded upon further in Chapters 4 and 5.

2.4.1 The wave nature of light

We begin with a brief description of the wave nature of light. In 1801, Thomas Young performed the classical double-slit experiment. It could be observed from his experiment that bright and dark fringes will appear after monochromatic light passes through the double slits, revealing the interference phenomenon, and also suggesting that the light can be described as a wave. Then, in 1861, James Clerk Maxwell described light as an electromagnetic wave, [10] and in 1873, he proposed the very well-known Maxwell's Equations to explain the behaviors and the relationship of electric field and magnetic field [2]. Based on the description of the Maxwell's Equation, if $\vec{E}$ is denoted as the vector of the electric field of light at the time t, it must satisfy the wave equation given by:

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{2.13}$$

where the ∇^2 denotes the Laplacian operator and c is the light speed in the vacuum. For linear polarized light (which means that the orientations of electric field vector and magnetic field vector are unchanged and perpendicular to each other, and both vectors are perpendicular to the light propagation direction z; and the polarization direction of a linear polarized light is equal to the orientation of electric field vector), then Equation 2.13 can be modified for the harmonic and linear polarized wave in the electric field:

$$\frac{\partial^2 \vec{E}}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{2.14}$$

The propagating electric field vector $\vec{E}$ at the time t at the position with the spatial vector $\vec{r} = (x, y, z)$ can be written as a complex function as follows:

$$E(x, y, z; t) = A \exp\left[j(\vec{k}\vec{r} - wt) \right] \tag{2.15}$$

where A is the amplitude of wave, j is the unit imaginary number equal to $\sqrt{-1}$, $|\vec{k}|$ is the wavenumber equal to $\frac{2\pi}{\lambda}$ (λ is the light wavelength), and $\vec{k}\vec{r} - wt$ is the phase of wave. The leads to the Helmholtz equation, which is a special case of the wave equation in which the time dependence of the wave is assumed to be sinusoidal and the speed of

propagation is equal to a constant, which is related to the frequency of the wave.

$$\nabla^2 E(\vec{r}, t) + k^2 E(\vec{r}, t) = 0 \tag{2.16}$$

The Helmholtz Equation is the fundamental equation for the propagation of mononchromatic light in free space and is the basis for the propagation of the Angular Spectrum, a very useful description of propagation, which is discussed below. The Angular Spectrum is used in Chapters 4, 5, and 6 as we attempt to trace the lines of flux in a focusing laser and to use these lines to predict the forces on an optically trapped particle. Before discussing the Angular Spectrum of a wavefield, we must first introduce the Fourier transform.

2.4.2 The Fourier transform in Optics

The Fourier transform [61] is a common and effective mathematical method, that can convert a signal in the time/space domain into a signal in the frequency domain. The Fourier transform has many applications in signal processing, including light wavefield analysis and digital image processing, which are required in this thesis. In the study of Fourier optics, [28] it is assumed that a light wavefield is composed by the superposition of a series of plane waves. This assumption is a little similar to the Huygens-Fresnel principle, [8, 28] but the difference is that the Huygens-Fresnel principle assumes that the light wavefront is made up of the superposition of a series of spherical waves. We first define a complex wavefield function of light $g(x, y)$ in the space domain that contains an amplitude field $A(x, y)$ and a phase field $\phi(x, y)$:

$$g(x, y) = A(x, y) \exp(j\phi(x, y)) \tag{2.17}$$

The 2D Fourier transform of $g(x, y)$ is written as follows:

$$G(k_x, k_y) = \iint_{-\infty}^{\infty} g(x, y) \exp[-j2\pi(xk_x + yk_y)] \mathrm{d}x \mathrm{d}y \tag{2.18}$$

where $G(k_x, k_y)$ is the spectrum in the spatial frequency domain, and (k_x, k_y) is the spatial frequency coordinate that is associated with the propagation direction of plane wave $\exp[-j2\pi(xk_x + yk_y)]$ in the space domain. The wavefield, $g(x, y)$, could also be recovered from its spectrum, $G(k_x, k_y)$, using the inverse Fourier transform:

$$g(x, y) = \iint_{-\infty}^{\infty} G(x, y) \exp[j2\pi(xk_x + yk_y)] \, \mathrm{d}k_x \, \mathrm{d}k_y \tag{2.19}$$

The frequency spectrum $G(k_x, k_y)$ can be considered as the weights for the different plane waves $\exp[j2\pi(xk_x + yk_y)]$. Thereby, the light wavefield $g(x, y)$ is equivalent to the superposition of these multiple plane waves multiplying their weights respectively. In practical simulations, it is convenient to represent the signals using a discrete set of numbers that can be stored in computer memory and subject to numerical computation. The continuous 2D signal in x and y will be 'discretized' into a 2D array of complex numbers by the process of sampling [61]. The resultant digital (complex) im-

age is the starting point for many of the algorithms developed in Chapters 4-6 that use numerical processing. One key element that is used is the Discrete Fourier Transform (DFT). Assuming that the digital image consists of $N_x \times N_y$ pixels (we use the term pixels and sampling intervals interchangeably) in two dimensions, each pixel location could be considered as a spatial coordinate, and the continuous Fourier transform provided in Equation 2.18 reduces to the DFT as follows:

$$G(m_x\Delta_x, m_y\Delta_y) = \sum_{-\frac{N_x}{2}}^{\frac{N_x}{2}-1} \sum_{-\frac{N_y}{2}}^{\frac{N_y}{2}-1} g(n_x\delta_x, n_y\delta_y) \exp[-j2\pi(\frac{n_x m_x}{N_x} + \frac{n_y m_y}{N_y})] \tag{2.20}$$

Different from the notation used for the continuous variables x, y, k_x, k_y in Equation 2.18, the new notation in Equation 2.20 is obtained using the following substitutions:

$$\begin{aligned} x &\Rightarrow n_x\delta_x, \quad n_x \in (-N_x/2, N_x/2-1) \\ y &\Rightarrow n_y\delta_y, \quad n_y \in (-N_y/2, N_y/2-1) \\ k_x &\Rightarrow m_x\Delta_x, \quad m_x \in (-N_x/2, N_x/2-1) \\ k_y &\Rightarrow m_y\Delta_y, \quad m_y \in (-N_y/2, N_y/2-1) \end{aligned}$$

where δ_x, δ_x are the sampling intervals used in the 2 dimensions of the spatial domain and are equal to the pixel width. Δ_x, Δ_y are the sampling intervals used in the 2 dimensions of the spatial frequency domain. Assuming the region in spatial domain is limited by the $(N_x\delta_x, N_y\delta_y)$, then the corresponding region in the spatial frequency domain is $\left(\frac{1}{\delta_x}, \frac{1}{\delta_y}\right)$, and the sampling intervals Δ_x, Δ_y in the spatial frequency domain are equal to $\left(\frac{1}{N_x\delta_x}, \frac{1}{N_y\delta_y}\right)$. The sampling intervals δ_x, δ_x must be chosen to be sufficiently small to satisfy the Nyquist condition and therefore to ensure that aliasing will not occur and accurate interpolation of the continuous signal can be achieved from the discrete samples. This is discussed further in Chapter 4 and in particular in Appendix 4.7.1.

Similarly, based on the inverse Fourier Transform Equation (2.19), the Inverse Discrete Fourier transform (IDFT) can be achieved using the discrete samples using the Nyquist-Shannon interpolation formula. This is converting a discrete spectrum into a discrete image in a spatial domain could also be deduced:

$$g(n_x\delta_x, n_y\delta_y) = \sum_{-\frac{N_x}{2}}^{\frac{N_x}{2}-1} \sum_{-\frac{N_y}{2}}^{\frac{N_y}{2}-1} G(m_x\Delta_x, m_y\Delta_y) \exp[j2\pi(\frac{n_x m_x}{N_x} + \frac{n_y m_y}{N_y})] \tag{2.21}$$

DFT and IDFT could be calculated by a Fast Fourier Transform (FFT) algorithm and an Inverse Fast Fourier Transform (IFFT) algorithm [61] to increase the calculation speed greatly.

2.5 Flux

2.5.1 Propagation of the Angular Spectrum

Propagation of the angular spectrum [62] is described as follows:

$$u_z(x_z, y_z) = \mathrm{FT}^{-1}\left\{\mathrm{FT}\left\{u_0(x_0, y_0)\right\} H_z(f_x, f_y)\right\} \tag{2.22}$$

where the transfer function H_z is defined as follows:

$$H_z(f_x, f_y) = \begin{cases} \exp\left(j2\pi z\sqrt{\frac{1}{\lambda^2} - {f_x}^2 - {f_y}^2}\right), \text{for} \frac{1}{\lambda^2} \geq {f_x}^2 + {f_y}^2 \\ \exp\left(-2\pi z\sqrt{{f_x}^2 + {f_y}^2 - \frac{1}{\lambda^2}}\right), \text{for} \frac{1}{\lambda^2} \leq {f_x}^2 + {f_y}^2 \end{cases} \tag{2.23}$$

where $u_0(x_0, y_0)$ is the initial wavefield defined in a planar coordinate system (x_0, y_0); $u_z(x_0, y_0)$ is the diffracted wavefield in coordinate system (x_z, y_z) following propagation a distance z. FT and FT^{-1} are the operators for the Fourier transform as defined in the previous section. In Equation 2.23, the transfer function, H_z has an infinite support in (f_x, f_y), which tends to zero at $1/\lambda^2 = {f_x}^2 + {f_y}^2$. However, it should be noted that the support of H_z that needs to be considered in the propagation of the angular spectrum defined in Equation 2.22 is typically limited by the support of the signal, which we define as $(\Delta f_x, \Delta f_y)$.

2.5.2 Physical interpretation of the the angular spectrum

We defined the propagation of the angular spectrum in terms of the scalar wavefield $u_0(x_0, y_0)$. Here, we review the physical interpretation of the angular spectrum and optical propagation. This interpretation is based on the concept of the the simple plane wave propagating with wave vector $\vec{k}$ where $\vec{k}$ has magnitude $2\pi/\lambda$ and has direction cosines (α, β, γ), as illustrated in Fig 2.9. Such a plane wave has a complex representation of the form

$$p(x, y, z; t) = \exp\left[j(\vec{k} \cdot \vec{r} - 2\pi\gamma t)\right] \tag{2.24}$$

where $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$ is a position vector (the $\hat{}$ symbol signifies a unit vector), while, $\vec{k} = \frac{2\pi}{\lambda}\left(\alpha\hat{x} + \beta\hat{y} + \gamma\hat{z}\right)$. Dropping the time dependence, the complex phase amplitude of the plane wave across a constant z-plane is given by:

$$p(x, y, z) = \exp\left[j\vec{k} \cdot \vec{r}\right] = \exp\left[j\frac{2\pi}{\lambda}(\alpha x + \beta y)\right] \exp\left[j\frac{2\pi}{\lambda}\gamma z\right] \tag{2.25}$$

where the direction cosines are interrelated through:

$$\gamma = \sqrt{1 - \alpha^2 - \beta^2} \tag{2.26}$$

Thus across the plane $z = 0$, a complex-exponential function $\exp\left[j2\pi(f_x x + f_y y)\right]$ may be regraded as representing a plane wave propagating with direction cosines

$$\alpha = \lambda f_x \qquad \beta = \lambda f_y \qquad \gamma = \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \tag{2.27}$$

In the Fourier decomposition of u_0, the complex amplitude of the plane-wave component with spatial frequencies (f_x, f_y) can be rewritten in terms of the direction cosines as follows:

$$U_0\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) = \iint_{-\infty}^{+\infty} u_0(x, y) \exp\left[-j2\pi\left(\frac{\alpha}{\lambda}x + \frac{\beta}{\lambda}y\right)\right] d_x d_y \tag{2.28}$$

which is called the angular spectrum of the disturbance $u_0(x, y)$. Optical propagation can be interpreted in terms of this angular spectrum by simply adding a delay (which is a function of the plane wave direction) to each of these angular plane waves as follows:

$$U_z\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) = U_0\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) \exp\left(\frac{j2\pi z\gamma}{\lambda}\right) \tag{2.29}$$

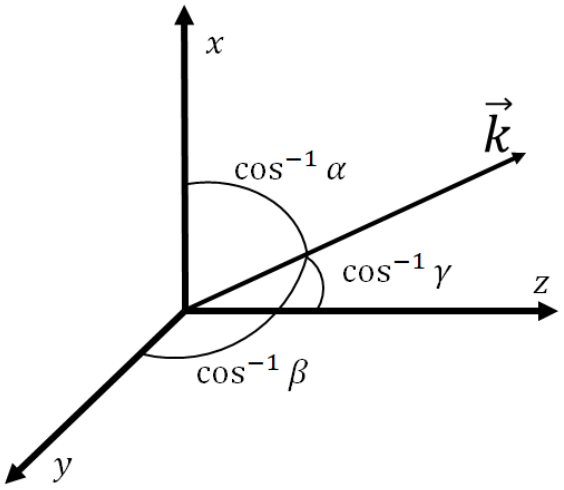

Figure 2.9: Illustration of the wave vector $\vec{k}$ for a plane wave uniquely defined by the angle cosines α, β, and γ.

2.5.3 The concept of a ray as a flux trace

A key part of Chapter 5 and Chapter 6 is the concept of a ray in terms of wave optics. The geometrical ray can be related to wave optics through a mathematical function named the ikonal [8]. The Eikonal function describes a surface in space where the wave has the same phase and can also be interpreted as the path that the ray takes through space such that a constant phase is maintained along that path. The disturbance $u(\vec{r})$ (that we have previously described with the notation $u_z(x, y)$) can be rewritten as follows:

$$u(\vec{r}) = a(\vec{r}) \exp\left[-\frac{j2\pi}{\lambda} S(\vec{r})\right] \tag{2.30}$$

where $a(\vec{r})$ is the real-valued amplitude and $-j2\pi/\lambda S(\vec{r})$ is the phase of the wave; the refractive index n of the medium is contained in the definition of S, which is called the Eikonal function. Surfaces defined by $S(\vec{r})$ = constant are called wavefronts of the disturbance. The flux (i.e. the direction of power flow) and the direction of the wave vector $\vec{k}$ are both normal to these wavefronts at every point $\vec{r}$ in an isotropic medium. A ray, therefore, is defined as a trajectory or a path through space that begins at any point on the wavefront and moves through space with the wave, always remaining perpendicular to the wavefront at every point on the trajectory. Thus a ray traces out the flux in an isotropic medium. If the disturbance $u(\vec{r})$ is to represent an optical wave, it must satisfy the scalar wave equation

$$\nabla^2 u - \frac{n^2}{c^2}\frac{\partial^2 u}{\partial t^2} = 0 \tag{2.31}$$

where ∇^2 is the Laplacian operator, n represents the refractive index of the dielectric medium within which light is propagation, and c represents the velocity of light in vacuum. Substituting u as defined in Equation 2.30 into the Helmholtz yields the following equation that must be satisfied by both $a(\vec{r})$ and $S(\vec{r})$:

$$\begin{aligned}&\left(\frac{2\pi}{\lambda}\right)^2\left[n^2 - |\nabla S|^2\right]a + \nabla^2 a \\ &-\left(\frac{j2\pi}{\lambda}\right)\left[2\nabla S\cdot\nabla a + a\nabla^2 S\right] = 0\end{aligned} \tag{2.32}$$

Examining the real part only reveals the following relationship:

$$|\nabla S|^2 = n^2 + \left(\frac{\lambda}{2\pi}\right)^2\frac{\nabla^2 a}{a} \tag{2.33}$$

Setting $\lambda \to 0$, produces the so-called Eikonal equation, which is a fundamental equation in geometrical optics:

$$|\nabla S(\vec{r})|^2 = n^2(\vec{r}) \tag{2.34}$$

This equation serves to define the wavefront in terms of the Eikonal function, S. Once the wavefronts are known, the trajectories defining rays can be determined.

2.5.4 Rays and local spatial frequency

It is also possible to approximately define the ray trajectory in terms of the local spatial frequency of the disturbance u, and in some cases, such as for a plane wave, this definition is exact. Since each of the spatial frequency components of the angular spectrum will have an infinite support over (x, y), it is not possible to directly relate a specific point in a disturbance with a single spatial frequency for the general case. However, if the disturbance is a wavefront described by slowly varying amplitude and phase functions $a(x, y)$ and $\phi(x, y)$, each point in the disturbance can be approximately interpreted as being related to a single frequency. Goodman [62] reports that this interpretation is accurate if $\phi(x, y)$ varies sufficiently slowly such that its is well ap-

proximated by only three terms of its Taylor series expansion about any point, i.e. a constant term and two first-partial-derivative terms. We begin by rewriting the disturbance once again as follows:

$$u_z(x,y) = a_z(x,y)\exp\left[j\phi_z(x,y)\right] \tag{2.35}$$

For this discussion we assume that the amplitude distribution. The local spatial frequency of u_z is defined as a frequency pair (f_{lx}, f_{ly}) as follows:

$$f_{lX} = \frac{1}{2\pi}\frac{\partial}{\partial x}\phi_z(x,y) \qquad f_{lY} = \frac{1}{2\pi}\frac{\partial}{\partial y}\phi_z(x,y) \tag{2.36}$$

In Section 2.5.2 we demonstrated that u_z could be decomposed by means of a Fourier transform into a collection of plane-wave components traveling in different directions, each defined by unique wave vector with direction cosines (α,β,γ) defined by Equation 2.26. The spatial frequencies defined through the Fourier decomposition exist at all points in space and cannot be regarded as being localized; however, for slowly varying disturbances, the definitions of the previously defined local spatial frequencies (f_{lx}, f_{ly}) can be invoked and these can be interpreted as defining the local direction cosines $(\alpha_l,\beta_l,\gamma_l)$ of the wavefront through the relations:

$$\alpha_l = \lambda f_{lx} \qquad \beta_l = \lambda f_{ly} \qquad \gamma_l = \sqrt{1-(\lambda f_{lx})^2 - \lambda f_{ly})^2} \tag{2.37}$$

These local direction cosines are the direction cosines of the ray through the (x,y) plane at each point. This definition of a ray, is consistent with the definition of a ray that was given earlier in terms of the Eikonal function, and is the basis for the nonlinear ray tracing method proposed in Chapter 5, and which is applied to predict trapping forces in Chapter 6. We strongly encourage the reader to review this section on flux befor proceeding to read Chapter 5.

2.6 Phase unwrapping

Phase unwrapping is a widely used technique in digital image processing and optical metrology. It is mentioned in the introduction chapter because we are required to apply phase unwrapping when developing the method to trace the flux lines in Chapter 5. Phase unwrapping is primarily used to remove phase discontinuities caused by measurement errors or optical non-uniformities, thereby achieving more accurate phase measurements and three-dimensional (3D) reconstruction of surfaces [63]. The process of phase wrapping/unwrapping is illustrated in Fig. 2.10 taken from Ref 64. In digital image processing, phase unwrapping works by detecting differences in phase between adjacent pixels in an image and correcting the phase values in regions where there are phase jumps. This process is typically performed in the complex plane, as phase is a periodic value and the difference in phase between adjacent pixels may be greater than one period. In optical metrology, phase unwrapping is commonly used for 3D surface reconstruction and surface morphology analysis. For example, when us-

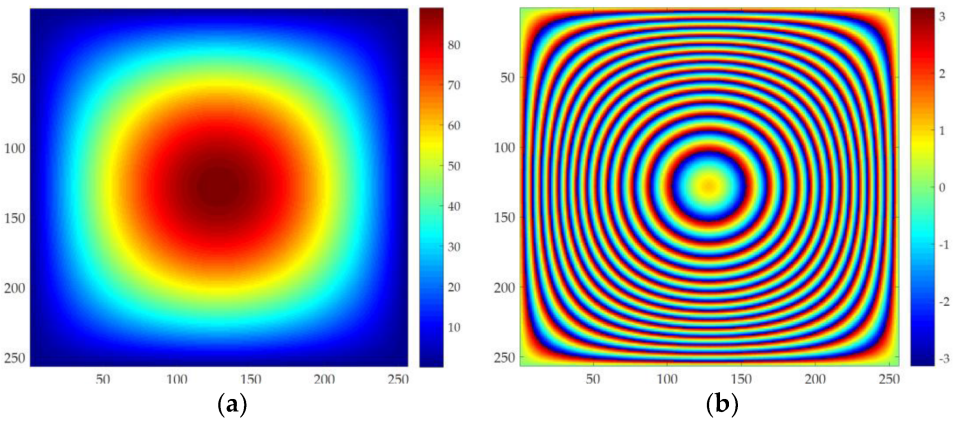

Figure 2.10: A description of phase wrapping and unwrapping [64]. (a) unwrapped phase result. (b) wrapped phase

ing interferometric measurement techniques to measure the surface of an object, the geometry of the surface is reflected by changes in phase. However, due to the limitations of interferometric measurement techniques and imperfections in optical equipment, phase discontinuities and distortions may occur. By applying phase unwrapping techniques, these errors can be removed, resulting in more accurate 3D surface reconstruction. There are many different approaches to phase unwrapping, including path-dependent methods, quality-guided methods, and network-based methods. However, in Chapter 5 we are able to implement phase unwrapping using an extremely simply technique that simply checks for 2π discontinues and correct for them.

2.7 Wavefront distortion and the Zernike polynomials

Aberration refers to defects in image formation in optical systems, [8,65] and is particularly relevant in Chapter 4 and 5. Geometrical aberrations can be broken into two classes: monochromatic aberration and chromatic aberrations. The former, which is more relevant in this thesis includes spherical aberration, coma, astigmatism, field curvature and distortion. In wave optics, aberrations relate to distortion of the wavefront and can be defined in terms of the difference between an ideal spherical wave emitted by a point source and a similarly ideal spherical wave having first passed through the optical system and consequently suffered distortion. Wavefront aberrations can be expressed succinctly in terms of the Zernike polynomials. Before proceeding to define these polynomials and their unique properties, we begin with a brief overview of some of the most commonly encountered aberrations. The Zernike polynomials are a set of orthogonal two-dimensional polynomials defined over the unit circle. These are particularly useful in describing the wavefront distortions that can exist for optical systems with circular pupils. The odd and even Zernike polynomials are given by:

$$\begin{aligned} Z_n^{-m}(\rho,\phi) &= R_n^m(\rho)\sin(m\phi) \\ Z_n^m(\rho,\phi) &= R_n^m(\rho)\cos(m\phi) \end{aligned} \tag{2.38}$$

which are defined in terms of the polar coordinates: ϕ, the azimuthal angle, and ρ the radial distance. Here, n and m are integers that uniquely define the polynomial. The factor R_m^n is defined as follows:

$$R_n^m(\rho) = \begin{cases} \sum_{l=0}^{(n-m)/2} \frac{(-1)^l (n-l)!}{l!\left(\frac{n+m}{2}-l\right)!\left(\frac{n-m}{2}-l\right)!} \rho^{(n-2l)} & \forall (n-m) \text{ even} \\ 0 & \forall (n-m) \text{ odd} \end{cases} \tag{2.39}$$

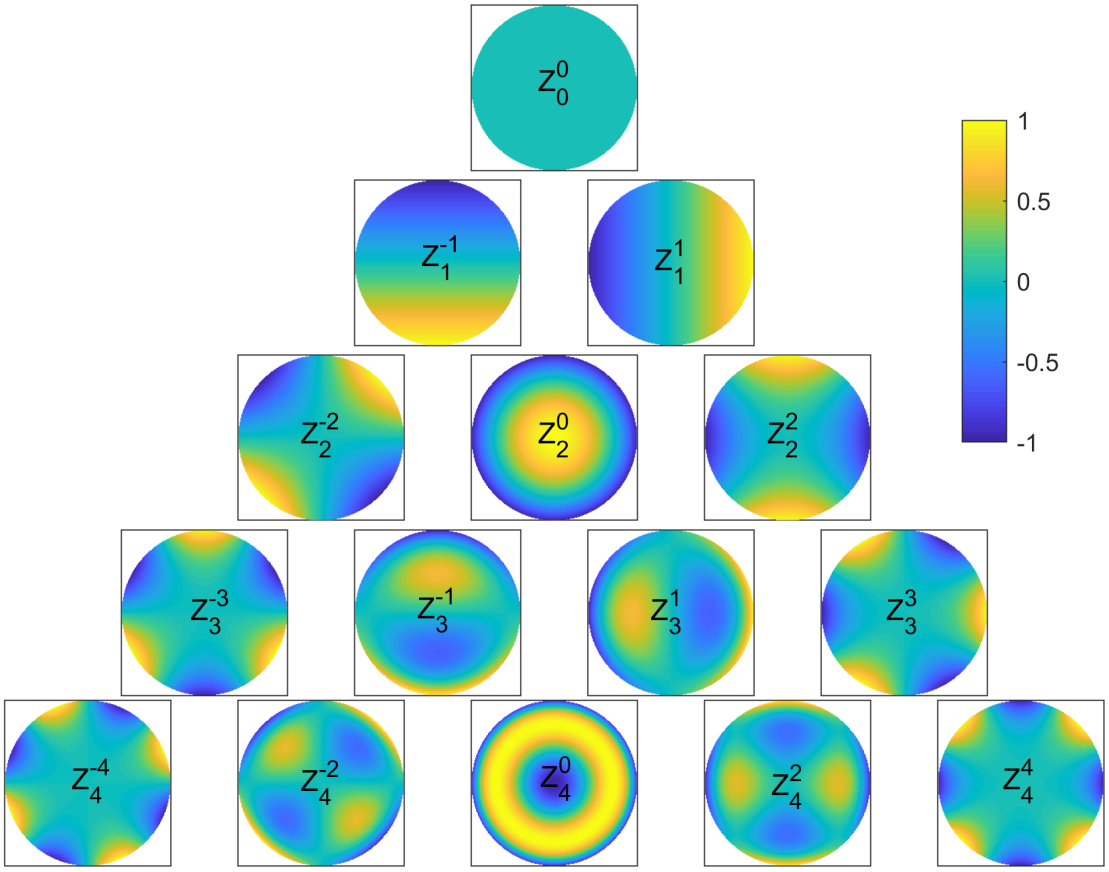

Figure 2.11: The first 15 zernike polynomials , ordered vertically by radial degree and horizontally by azimuthal degree.

These polynomials have an important mathematical property. They form a complete and orthogonal basis set over the unit circle and therefore, any given function over this unit circle can be represented as a weighted (infinite) sum of these polynomials. Therefore, any function defined over the unit circle , $W(\rho,\phi)$, can be defined in terms of these polynomials as follows:

$$W(\rho,\phi) = \sum_{n=0}^{\inf} \sum_{m=-n}^{m=-n} w_{nm} Z_n^m(\rho,\phi) \tag{2.40}$$

where w_{nm} represents the weights of the polynomials in the infinite summation. When wavefront error exists in an optical system that contains aberration, it is possible to account for this aberration as a phase delay distortion of the pupil function of the optical system. Whereas an ideal diffraction limited optical system can be described as having a frequency response or pupil function defined as circle function, $P(\rho,\phi)$, an

Table 2.1: **Various Zernike Aberrations**

Radial degree m/ n	0	1	2	3	4
0	Z_0 Constant				
1		$2r\cos\theta$ $2r\sin\theta$ Tip and Tilt			
2	$\sqrt{3}(2r^2-1)$ Defocus		$Z_4=\sqrt{6}r^2sin(2\theta)$ $Z_5=\sqrt{6}r^2cos(2\theta)$ Astigmatism		
3		$Z_6=\sqrt{8}(3r^3-2r)sin\theta$ $Z_7=\sqrt{8}(3r^3-2r)cos\theta$ Coma		$Z_8=\sqrt{8}r^3sin(3\theta)$ $Z_9=\sqrt{8}r^3cos(3\theta)$ Trefoil	
4	$Z_{10}=\sqrt{5}(6r^4-6r^2+1)$ Spherecial		$Z_{11}=\sqrt{10}(4r^4-3r^2)cos(2\theta)$ $Z_{12}=\sqrt{10}(4r^4-3r^2)sin(2\theta)$ Secondary Astigmatism		$Z_{13}=\sqrt{10}r^4sin(4\theta)$ $Z_{14}=\sqrt{10}r^4cos(4\theta)$ Quadrafoil

aberrated optical system can be described as having the same pupil function with a spatially varying phase delay $kW(\rho,\phi$, where $k = 2\pi/\lambda$ and λ is the wavelength of the light. Therefore, the pupil function, P' of an abberrated system is defined as follows:

$$P'(\rho,\phi) = P(\rho,\phi)\exp jkW(\rho,\phi) \tag{2.41}$$

where W can be described as a weighted sum of Zernike polynomials as described in Equation 2.40. In Fig. 2.11 the first fifteen Zernike polynomials are illustrated and each of these polynomials is defined in Table 2.1. The rotational symmetry between polynomials Z_n^{-m} and Z_n^m is clearly evident; one is simply a $pi/4$ rotation of the other. Only the low order polynomials Z_0^0 (constant), $Z_{\pm1}^1$ (tip/tilt), Z_0^2 (defocus), $Z_{\pm2}^2$ (vertical and oblique astigmatism), $Z_{\pm1}^3$ (spherical aberration) have geometrical ray optical interpretation. The others such as $Z_{\pm1}^3$ (treofil) and $Z_{\pm4}^4$ (quadrafoid) cannot easily be explained in terms of ray optics.

2.8 Conclusion

In this chapter, a brief literature review has been provided for some of the most important concepts that appear in this thesis, particularly on optical trapping and flux. The background theory has been introduced that underpin the contributions that appear in Chapter 3-6. Only a brief review of the trapping forces acting on a trapped particle were provided; in the next chapter we provide a full derivation of Ashkin's model, and we proceed to augment this model to account for particles that absorb light. This is followed in Chapters 4-5 in which we develop a method that can be used to trace the flux lines through the focusing laser; the introduction on Flux above is particularly relevant for these chapters. Finally in Chapter 6, we return to Ashkin's extended model and replace the rays with the flux lines, demonstrating for the first time how the spin on trapped particles can be accounted for using the ray model.

Chapter 3

A new Spin on Ashkin's Laser Trapping Forces in the Ray Optics Regime

The work in this chapter has been prepared for a submission to a journal with the following title: Qin Yu, Bryan Hennelly, "A new Spin on Ashkin's Laser Trapping Force in the Ray Optics Regime", To be submitted to Optics and Lasers in Engineering with the abstract reproduced below:

We repeat the derivation of the forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime as first reported by Ashkin; in this case, however, we investigate the impact of partial absorption of light by the sphere on these trapping forces. Our model indicates that the effect of absorption is to reduce the axial and transverse trapping forces as a function of absorption, as well as changing the axial trap position. The trap stiffness is predicted to remain relatively unchanged even for high levels of absorption. The model also predicts that a rotational force is imparted on the sphere as a consequence of absorption. The torque associated with a single ray is first derived, which is shown to increase as a function of both absorption, radius, and the angle of incidence. This enables a derivation of the torque acting on a trapped sphere as it is laterally displaced from the trapping position, which is found to increase linearly with respect to the transverse displacement of the sphere. The torque is derived to be zero for the case of no absorption, and also for the case of axial displacement of the sphere, regardless of absorption. The model reduces exactly to Ashkin's model for the case of zero absorption.

3.1 Introduction

In this chapter, we begin the journey of this thesis by re-deriving the forces acting an a dielectric sphere in an optical trap and extending this approach to predict rotational forces for the first time. We shown how Ashkin's ray-optics model is easily augmented to account for the effect of absorption of some portion of the trapping lasers power within the particle. The result is a torque that is proportional to the absorption coefficient of the material of the particle. In the two chapters that follow, the journey of

the thesis will then lead us into developing new methods of ray tracing that will allow us to account for spinning cork-screw rays associated with a Laguerre Gaussian laser mode, which is known to impart an orbital angular momentum on a trapped particle causing it to spin. This journey will eventually lead back to the equation developed in this present chapter such that the spinning forces imparted by a TEM_{01} laser can be estimated using the ray-optics model for the first time.

Ashkin first demonstrated the effect of optical forces in the early 1970s [5,66] which led to the development of the optical trap [7,37,67]. Since their invention, optical traps have been shown to have widespread applications, particularly in the field of biology [44, 56, 68–75], whereby single cells, bacteria, virus, proteins, and even DNA can be trapped in the focus of a laser beam and their position manipulated in three dimensions. The use of holographic traps have enabled manipulation of multiple particles in three dimensions by digitally controlling a spatial light modulator [20], which has further extended the application of optical traps.

For particles of size that is greater than the wavelength of light, Mie scattering results in a transfer of the linear optical momentum of light to the particle via refraction of the focussing laser beam; for this case, the forces acting on the particle can be calculated based on ray optics [7,21,22]. Ashkin presented the first comprehensive model to calculate such forces [21] by first modeling the force imparted by a single ray, and then integrating the forces imparted by a set of converging rays for a particle. He showed that as the particle moves axially and transversely around the focal point of the laser, it will experience a force that pushes it back towards the focal point. This condition is satisfied if the refractive index mismatch between the surrounding medium and that of the particle is favorable, and if the numerical aperture of the focusing optic is sufficiently high. Although other models have been developed to calculate trapping forces in the Mie regime [76, 77] based on Gaussian optics, Ashkin's model is by far the most cited and arguably the most accurate. It is this model that we wish to extend in this chapter by accounting partial absorption by the particle for the first time.

Absorption is often avoided in optical trapping, in particular for application in biology; absorption leads to heating and this can adversely affect biological material. Living cells for example cannot be analysed in an optical trap in the visible wavelength band. For this reason, wavelengths in the 750–1200 nm region are preferred as these wavelengths are not absorbed in high levels by biological materials. This window avoids the high absorption of proteins for wavelengths <750 nm and of water for >1250 nm [56]. Perhaps this the reason why there have been no previous attempts to augment Ashkin's model to account for absorption. It should be noted that outside of biological applications, absorption is more acceptable; for example the use of optical trapping to investigate ions [78], aerosols [79], or beads connecting to DNA, [79,80] can use visible traps that will be partially absorbed by the particle.

Rather than avoid absorption, an optical spanner or wrench [81–83] embraces it. This type of trap employs Laguerre-Gaussian laser spatial modes, which have been shown to carry optical angular momentum [84,85] that can be transferred to the trapped particle so long as partial absorption takes place. This work is important in the context of the contribution presented here; for an an optical spanner, transfer of rotational

force is directly related to the absorption coefficient of the particle. In this chapter, we discover that a rotational force is imparted by a a single ray due to refraction, so long as partial absorption takes place. We make no effort to link these independent theoretical predictions of absorption related rotational trapping forces; however, we feel that more investigation could possibly bridge these models.

The breakdown of this chapter is as follows. We begin in Section 3.2 with a review of Ashkins model for the force imparted by a single ray, which is extended in 3.3 to account for partial absorption of the ray; this includes the derivation of a rotational force based on refraction for the first time. In Section 3.4 the new model is used as the basis to calculate axial and transverse trapping forces for a number of different particles with varying size and absorption coefficients and these results are compared with those of Ashkin. Finally a brief conclusion is offered.

3.2 Review of Ashkin's Forces from a Single Ray

Ashkin first proposed a method to quantify the forces acting on a dielectric sphere in an optical trap in Refs [21, 86], in the ray optics regime, which is based on earlier contributions, which qualitatively described optical trapping using ray optics [7]. This model is widely accepted to be accurate for particles of a scale that is significantly greater than the wavelength of the light. In the ray optics regime, a wavefield is decomposed into a set of individual rays with varying properties of intensity, angle, and polarization. Each ray travels along a straight line in a homogeneous medium and is incident on the sphere at some angle, and consequently imparting some force. Taking into account many converging rays that are focused by a high numerical aperture microscope objective, and summing the individual forces imparted on the particle by of each of these rays, provides an estimate of the total force acting on the particle.

At the core of Ashkin's approach is modelling the force imparted on the sphere by a single ray that is incident at an arbitrary angle. This model is, in essence, based on the transmission, reflection, and refraction of the ray as it sequentially intersects the interfaces of the particle and the surrounding medium. The degree of reflection, refraction, and transmission at each interface is defined according to the Fresnel coefficients, which are related to the state of polarization, as well as Snell's law. The net result of this sequence of scattering, results is a transfer of momentum to the particle. In Fig. 3.1(a), we illustrate Ashkin's model; A ray is shown to be incident on the surface with a force vector $\vec{F}_0$, where the magnitude of this force is given by the Equation 3.1:

$$\left|\vec{F}_0\right| = \frac{n_1 P}{c} \tag{3.1}$$

where P denotes the optical power associated with the ray, n_1 is the refractive index of the surrounding medium, and c is the speed of light. For simplicity, this ray is traveling in the positive z direction and has a component only along this axis, i.e.:

$$\vec{F}_0 = \left|\vec{F}_0\right|\vec{z} \tag{3.2}$$

The angle between the ray and the surface normal is given by $(\pi - \theta)$, which is defined relative to the z-axis following the convention set by Ashkin [21]. This ray is transmitted and reflected at this first interface according to the Fresnel coefficients, which are polarisation dependent:

$$R_s = \left\{ \frac{n_1 \cos\theta - n_2 \cos r}{n_1 \cos\theta + n_2 \cos r} \right\}^2 \tag{3.3}$$

$$R_p = \left\{ \frac{n_1 \cos r - n_2 \cos\theta}{n_1 \cos r + n_2 \cos\theta} \right\}^2 \tag{3.4}$$

where Equations 3.3 and 3.4 are the reflection coefficients for s and p polarisation components, respectively, and the corresponding transmission coefficients are given by $T_s = 1 - R_s$ and $T_p = 1 - R_p$. The reflected and transmitted rays have powers of PR and PT, respectively (where R represents R_s or R_p as the case may be). The force vector associated with the reflected ray is shown in the figure as $\vec{F}_1$, which has a magnitude $\left|\vec{F}_0\right| R$ and reflects at an angle $\pi + 2\theta$. The transmitted ray is refracted at an angle r, which is defined by Equation 3.5:

$$\frac{\sin r}{\sin\theta} = \frac{n_1}{n_2} \tag{3.5}$$

This transmitted ray is then incident at the second interface at angle r with respect to the surface normal, where once again it undergoes reflection and transmission. In this case, the reflected ray remains inside the sphere with a power of PTR, and the transmitted ray passes into the surrounding medium with power PT^2. The force vector associated with the transmitted ray is shown in the figure as $\vec{F}_2$, with magnitude $\left|\vec{F}_0\right| T^2$ and angle $(2\theta - 2r)$.

The reflected ray is then incident at the third interface, again at angle r with respect to the surface normal. Again it undergoes reflection and transmission. As for the case of interface 2, the reflected ray remains inside the sphere and the power is again scaled by the factor R. The transmitted ray passes into the surrounding medium with powerPRT^2. The force vector associated with the transmitted ray $\vec{F}_3$, with magnitude $\left|\vec{F}_0\right| RT^2$ and angle $(\pi + 2\theta - 4r)$. This process repeats an infinite number of times, with each newly reflected ray being scaled by a factor R and each newly transmitted ray being scaled by a factor, R.

The total force, therefore, that is acting on the sphere is given by the sum of the forces entering the sphere minus the forces leaving. Given that only one of these vectors, $\vec{F}_0$ is incident on the sphere, while all others are transmitted out of the sphere, the total force acting on the sphere, $\vec{F}$, is given by:

$$\vec{F} = \vec{F}_0 - \sum_{m=1}^{\infty} \vec{F}_m \tag{3.6}$$

This force vector can be decomposed into horizontal and vertical components as follows:

$$\vec{F} = \left|\vec{F}\right| \cos\phi\, \vec{z} + \left|\vec{F}\right| \sin\phi\, \vec{y} = F_s\, \vec{z} + F_g\, \vec{y} \tag{3.7}$$

where F_s and F_g are the scattering and gradient forces for a single ray. Ashkin defines these forces to act in the same direction as the incident ray, and perpendicular to it,

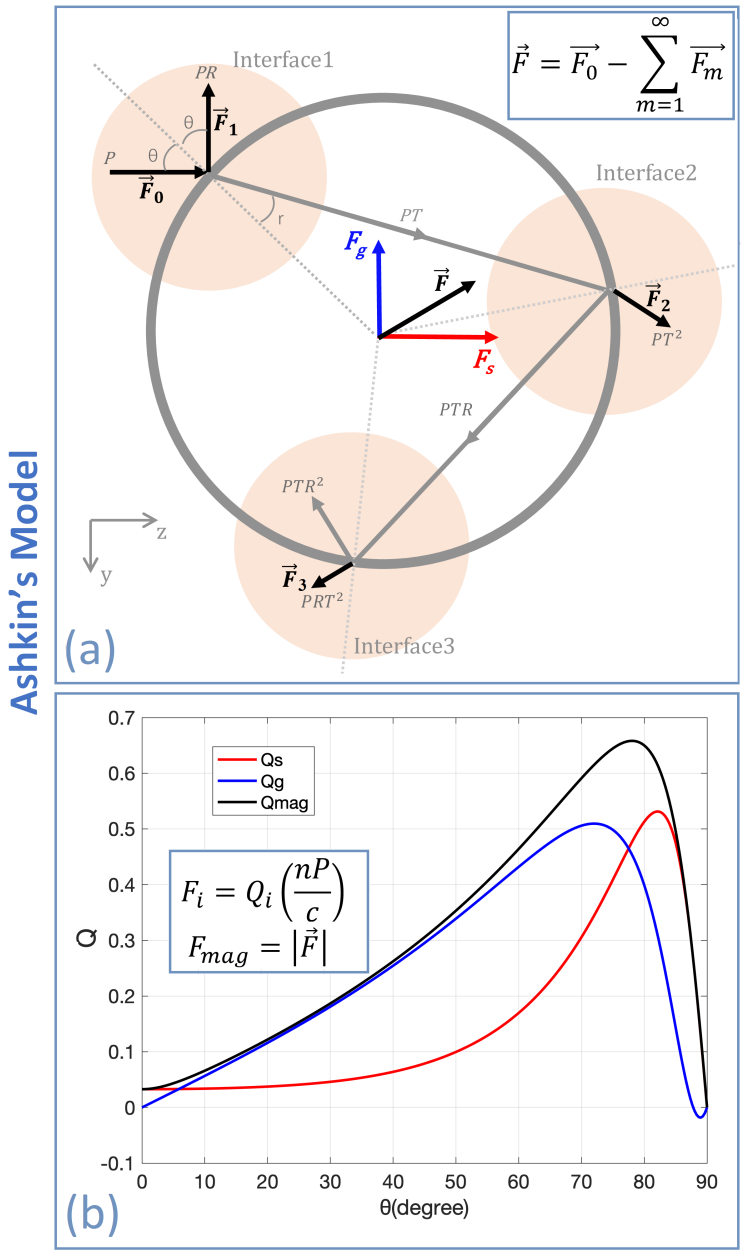

Figure 3.1: (a) Illustration of Ashkin's model for calculation of the force of a single ray with power P acting on a sphere. The incident ray with force vector $\vec{F}_0$ is incident on the sphere at angle $(\pi - \theta)$ and generates a sequence of reflection and transmission/refraction events within and out of the sphere. The total force acting on the sphere, $\vec{F}$, is given by the incident ray force, minus the forces of all rays leaving the sphere. (b) The magnitude of $\vec{F}$ as a function of θ, as well as the values of its z and y components, namely the scattering and gradient forces. Forces are shown as a fraction of $|\vec{F}_0|$, denoted as Q_i

respectively. The scattering and gradient forces can be derived by summing the sine and cosine components of each of the terms in Equation 3.6:

$$F_s = \frac{n_1 P}{c}\left[1 - R\cos(\pi + 2\theta) - \sum_{n=0}^{\infty} T^2 R^n \cos(2(\theta - r) + n(\pi - 2r))\right] \tag{3.8}$$

$$F_g = \frac{n_1 P}{c}\left[0 - R\sin(\pi + 2\theta) - \sum_{n=0}^{\infty} T^2 R^n \sin(2(\theta - r) + n(\pi - 2r))\right] \tag{3.9}$$

where the first two terms in each expression correspond to vectors F_0 and F_1 and the infinite sum relates to vectors F_m for $m \geq 2$, where $n = m - 2$. We note that only the first three interfaces are shown in Fig 3.1, although all reflection/transmissions are accounted for quantitatively in Equation 3.8 and Equation 3.9. It should be noted that although the summation has an infinite number of terms, only the first few terms have appreciable weighting in the sum owing to the repeated scaling by R. The expressions above can be simplified as follows:

$$F_s = \frac{n_1 P}{c}\left[1 + R\cos 2\theta - T^2 \frac{\cos(2\theta - 2r) + R\cos 2\theta}{1 + R^2 + 2R\cos 2r}\right] = \frac{n_1 P}{c} Q_s \tag{3.10}$$

$$F_g = \frac{n_1 P}{c}\left[R\sin 2\theta - T^2 \frac{\sin(2\theta - 2r) + R\sin 2\theta}{1 + R^2 + 2R\cos 2r}\right] = \frac{n_1 P}{c} Q_g \tag{3.11}$$

A proof of these closed-form solutions is provided in Appendix A. In Fig. 3.1(b), the relationship between the various forces and the incident angle, θ, is shown. Following from Ashkin, rather than calculate the force, which will vary depending on the incident ray power, P, the forces are shown as a fraction of the force of the incident ray, and this fraction is denoted as Q:

$$F_i = Q_i\left(\frac{n_1 P}{c}\right) \tag{3.12}$$

where i can be s, g, or mag (where $F_{mag} = \left|\vec{F}\right|$), for the values of the scattering and gradient forces, and the magnitude of the total force acting on the sphere, respectively. Similar to the same result shown in Ref [21], this plot is based on circularly polarised light for which the value of R in Equations 3.10 and 3.11 is taken to be $R = (R_s + R_p)/2$ and n_1 and n_2 are taken to be 1.33(refractive index of water) and 1.6(refractive index of Oil). Negative angles of θ are not shown in Fig. 3.1(b) as the magnitude of all three forces is symmetrical about the zero angle. It should be noted that the direction of the scattering force, F_s, is always in the positive z-direction, while the direction of F_g is in the positive y-direction for positive θ and in the negative y-direction for negative θ. The scattering force is significantly stronger for high angles, while the gradient force can be seen to increase approximately linearly as a function of angles before decaying rapidly.

3.3 Extending Ashkin's model to include Absorption

In this section, we build on Ashkin's model reviewed in Section 3.2. In this case, however, the model is extended to include absorption of the ray power within the dielectric sphere. In simple terms, the power of the ray is scaled by a factor α as it propagates from one interface to the next within the sphere. This factor is given by the product of the absorption co-efficient of the material, A, with units of $(1/m)$, and the distance the ray propagates between successive interfaces, which relates to the refraction angle, r, and the radius of the sphere, ρ as defined in Equation 3.13:

$$\alpha = 2A\rho\cos(r) \tag{3.13}$$

The extended model is illustrated in Fig. 3.2(a) in which, once again, only the first three interfaces are shown.

Note that the first transmitted ray, with power PT, arrives at the second interface with a reduced power $(1-\alpha)PT$. The resultant reflected ray, with power $(1-\alpha)PTR$, is scaled by the same amount as it propagates to the third interface, and arrives with power $(1-\alpha)^2PTR$, and this process continues for all internally reflected rays. Therefore, the first two force vectors, F_0 and F_1 remain unchanged with respect to Ashkin's model, while all other force vectors, F_m are scaled by a factor $(1-\alpha)^{m-1}$. The total force acting on the dielectric sphere is still given by Equation 3.6.

3.3.1 Re-deriving the scattering and gradient forces

The scattering and gradient forces can be calculated in the same manner as before, by summing the z and y components of each of the force vectors in Equation 3.6:

$$F_s = \frac{n_1 P}{c}\left[1 - R\cos(\pi+2\theta) - \sum_{n=0}^{\infty}(1-\alpha)^{n+1}T^2R^n\cos(2(\theta-r)+n(\pi-2r))\right] \tag{3.14}$$

$$F_g = \frac{n_1 P}{c}\left[0 - R\sin(\pi+2\theta) - \sum_{n=0}^{\infty}(1-\alpha)^{n+1}T^2R^n\sin(2(\theta-r)+n(\pi-2r))\right] \tag{3.15}$$

where the first two terms in each expression correspond to vectors F_0 and F_1, and these terms are identical to the first two terms in Equations 3.8 and 3.9 since these terms are independent of absorption. However, all terms in the infinite sum, which relate to vectors F_m for $m \geq 2$, are scaled by an amount $(1-\alpha)^{m-1}$. For simplification, the symbol n has been substituted in these expressions, where $n = m-2$. As in the previous section, these infinite summations can be defined in terms of a closed form solution as shown below in Equations 3.16 and 3.17:

$$F_s = \frac{n_1 P}{c}\left[1 + R\cos 2\theta - (1-\alpha)T^2\frac{\cos(2\theta-2r)+(1-\alpha)R\cos 2\theta}{1+(1-\alpha)^2R^2+2(1-\alpha)R\cos 2r}\right] \tag{3.16}$$

Gradient force

$$F_g = \frac{n_1 P}{c}\left[R\sin 2\theta - (1-\alpha)T^2\frac{\sin(2\theta-2r)+(1-\alpha)R\sin 2\theta}{1+(1-\alpha)^2R^2+2(1-\alpha)R\cos 2r}\right] \tag{3.17}$$

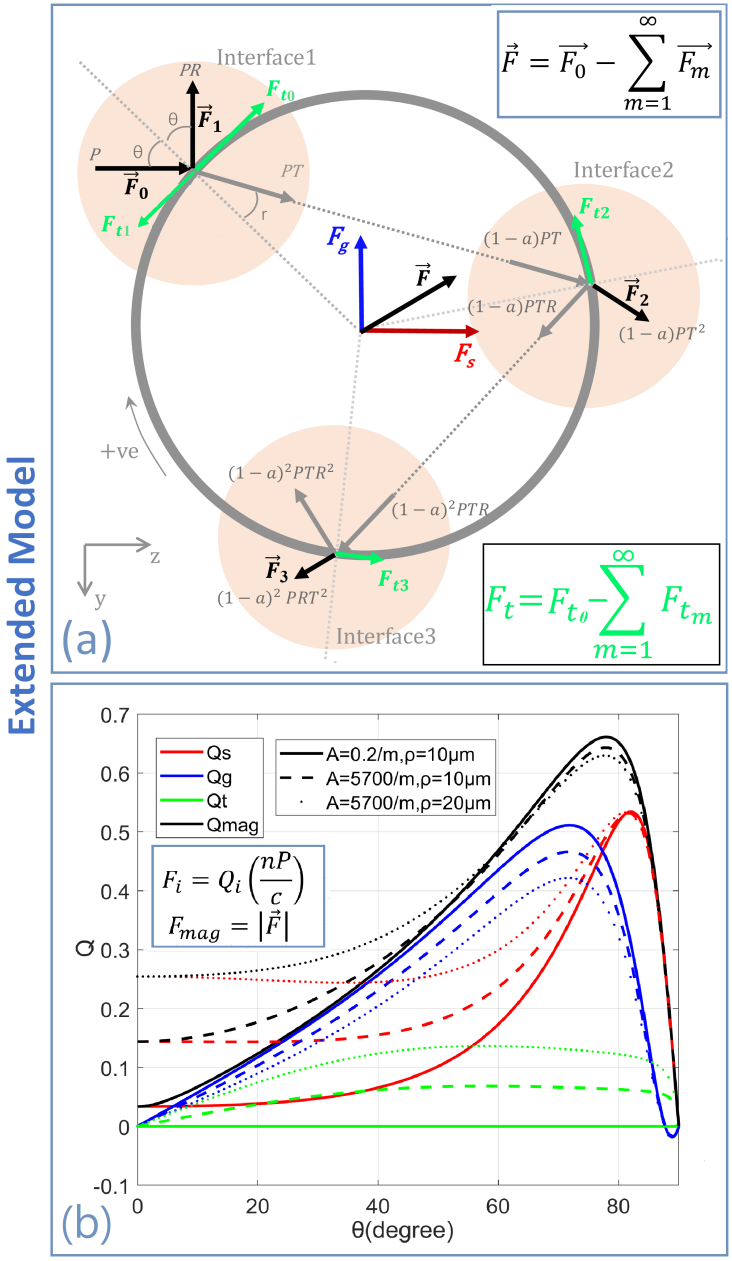

Figure 3.2: (a) Illustration of the extended model proposed in this chapter for the calculation of the forces imparted by a single ray with power P acting on a sphere. Once again, the incident ray generates a sequence of reflection and transmission/refraction events. In this extended model, the partial absorption of the ray is taken into account as it propagates from one interface to the next. As for Ashkin's model, the total force acting on the sphere, $\vec{F}$, is given by the incident ray force, minus the forces of all rays leaving the sphere. In this case, however, it is found that rotational forces can be derived acting at each interface. (b) Values of the scattering force Q_s, gradient force Q_g, and magnitude of the total force Q_{mag} for a single ray hitting a dielectric sphere with relative index of refraction $n = 1.2$ at an angle θ for three different cases: (i) a sphere of radius 10μm with an absorption coefficient of $A = 0.2/\text{m}$, (ii) a sphere of radius 10μm, with an absorption coefficient of $A = 5700/\text{m}$ and (iii) a sphere of radius 10μm, with an absorption coefficient of $A = 5700/\text{m}$. Also shown in the figure is the rotational force, denoted as a fraction Q_t, which increases a function of absorption.

We note that for the case of $\alpha = 0$, Equations 3.16 and 3.17 reduce to the same form as the corresponding expressions defined by Ashkin, see Equations 3.8 and 3.9. In Fig. 3.2(b), the relationship between the various forces and the incident angle, θ, is shown for three cases: (i) a sphere of radius 10μm with an absorption coefficient of $A = 0.2/\text{m}$, (ii) a sphere of radius 10μm, with an absorption coefficient of $A = 5700/\text{m}$ and (iii) a sphere of radius 10μm, with an absorption coefficient of $A = 5700/\text{m}$. We note that these two values for the absorption coefficient are practically real; BK7 and BG3 glass for example, have absorption coefficients of $A = 0.2/\text{m}$ and $A = 5700/\text{m}$, respectively. As for Fig. 3.1(b), the calculation presented in Fig. 3.2(b) is based on circularly polarised light for which n_1 and n_2 are taken to be 1.33 and 1.6.

The lower absorption case, for which $A = 0.2/m$, is almost identical to Ashkin's previous result, which is equivalent to $A = 0$ in the extended model. For this case, the values of α and the total absorption of the ray power is negligible for all angles, θ; therefore, the gradient and scattering forces are equivalent to Ashkin's case for all angles θ. For the two higher levels of absorption, there are striking differences with respect to Askin's result. It can be seen that as α increases, the value of the gradient force reduces across all values of θ and the value of the scattering force increases, especially for lower angles of θ. As α increases, the total force is seen to reduce marginally for high angles of $\theta > 60°$ and to increase significantly for lower angles of $\theta < 40°$.

For completeness, the relationship between the parameter α and the total amount of power that is absorbed from the input ray, P_α, is defined as follows:

$$\begin{aligned} P_\alpha &= \left|\vec{F}_0\right| - \sum_{m=1}^{\infty} \left|\vec{F}_m\right| \\ &= PT\left[1 - (1-\alpha)T\sum_{n=0}^{\infty}(1-\alpha)^n R^n\right] \\ &= PT\left[1 - \frac{T}{(1-\alpha)^{-1} - R}\right] \end{aligned} \tag{3.18}$$

Therefore, the fraction of power that is absorbed from the incident ray, Q_α, is related to the angle θ and is given by:

$$Q_\alpha = T\left[1 - \frac{T}{(1-\alpha)^{-1} - R}\right] \tag{3.19}$$

3.3.2 Torque - A rotational force emerges as a function of absorption

Ashkin's model predicts the total force acting on the sphere is given by the sum of forces in and out of the sphere as defined by Equation 3.6. The extended model uses the same same definition for the total force; however, in this case, the forces of $\vec{F}_2$ and higher are reduced due to absorption of the ray within the sphere, which results in changes to the scattering and gradient forces as discussed in the previous subsection.

In parallel to the total force acting on the sphere, a separate rotational force can also be derived by adding the components of the forces $\vec{F}_0 \rightarrow \vec{F}_\infty$ that are tangential to the surface of the sphere. Note that all of these vectors are angled at θ relative to the surface normal, which simplifies the resulting derivation; the component of each

force is scaled by $\sin\theta$. The total rotational force, F_t, is therefore given by the tangential component of $\vec{F}_0$ (denoted as F_{t0}) minus the tangential components of all of the other forces ($\sum_{n=1}^{\infty} F_{tn}$) as follows:

$$\begin{aligned} F_t &= \frac{n_1 P}{c}\left[\sin\theta - R\sin\theta - T^2\sin\theta\sum_{n=2}^{\infty}(1-\alpha)^{n-1}R^{n-2}\right] \\ &= \frac{n_1 P}{c}T\sin\theta\left[1-\frac{T}{(1-\alpha)^{-1}-R}\right] = \frac{n_1 P}{c}Q_t \end{aligned} \tag{3.20}$$

We note at the outset, that for the case of no absorption, i.e. $\alpha \to 0$, $F_t \to 0$ and the extended model reduces to Ashkin's model. For this case, F_{t0} is exactly counter balanced by the opposing forces $\sum_{n=1}^{\infty} F_{tn}$. For $\alpha > 0$ these opposing forces are weakened and are no longer strong enough to counter balance $\vec{F}_{t0}$. In order to further elucidate the origins of F_t, the rotational forces $F_{t0} \to F_{t3}$ at the first three interfaces are shown in Fig. 3.3. The rotational forces originating from the first four rays are defined below in Equation 3.21:

$$\begin{aligned} Q_{t0} &= \sin(\theta) \\ Q_{t1} &= R\sin(\theta) = RQ_{t0} \\ Q_{t2} &= (1-\alpha)T^2\sin(\theta) = (1-\alpha)T^2Q_{t0} \\ Q_{t3} &= R(1-\alpha)^2T^2\sin(\theta) = R(1-\alpha)^2T^2Q_{t0} \\ Q_{tn} &= R^{n-2}(1-\alpha)^{n-1}T^2\sin(\theta) = R^{n-2}(1-\alpha)^{n-1}T^2Q_{t0} \end{aligned} \tag{3.21}$$

and are illustrated in Fig. 3.3(a)→(d). Q_{t0} has the same sign as the angle θ, i.e. the rotational force due to $\vec{F}_0$ acts in a clockwise direction for positive θ, while the rotational force due to all rays leaving the sphere act in the opposite direction and are illustrated as such in Fig. 3.3(b)→(e). The first two rotational forces, which are associated with the input and first reflected rays are independent of absorption and have the same values for the three different spheres that are investigated in this chapter. However, as absorption increases, it can be seen that the remaining rotational forces weaken. In Fig. 3.3(e) Q_{t0} is shown alongside the total opposing force $\sum_{n=1}^{\infty} F_{tn}$ for the same three cases of absorption introduced in the earlier section. For the low absorption case, which is equivalent to a glass BK7 sphere, it is found that the torque is negligible and the extended model is equivalent to Ashkin's model both in terms of predicting the rotational force as well as the total force acting on the sphere. However, for the case of high absorption, which is equivalent to BG3, an appreciable level of torque is found. For both spheres it is observed that F_t increases as a function of θ, and it is found that these values are higher for the larger sphere of radius $\rho = 20\mu$m owing to the increased path length and consequently higher levels of absorption. It is notable that Q_{t2} is the strongest force that opposes Q_{t0}, for all but the highest angles and changes to Q_{t2} resulting from absorption, contribute to the majority of the total rotational force that manifests. Q_{t3} and higher are significantly weaker owing to the repeated multiplication with $(1-\alpha)R$ and contribute only to the rotational force when θ is large.

One final point of note in this section is the relationship between the total fraction

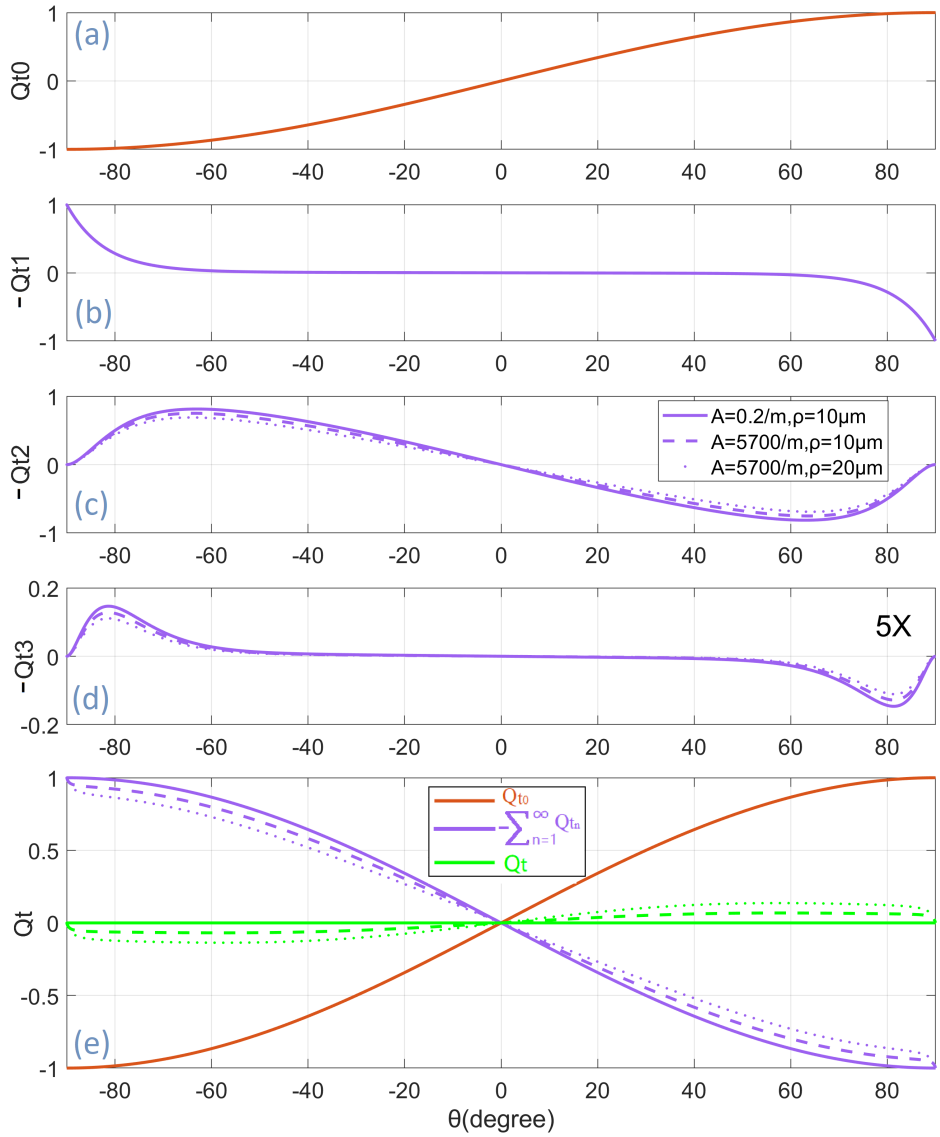

Figure 3.3: This figure illustrates the relative amplitudes of the rotational forces imparted by a single ray on the sphere, associated with the first three interfaces. (a) Illustration of Q_{t0}, the force associated with the input ray force $\vec{F}_0$. Similarly, the rotational force from the reflected ray at interface 1 is shown in (b), and for the transmitted rays at interface 2 and 3 in (c) and (d). Each of these rotational forces shown in (a)-(d) are the tangential components of the forces $\vec{F}_0 \rightarrow \vec{F}_3$ illustrated in Fig. 3.2(a). (e) Illustrate of $\mathrm{Q}_t, the remaining force of \mathrm{Q}_{t_0} - \sum_{n=1}^{\inf} Q_{t_n}$

of power absorbed by the sphere Q_α (see Equation 3.19), and the fraction of the force of the input ray contributing to rotation Q_t (see Equation 3.20). These two values are simply related by $\sin\theta$.

3.4 Calculation of Trapping Forces on a Sphere

In the previous section, an extended model was derived for the forces acting on a dielectric sphere resulting from a single ray incident at angle θ. This model includes the gradient F_g and scattering F_s forces affecting the bulk motion of the sphere, as well as a rotational force F_t. In order to solve for the trapping forces acting on the sphere in a laser trap, forces resulting from the individual focusing rays must be solved and integrated. Using the same approach as Askin we consider the computation of the force of a gradient trap on a sphere when the focus f of the trapping beam is located along the Z axis at a distance S above the center of the sphere at 0, as shown in Appendix 2 and Fig. 2. in Ref [21] for further details; the description below uses the same symbols as Ashkin.

Initially, a circle is selected representing the area of the focusing optic. The circle is sampled uniformly in two dimensions, and each sampling point represents the origin of a single ray, and all rays propagate to a common focal point. The origin of each ray is defined in terms of a polar coordinate (r,β). In order to calculate the aggregate forces acting on the sphere, resulting from each individual ray, we follow a sequence of steps, most of which are based on solving trigonometric equations. For a single ray, the first step is to calculate the angle of incidence, θ, on the sphere surface. The second step is to calculate, μ, the angle between the plane of incidence (defined by the line of the ray and a vertical line passing through the ray origin) and the vertical plane (the YZ plane). Both horizontally and vertically polarised light must be considered separately. The third step is to calculate The values of F_s, F_g, and F_t, defined in Equations 3.16, 3.17, and 3.20, for light that is $s-$ and $p-$ polarised, i.e. for $R = R_s, T = T_s$ and separately for $R = R_p, T = T_p$. This provides six separate results denoted as F_{s_s}, F_{g_s}, F_{t_s}, F_{s_p}, F_{g_p}, and F_{t_p}. The fourth step deals with light polarised horizontally relative to the Y-axis. For this case, f_s and f_p, which are the fractions of light perpendicular and parallel to the plane of incidence, respectively, are calculated as follows:

$$\begin{aligned} f_s &= (\cos\beta\cos\mu + \sin\beta\sin\mu)^2 \\ f_p &= (\cos\beta\sin\mu - \sin\beta\cos\mu)^2 \end{aligned} \tag{3.22}$$

The scattering, gradient, and rotational forces resulting from horizontally polarised light is, therefore, given by:

$$\begin{aligned} F_{s_h} &= f_s F_{s_s} + f_p F_{s_p} \\ F_{g_h} &= f_s F_{g_s} + f_p F_{g_p} \\ F_{t_h} &= f_s F_{t_s} + f_p F_{t_p} \end{aligned} \tag{3.23}$$

The fifth step deals with vertically polarised light. For this case, the values of f_s and f_p, defined above are swapped. Therefore, using the values f_s and f_p in Equation 3.22, the

scattering, gradient, and rotational forces resulting from vertically polarised light are defined as follows:

$$\begin{aligned} F_{s_v} &= f_p F_{s_s} + f_s F_{s_p} \\ F_{g_v} &= f_p F_{g_s} + f_s F_{g_p} \\ F_{t_v} &= f_p F_{t_s} + f_s F_{t_p} \end{aligned} \tag{3.24}$$

The total scattering, gradient, and rotational forces are given by the sum of these two sets of values:

$$\begin{aligned} F_s &= n_v F_{s_v} + n_h F_{s_h} \\ F_g &= n_v F_{g_v} + n_h F_{g_h} \\ F_t &= n_v F_{t_v} + n_h F_{t_h} \end{aligned} \tag{3.25}$$

where n_v and n_h are the fractions of light that is vertically and horizontally polarised, where $n_h + n_v = 1$. The sixth step involves decomposing F_s, F_g into X, Y, and Z components, and F_t into and X and Y component acting at the top point of the sphere. The final step involves repeating each of Steps $1 \rightarrow 6$ for all rays individually and adding the final X, Y, and Z components of the scattering and gradient forces together. Similarly for the rotational force, the X- and Y-components are superimposed.

3.4.1 Results

We consider only the simple case of circularly polarised light, for which $f_s = F_p = n_h = n_v = 0.5$. In this case Equation 3.25 reduces to:

$$\begin{aligned} F_s &= (F_{s_s} + F_{s_p})/2 \\ F_g &= (F_{g_s} + F_{g_p})/2 \\ F_t &= (F_{t_s} + F_{t_p})/2 \end{aligned} \tag{3.26}$$

Taking the same parameters as Ashkin in Ref [21], i.e., $n_1 = 1.33$, $n_2 = 1.6$, and a numerical aperture of 1.3, the trapping forces are calculated for three spheres with the radius and absorption properties described in Fig. 3.2.

In Fig. 3.4 (a) and (c) the trapping forces in the axial and lateral directions are shown, which are calculated using Ashkin's method [21] (i.e. using Equations 3.10 and 3.11 to calculate the forces for each ray or equivalently Equations 3.16 and 3.17 with $\alpha = 0$). For the lateral case, the Z position is taken to be 0, i.e. the line passes horizontally through the lens focus. Since Ashkin's method [21] is independent of absorption and is relative to the radius, the same result is obtained for all three spheres. Note, these results are identical to those obtained by Ashkin (see Fig.5 and Fig.7 in Ref [21]) and are reproduced here for comparison with the equivalent results generated using the extended method (i.e. using Equations 3.16 and 3.17 with $\alpha \neq 0$); these results are shown in Fig 3.4 (b) and (d).

It can be seen that the lowest absorption case, for which $A = 0.2/m$ (corresponding to BK7 glass) is identical to the result of Ashkin. However, for the higher absorption coefficient of $A = 5700/m$ (corresponding to BG3) significant differences arise. We consider first the axial trapping forces shown in Fig. 3.2 (b). Notably, the ampli-

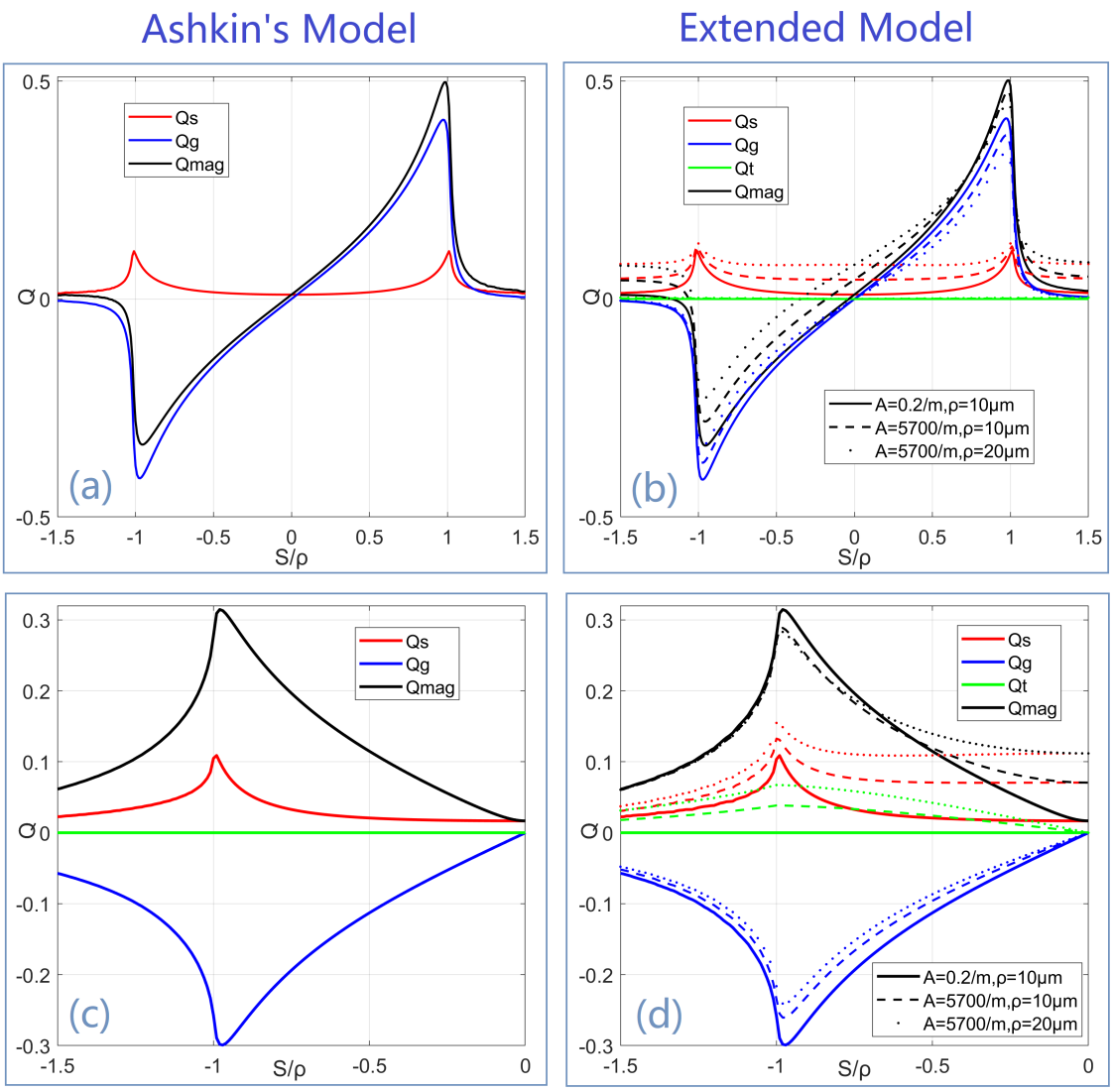

Figure 3.4: Trapping forces on the three spheres with various radius and absorption coefficient discussed in Fig 3.2: Results are shown using Ashkin's method [21] to calculate the forces on the particle as it is displaced (a) axially and (c) laterally, with respect to the focal point of the lens. The corresponding set of results is shown for when axial (b) and lateral (d) displacement, are calculated using the extended method, proposed in this chapter. Most notably we see a change in the axial trapping position as a result of absorption, as well as the emergence of a rotational force, Q_t, which manifests as a function of lateral displacement of the particle from the focal point of the lens. This rotational force increases as a function of the absorption coefficient A, and radius, ρ.

tude of the gradient force, Q_g, reduces marginally as a function of absorption, and this reduction is doubled for the particle of 20μm radius when compared with the particle of 10μm radius. The change to the scattering force, Q_s, is more appreciable, particularly for smaller displacement along the Z-axis; at the zero position the scattering force increases by approximately 400% for the smaller sphere, and 800% for the larger. The smallest increase occurs at the two maxima at $\pm\rho$ and as α increases, the scattering force appears to increase and flatten overall positions in Z. The total trap-

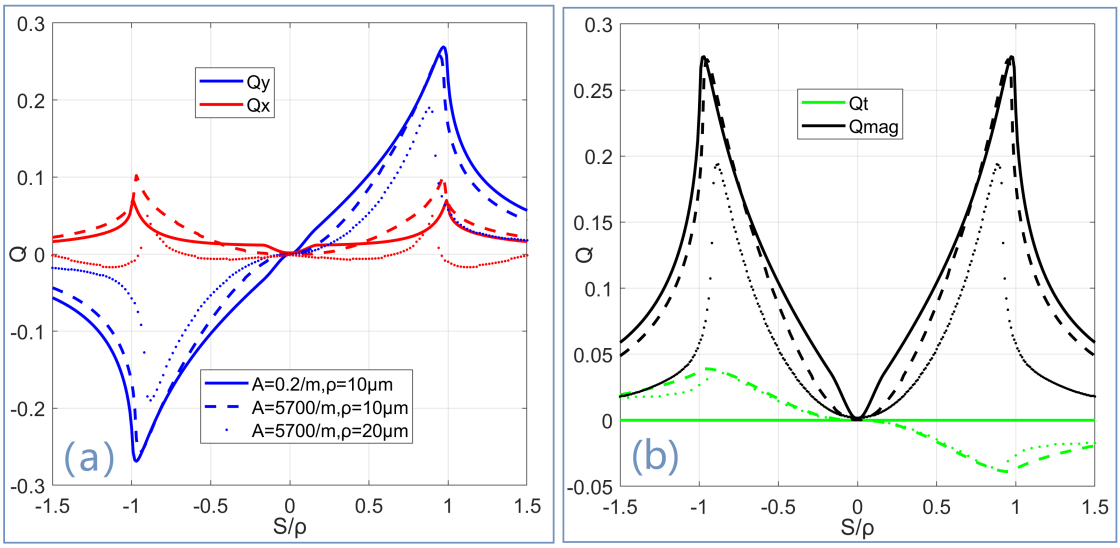

Figure 3.5: (a) The combined Y- and Z- components of the scattering and gradient forces acting on the three particles when they are displaced laterally around their axial trapping positions (b) The corresponding total force and the rotational force acting on the particles.

ping force, Q_{mag}, also changes; while the trap strength and stiffness appear to change only marginally, the trapping position changes significantly. For the smaller particle, the trapping position shifts by $\approx \rho/5$ (or 2.02μm), while for the larger one, it shifts by $\approx \rho \times 2/5$ (or 7.85μm). In units of radius, the shift in the axial trapping position appears to be approximately linear as a function of α. Notably, the rotational force, Q_t is zero at all particle positions along the axis due to the cancellation of forces from symmetrically opposing rays.

We consider now the lateral trapping forces shown in Fig. 3.2 (d) for the Y-axis centered at $Z = 0$ with respect to the focal point; only negative displacement is shown with symmetrical results for $Y > 0$. Once again, the amplitude of the gradient force, Q_g, reduces marginally as a function of absorption while the scattering force, Q_s, is seen to increase significantly as absorption increases, particularly for small movement along the Y-axis. The total force Q_{mag} is broadly unaffected for displacement $> 0.5\rho$ and for smaller movements, it increases significantly. The value of Q_t is negligible for the particle with the lower absorption coefficient; however, for the two particles with higher absorption coefficient, Q_t is seen to increase as a function of displacement reaching a maximum value at $Z = -\rho$ before decreasing for larger displacement. The rotational force is seen to be larger for the larger particle. We note that these lateral forces are calculated for the Y-axis passing trough the focal point of the lens at $Z = 0$. At this position the gradient force has a contribution only in the Y-direction while the scattering force has a contribution only in the Z-direction. Since this is not the axial trapping position for any of the three particles, the force on the particles is not equal to zero at $Y = 0$. We have taken this approach in order to compare with the results of Ashkin in his seminal paper [21]. The correct approach is to calculate the lateral forces for all three particles along a line parallel to the Y-axis, passing through the appropriate trapping

position along the Z-axis (i.e. at $Z = -0.42\mu m$, $Z = -2.02\mu m$, and $Z = -7.85\mu m$ for the three particles of interest here; note for the case of no absorption the trapping position is found to be $Z = -0.4\mu m$). Taking this approach, the total trapping force Q_{mag} must equal zero at a displacement of $Y = 0$. For completeness, we provide this result in Fig. 3.5. Note that for this case the gradient and scattering forces each have contribution in both the Y- and Z-directions. The total forces in the Y- and Z-directions are shown for the three particles in Fig. 3.5(a), while the total force (root sum of squares of forces in (a)) are shown in Fig. 3.5(a) along with the torque. It is found that trap strength is much the same for the similarly sized particles, irrespective of absorption. However, for the particle with higher absorption there is a notable reduction in the trap stiffness. For the larger particle there is a further reduction in stiffness and in this case there is also an appreciable reduction in trap strength. Interestingly, the relationship between torque and lateral displacement is similar for the two particles with higher absorption and does not vary significantly as a result of radius. This torque acts only in the Y-plane.

3.5 Conclusion

In this chapter, we have proposed an extension of Ashkin's model for deriving the the forces imparted on a dielectric sphere by as single ray. The only conceptual change that we introduce is the inclusion of a parameter to account for partial absorption of the rays as it passes through the particle as a function of the materials absorption coefficient. Based on this extended model for a single ray, we calculate the axial and lateral trapping forces for a number of different spheres with various radius and absorption coefficients. In particular, we examine two cases of absorption coefficient corresponding to low (0.2/m - BG7) and high (5700/m -BG3) absorption. The extended model predicts that while the trap strength and stiffness are only slightly changed, the trapping position will change significantly as a function of absorption. This manifests from a significant increase in the scattering force for small axial displacements. The most striking result is the emergence of a strong rotational force which increases a function of lateral displacement and absorption. For the particle with low or negligible absorption coefficient, this rotational force reduces to zero and the extended model provides identical results to Ashkin.

It must be noted that rotational forces imparted by an optical trap is not new, and has been observed by many authors for the case of lasers with Laguerre-Gaussian (L-G) spatial modes; such optical traps are commonly referred as an optical spanners or wrench. However, this chapter offers the first attempt to predict a rotational force caused from a single ray caused by refraction, and for a TEM_{00} spatial mode. Interestingly, partial absorption is also required for an optical spanner to impart torque; partial absorption of a high order L–G mode is known to result in the transfer of angular momentum to the particle in the trap. This transfer of angular momentum is predicted to be lh multiplied by the number of photons absorbed. This chapter poses a fascinating question: could the equations derived in this Chapter possibly be used to predict the torque imparted by a laser with a L-G spatial mode. This answer to this question is not

trivial. Classically Ashkin predicts the forces on a particle trapped with such a laser by simply modulating the 'power' associated with each ray with the donut intensity profile that is characteristic of the Laguerre polynomial. The force associated with each ray is weighted by this intensity before the integration. Such a simple approach cannot account for the spin that is imparted physically. In the two chapters that follow we develop a new method known as non-linear ray tracing that can model the rays of the focusing Laguerre-Gaussian laser to be spinning in a corkscrew fashion. In the final chapter we use this model, combined with the equations developed in this chapter to estimate the spinning forces acting on a L-G trap for the first time.

A final point of note is the absence of experimental results to verify the predictions of the extended model. This must be the subject of future work; while it may be possible to confirm the change in trapping position relating to increased absorption with relative ease, it may prove more difficult to accurately measure rotational force as the particle is displaced laterally.

Chapter 4

Nonlinear Ray Tracing in Focused Fields. Part 1: Calculating 3D Complex Wave-fields

The work in this chapter has been prepared for a submission to a journal with the following title: Qin Yu, Bryan Hennelly, "Nonlinear Ray Tracing in Focused Fields, Part1: Calculating 3D Complex Wavefields", To be submitted to Applied Optics with the abstract reproduced below:

In this two part chapter, we develop a method to trace the lines of flux through a three-dimensional wavefield by following a direction that is governed by the derivative of the phase at each point, a process that is best described as flux tracing but which we interchangeably name as 'non-linear ray tracing'. In this first part, we focus on the high-speed calculation of three-dimensional complex wavefields. The basis of this calculation is the Angular Spectrum method, a well known numerical algorithm that can be used to efficiently and accurately calculate diffracted fields for numerical apertures <0.7. It is known that this approach yields identical predictions to the first Rayleigh-Sommerfeld solution. It has recently been shown that the Angular Spectrum method can efficiently calculate highly focused diffraction patterns with high sampling rates by decomposing the input image into separate tiles and superimposing them together. In this chapter, we provide a rigorous derivation of this algorithm using the first principles of sampling theory. We employ the Angular Spectrum to develop two algorithms that generate the 3D complex wavefield in the region of focus of a lens. The first algorithm is based on the thin lens approximation and the second is based on the ideal lens, which can be modeled an optical Fourier transform. Both algorithms are investigated to calculate focused laser beams with TEM_{00} and TEM_{01} laser profiles. For the TEM_{00} case, we also simulate the effects of various types of optical aberration, as described by the Zernike polynomials, on the three-dimensional focused wavefield.

4.1 Introduction

We now deviate away from calculating trapping forces as seen in the previous chapter and over the course of the next two chapters we develop a method to trace the flux-lines, i.e. the fluid movement of power through the propagation of the laser. Once the method for tracing these 'non-linear rays' has been developed in Chapters 4 and 5, we can return to the Equations developed in Chapter 3 and calculate the trapping forces acting on a sphere that are caused, not by linear rays converging to a single point at the focus, but instead by the non-linear flux lines that bend around the focal point.

The objective of this chapter and the next is to develop a set of computationally efficient methods that accurately trace the flux lines in the focal region of a lens. These flux lines, which we interchangeably call 'non-linear rays' are discussed in much more detail in the next chapter. Briefly, the concept of wave optics and geometrical ray optics can be related using the Eikonal function [8], which can be related to the derivative of the phase of the wavefront. By following a path that is guided by the derivative of the phase of the wavefield in three dimensions, these flux lines can be traced. Indeed these non-linear rays and the classical rays are one and the same if the wavefront is slowly varying. A physical interpretation of the flux lines and their relationship to classical ray optics is provided in the next chapter as is a detailed discussion of the method used to trace the lines. A necessity of this method is the availability of the samples of the converged wavefield in the focal region of the lens over a three-dimensional grid. The computation of this 3D sample distribution in an accurate and timely manner is the subject of this chapter, which forms a necessary prerequisite for the method outline in the next chapter.

More specifically, in this chapter, we focus on numerical methods that can accurately calculate a propagating complex wavefield over a three-dimensional volume, and the associated sampling requirements and computational efficiencies are considered in some detail. We limit this study to focusing fields, for which the most rapid change of flux can be expected. Two distinct numerical algorithms are developed for the case of a thin lens and an ideal lens, respectively, both of which iteratively apply the angular spectrum method over a range of distances, albeit in different manners. In the next chapter, these numerical algorithms form the basis of a nonlinear ray tracing algorithm, whereby the 3D wavefield is sampled in a uniform grid in the focal region of the lens, and the 'ray' is guided by the derivative of the phase, which is calculated at a sequence of points traveling through this grid. Interestingly the nonlinear rays appear to propagate around regions of low intensity. Particular attention is given in both this chapter and the next, to convergent lenses that contain aberration, several of which cannot readily be interpreted in terms of linear geometrical ray optics but as demonstrated in the next chapter can be interpreted more easily in terms of the nonlinear rays. The concept of flux is explored in the next chapter; in this chapter, we focus only on calculating a focusing wavefield over a uniform 3D grid of sampling points.

At the heart of this chapter is the concept of numerical propagation. Numerical propagation algorithms that can accurately calculate diffraction patterns at variable distances within an homogeneous isotropic medium are an active area of research,

with far reaching applications including in the areas of quantitative phase microscopy [87] computer generated holography [88], confocal microscopy [89]computation of the point spread function in lens design and lithography [90, 91] and digital holography [92]. Solutions based on paraxial scalar optics include various discretisations of the Fresnel and Linear Canonical transforms [93] continue to find application in quantitative phase imaging. However, algorithms that are founded on nonparaxial scalar optics are, in general, preferable since they offer superior performance in terms of accuracy, and can be numerically implemented at least as efficiently as paraxial based algorithms.

These algorithms include the Angular Spectrum method (ASM) [62, 94] and the Rayleigh-Sommerfeld method (RSM) [95–97] which can both provide highly accurate solutions for numerical apertures up to at least 0.6 [90]; For higher numerical apertures, algorithms based on the more intricate vectorial model are required [91]. Indeed the first Rayleigh–Sommerfeld diffraction integral and the propagation of the angular spectrum can be shown to provide identical solutions [62]. However, their discrete counterparts, the RSM and ASM have significantly different sampling considerations [95, 98–100]. In this chapter, we employ the ASM as the numerical propagation algorithm of choice, making use of recent advances in sampling theory to optimise the the computational memory efficiency of the ASM for focused fields [13, 14].

The contributions in this chapter are as follows:

- We develop two algorithms for the calculation of the diffraction pattern at the focal point of a lens over a 3D grid of sampling points. The first algorithm, which relates to the case of the thin lens approximation, iteratively applies the ASM, where the lens function is first 'pre-aliased' [13, 14] in order improve the computational efficiency. The second algorithm, which relates to the ideal lens involves the use of a single discrete Fourier transform to 'propagate' from the front focal plane of the lens to the back focal plane followed by an iterative application by the ASM.

- We extend these algorithms to include different laser spatial modes that are focused by the lens as well as different lens aberrations.

- We provide for the first time a rigorous proof using fundamental sampling theory, of the superposition-tile method introduced by Kanka [13] and further developed by Kelly, [14] which involves adapting the input to the ASM in order to increase its computational efficiency. This involves breaking the input into smaller 'tiles', adding these together, and using the result as input to the ASM. We also examine the computational efficiency of the approach for the first time, which we find is only improved if the ASM is to be applied iteratively over several distances for the same input.

- The Rayleigh-Sommerfeld aspheric thin lens function is introduced for the first time, which was found to produce superior focusing when compared to spherical and commercial aspheric lens functions.

The breakdown of this chapter is as follows: In Section 4.2.1, the ASM is described and the associated sampling conditions are examined. Particular attention is given

to the superposition-tile method, [14] which is renamed as the 'pre-aliasing' method in this chapter; this is accompanied by a rigorous proof of the underlying sampling theory in Appendix B. In Section 4.3, the two algorithms are defined to calculate the 3D diffraction pattern around the focal point of a thin lens and an ideal lens, respectively. In Sections 4.4 and 4.5, the algorithms are extended to account for the spatial mode of the laser and lens aberrations, respectively. This is followed by a brief discussion in Section 4.6.

4.2 The Angular Spectrum Method

4.2.1 Propagation of the Angular Spectrum

Propagation of the angular spectrum [62] is described as follows:

$$u_z(x_z, y_z) = \mathrm{FT}^{-1}\left\{\mathrm{FT}\left\{u_0(x_0, y_0)\right\} H_z(f_x, f_y)\right\} \tag{4.1}$$

where the transfer function H_z id defined as follows:

$$H_z(f_x, f_y) = \begin{cases} \exp\left(j2\pi z\sqrt{\frac{1}{\lambda^2} - {f_x}^2 - {f_y}^2}\right), \text{for} \frac{1}{\lambda^2} \geq {f_x}^2 + {f_y}^2 \\ \exp\left(-2\pi z\sqrt{{f_x}^2 + {f_y}^2 - \frac{1}{\lambda^2}}\right), \text{for} \frac{1}{\lambda^2} \leq {f_x}^2 + {f_y}^2 \end{cases} \tag{4.2}$$

where $u_0(x_0, y_0)$ is the initial wavefield defined in a planar coordinate system (x_0, y_0); $u_z(x_z, y_z)$ is the diffracted wavefield in coordinate system (x_z, y_z) following propagation a distance z. FT and FT^{-1} are the operators for the Fourier transform as defined in Appendix B. In Equation 4.2, the transfer function, H_z has an infinite support in (f_x, f_y), which tends to zero at $1/\lambda^2 = {f_x}^2 + {f_y}^2$. However, it should be noted that the support of H_z that needs to be considered in the propagation of the angular spectrum defined in Equation 4.1 is typically limited by the support of the signal, which we define as $(\Delta f_x, \Delta f_y)$.

4.2.2 The Angular Spectrum Method

The Angular Spectrum Method (ASM) can be used to calculate the samples of a the diffracted wavefield, u_z, using the samples of u_0. We stipulate that the two continuous wavefields have spatial support $(\Delta_{x_0}, \Delta_{y_0})$ and $(\Delta x_z, \Delta y_z)$ in space and both have spatial frequency support of $(\Delta f_x, \Delta f_y)$, which is a consequence of the properties of the propagation of the angular spectrum as described in Section 4.2.1. We note that the spatial supports at distance 0 and at z can be significantly different for the case of a converging or diverging wavefield. It is well known that a function cannot be truly bounded with finite support in both space and spatial frequency; however, as discussed in Section 4.2.1 this can be said to be approximately true in many cases. Both continuous functions u_0 and u_z can be recovered (via Shannon interpolation) from their discrete samples, if uniform sampling intervals (T_x, T_y) have been applied where the Nyquist condition is satisfied, i.e. $T_x \leq 1/\Delta_{f_x}$ and $T_y \leq 1/\Delta_{f_y}$.

The ASM is based on a direct 'discretization' of Equations 4.1 and 4.2 and is implemented using the Discrete Fourier Transform (DFT) in place of the Fourier transform operator in Equation 4.1 as follows:

$$u_z(n_xT_x, n_yT_y) = \\ \text{DFT}^{-1}\left\{\text{DFT}\left\{u_0(n_xT_x, n_yT_y)\right\} H_z(m_xT_{f_x}, m_xT_{f_x})\right\} \tag{4.3}$$

where DFT and DFT^{-1} are the operators for the discrete Fourier transform. In summary, the specific steps involved in implementing the ASM are as follows:

1. Firstly $u_0(x_0, y_0)$ is sampled to produce the discrete function $u_0(n_xT_x, n_yT_y)$ for $n_x = -N_x/2 \rightarrow N_x/2-1$ and $n_y = -N_y/2 \rightarrow N_y/2-1$, where these samples span the full support: $\Delta_{x_0} = N_xT_x$ and $\Delta_{y_0} = N_yT_y$.

2. Secondly, these $N_x \times N_y$ samples are input to a DFT to produce the discrete samples of the Fourier transform $U_0(f_x, f_y)$ denoted as $U_0(m_xT_{f_x}, m_yT_{f_y})$ for $m_x = -N_x/2 \rightarrow N_x/2-1$ and $m_y = -N_y/2 \rightarrow N_y/2-1$, where $T_{f_x} = 1/T_xN_x$ and $T_{f_y} = 1/T_yN_y$.

3. These discrete samples are multiplied by the discrete samples $H_z(m_xT_{f_x}, m_yT_{f_y})$ over the same range of m_x, m_y.

4. Finally, and inverse DFT is applied to the resultant samples to produce the samples of the propagated function $u_z(n_xT_x, n_yT_y)$ over the same range of n_x, n_y as the input.

We refer the reader to Appendix B, Fig.1 for a definition of the DFT as well as a detailed derivation of the sampling constraints of the ASM.

4.2.3 Pre-Aliasing for convergent wavefields

Recently, there have been several contributions on the numerical calculation of diffraction fields converging at the focal point of a lens [14, 95, 100]. Of particular interest in this chapter is the method developed by Kanka [13], and further investigated by [14], which exploits redundancy for the case of a small spatial support in the output field for convergent focusing fields; in such cases the input to the ASM can be broken up into smaller tiles, which are superimposed, and this smaller matrix serves as input to the ASM without any loss in accuracy. In this chapter we refer to this method as 'prealising' and a rigorous proof of its validity is provided in Appendix B using fundamental sampling theory. The method is summarised below.

Taking the definition of the ASM above, the discrete signals $u_0(n_x, n_y)$ and $u_z(n_x, n_y)$, which are input and output to the ASM must have the same support, $(\Delta_{x_0}, \Delta_{y_0}) = (N_xT_x, N_yT_y)$. However, for convergent wavefields, the propagated wavefield will have significantly smaller support: $(\Delta_{x_z}, \Delta_{y_z}) \ll (\Delta_{x_0}, \Delta_{y_0})$. In such case, a smaller number of samples $M_x \times M_y$ is required to fully represent the continuous function u_z, where $(\Delta_{x_z}, \Delta_{y_z}) = (M_xT_x, M_yT_y)$. It is possible to generate a discrete input function in Step 1

that will produce this reduced set as samples. In short, Step 1 in the algorithm described in Section 4.2.2 is replaced with the following:

1. As before, $u_0(x_0, y_0)$ is sampled to produce the discrete function $u_0(n_x T_x, n_y T_y)$ for $n_x = -N_x/2 \rightarrow N_x/2 - 1$ and $n_y = -N_y/2 \rightarrow N_y/2 - 1$, where again $\Delta_{x_0} = N_x T_x$ and $\Delta_{y_0} = N_y T_y$. However, in this case, an aliased version of this function, $u_0'(n_x, n_y)$, is generated to serve as input to Step 2:

$$\begin{aligned} u_0'(n_x T_x, n_y T_y) = \sum_{k_x=-\frac{P_x}{2}}^{\frac{P_x}{2}-1} \sum_{k_y=-\frac{P_y}{2}}^{\frac{P_x}{2}-1} & u_0(k_x T_x + k_x M_x T_x, n_y T_y + k_y M_y T_y) \\ & \forall n_x = -M_x/2 \rightarrow M_x/2 - 1 \\ & \forall n_y = -M_y/2 \rightarrow M_y/2 - 1 \end{aligned} \tag{4.4}$$

   where the integers P_x and P_y are the largest integers that satisfy the following relationship:

$$\begin{aligned} P_x &< \frac{N_x}{M_x} \\ P_y &< \frac{N_y}{M_y} \end{aligned} \tag{4.5}$$

Steps 2-4 are implemented as described in Section 4.2.2, where M_x and M_y are substituted for N_x and N_y. Even though this function is clearly aliased, it can serve as an input to Steps 2-4 to produce an accurate calculation of the samples of u_z such that the continuous function can be interpolated from these samples.

4.2.4 Sampling conditions in three dimensions for the ASM

Given the definition for the propagation of the angular spectrum in Section 4.2.1, the required sampling conditions can be defined in terms of Nyquist and Shannon sampling theory [61]. Firstly, the input function u_0 must be Nyquist sampled such that the sampling intervals $T_x \leq \left|\frac{1}{\Delta f_x}\right|$ and $T_y \leq \left|\frac{1}{\Delta f_y}\right|$. However, if the support $(\Delta f_x, \Delta f_y)$ is greater than the support of H_z defined in Equation4.2, which has radius $\frac{1}{\lambda}$, then this latter support will constrain the sampling intervals. More explicitly, the values for T_x and T_y should be chosen to satisfy the constraint:

$$T_{x,y} \leq \max\left(\frac{\lambda}{2}, \frac{1}{\Delta f_{x,y}}\right) \tag{4.6}$$

In general, $\frac{1}{\Delta f_{x,y}}$ will be the larger of the two values and in that case, the spatial frequency support of u_z will be the same as that of u_0 and, therefore, the same sampling intervals are applicable in both planes. This case is illustrated in Fig. 4.1(a). We note that there have been several contributions in the literature that relax the Nyquist constraint for the propagated wavefield, or indeed for any chirped wavefield, which can be related to the shape of the signals phase space distribution [101–103]. These generalized sampling theorems can tolerate the aliasing of the function in the Fourier domain

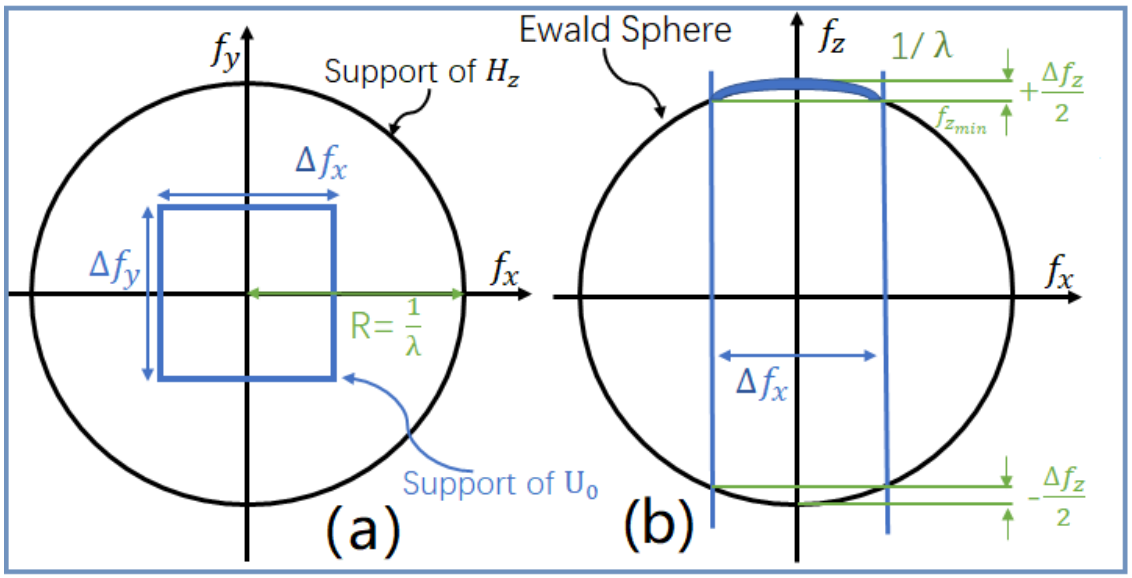

Figure 4.1: Illustration of the bandwidth of the three-dimensional function $u_z(x_z, y_z)$; (a) the bandwidth in x and y is in general determined by the bandwidth of the initial function u_0 and (b) the bandwidth in the z-direction is determined by the size of the surface area on the Ewald sphere over which the function exists.

caused by breaking of the Nyquist condition. However, the ASM cannot generally be applied to such a sub-Nyquist sampled function due to the multiplication in Equation 4.3; see Ref 95 for an excellent analysis of this subject.

The previous paragraph deals with the sampling conditions for the input/output functions for the ASM for the x- and y-dimensions. However, if a three dimensional wavefield was to be sampled, for example by repeated application of the ASM over a sequence of different distances, $n_z T_z$, where n_z is an integer, then the sampling interval, T_z also needs to be considered. Taking the Fourier transform over z of the function $u_z(x_z, y_z)$ as defined by Equation 4.1, it is straightforward to deduce that the three-dimensional Fourier transform of u_z exists only over a surface defined by the Dirac-delta functional $\delta\left(f_z - \sqrt{\frac{1}{\lambda^2} - f_x^2 - f_y^2}\right)$, which relates to the well-known Ewald sphere [8], as illustrated in Fig. 4.1 (b). Therefore, the maximum frequency in the z-direction is given by $f_{z,max} = \frac{1}{\lambda}$, which occurs at $f_x = f_y = 0$; and the minimum frequency in the z-direction occurs at $f_{x,y} = \min\left(\frac{1}{\lambda}, \frac{\Delta f_{x,y}}{2}\right)$. The first term in the minimum function relates to the case where the spatial frequency support is imposed by H_z, and the second term relates to the case where it is imposed by u_0; this latter case is the general case and is illustrated in Fig. 4.1. The minimum spatial frequency in f_z is given by:

$$\begin{aligned} f_{z,min} &= \sqrt{\frac{1}{\lambda^2} - \min\left(\frac{1}{\lambda^2}, \left[\left(\frac{\Delta f_x}{2}\right)^2 + \left(\frac{\Delta f_y}{2}\right)^2\right]\right)} \\ &= \max\left(0, \sqrt{\frac{1}{\lambda^2} - \left(\frac{\Delta f_x}{2}\right)^2 + \left(\frac{\Delta f_y}{2}\right)^2}\right) \end{aligned} \tag{4.7}$$

and the sampling interval in the z-direction is given by:

$$\begin{aligned} T_z &\leq \frac{1}{\Delta f_z} = \frac{1}{2(f_{z,max} - f_{z,min})} \\ &\leq \max\left(\lambda, \frac{0.5}{\frac{1}{\lambda} - \sqrt{\frac{1}{\lambda^2} - \left(\frac{\Delta f_x}{2}\right)^2 + \left(\frac{\Delta f_y}{2}\right)^2}}\right) \end{aligned} \tag{4.8}$$

Therefore, if the ASM is used to calculate the samples of a wavefield in three dimensions, sampling intervals of T_x and T_y should meet the constraint set out in Equation 4.6 and the sampling interval T_z should meet the constraint set out in Equation 4.8 such that the continuous three dimensional function $u_z(x_z, y_z)$ can be recovered via Shannon interpolation; i.e. by convolution with a three dimensional Sinc convolution as follows:

$$\begin{aligned} u_z(x_z, y_z) = &\sum_{\frac{-N_z}{2}}^{\frac{N_z}{2}-1} \sum_{\frac{-N_x}{2}}^{\frac{N_x}{2}-1} \sum_{\frac{-N_y}{2}}^{\frac{N_y}{2}-1} u_{n_z T_z}(n_x T_x, n_y T_y) \\ &\times \operatorname{sinc}\frac{\pi(x_z - n_x T_x)}{T_x} \operatorname{sinc}\frac{\pi(y_z - n_y T_y)}{T_y} \operatorname{sinc}\frac{\pi(z - n_z T_z)}{T_z} \end{aligned} \tag{4.9}$$

4.2.5 Computational efficiency of the ASM

It should be noted that the pre-aliasing step for convergent fields described in Section 4.2.3 requires the same Nyquist condition to be satisfied. The redundancy that is exploited relates to the small spatial support of u_z, which places a relaxed sampling condition on the DFT, which in turn translates to pre-aliasing of u_0. The pre-aliasing step is not required for convergent fields; it is always possible to avoid this step but the advantage of pre-aliasing is computational and memory efficiency. For example, numerical propagation between a lens with large radius and high numerical aperture to its focal plane, would necessitate an input field with a small sampling interval applied over a large area; the total number of samples for such a discrete function may be onerous in terms of temporary computer memory. However, following the reduction in size of the input matrix, u_0' defined in Equation 4.4, the requirement for memory is significantly relaxed. Although the efficiency of memory afforded by this technique is straight forward to calculate and has been highlighted by other authors, [14] there has been little discussion on the numerical efficiency of the technique.

It is natural to assume that the pre-aliasing step will translate to a significant computational saving to match the saving in memory. However, the saving in computational load, while still significant is much less than saving in terms of memory. The DFT is typically calculated using the fast Fourier transform (FFT) algorithm, [61] which will have a number of operations in the order of $O(M_x M_y \log_2(M_x M_y))$ for radix 2 FFT. Thus, if $(M_x, M_y) \ll (N_x, N_y)$ it can be expected that there will be a significant saving in terms of computational time for Steps 2-4 defined in Section 4.2.3. However, this is partially offset by the additional computational load introduced with with pre-aliasing

in the revised Step 1; the greater the saving in Steps 2-4, the greater the additional load in Step 1 due to the larger values of P_x and P_y in Equation 4.5. However, if the purpose is to apply the ASM repeatedly to the same input in order to calculate the field over a long sequence of different distances, then far greater computational savings can be made. For this case, Step 1 is implemented only once, while steps 2-4 are repeated iteratively. The computational efficiency of the ASM with and without the

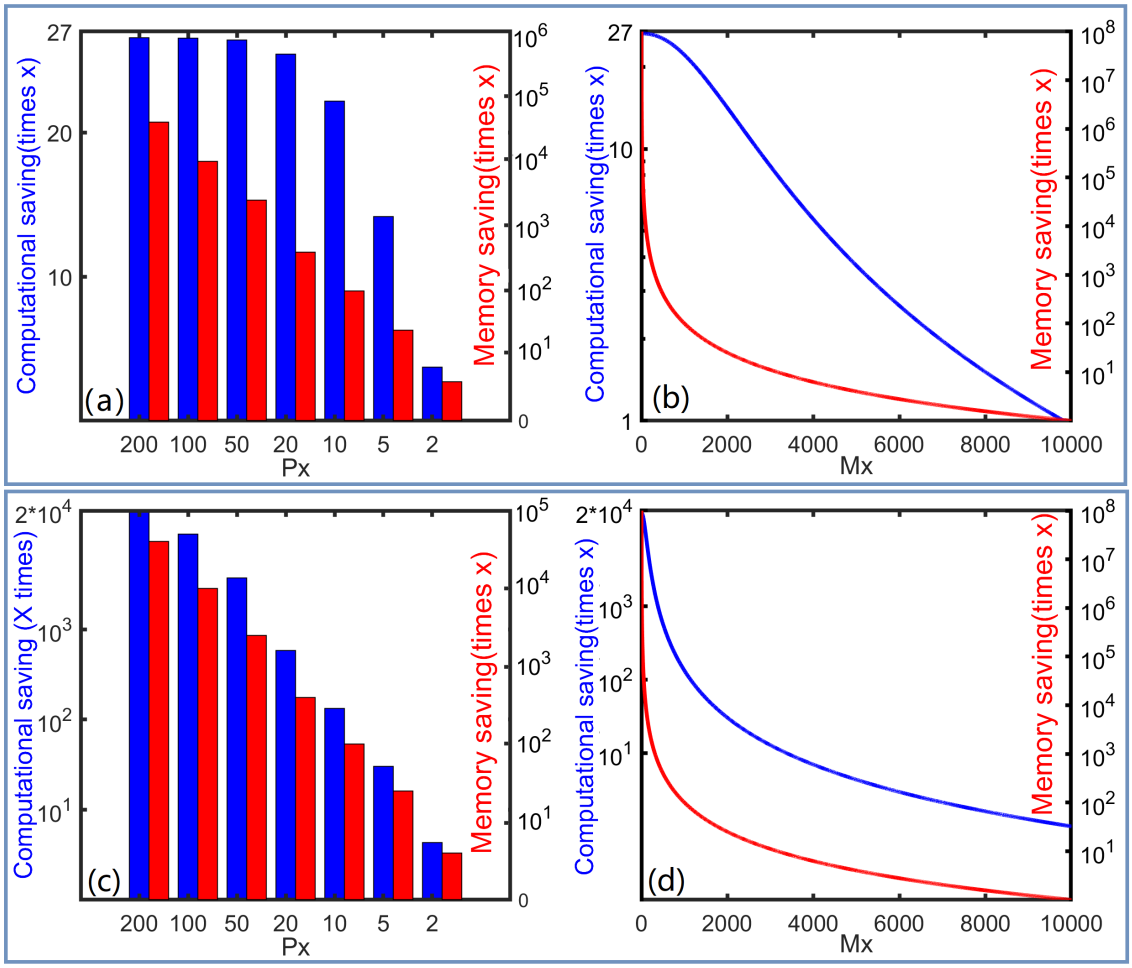

Figure 4.2: Computational and memory savings provided by pre-aliasing the input to the ASM where $N_x = N_y = 10000$; (a) computational and memory saving for the 2D case for a range of values P_x and (b) the same result over a continuous range of M_x; (c) and (d) show the same set of results for the 3D case where $N_z = 1000$.

pre-aliasing step can be quantitatively compared. With pre-aliasing, Step 1 will require a number of operations in the order of $O(P_x P_y M_x M_y)$. Therefore, the total number of operations without pre-aliasing is given by $O(N_x N_y \log_2(N_x N_y))$ and with pre-aliasing this becomes $O(P_x P_y M_x M_y + M_x M_y \log_2(M_x M_y))$. When the ASM is applied to the same input repeatedly in order to calculate the 3D diffraction distribution, the comparison of these efficiencies changes dramatically. Therefore, the total number of operations without pre-aliasing is given by $O(N_z N_x N_y \log_2(N_x N_y))$ and with pre-aliasing this becomes $O(P_x P_y M_x M_y + N_z M_x M_y \log_2(M_x M_y))$. To illustrate the practical importance of the pre-aliasing step we examine the case where then input u_0 is sampled

with $N_x = N_y = 10000$ samples. Taking $M_x = M_y = 50, 100, 200, 500, 1000, 2000, 5000$ (i.e. $P_x = P_y = 200, 100, 50, 20, 10, 5, 2$) the memory saving and the computational saving are shown in Fig. 4.2 (a); for the first five cases the computational saving is approximately the same at $\approx \times 22 - 26$. The memory saving, which is plotted on a $\log_{10}$ scale, is significantly higher and varies over a much wider range from 40000 for $P_x = 200$ down to 4 for $P_x = 2$. These savings are plotted over a continuous range of M_x in Fig. 4.2 (b). For the three-dimensional case, significantly greater computational savings are made, while memory saving remains the same. Taking the same parameters as above and now taking $N_z = 1000$, the savings are illustrated in Fig. 4.2 (c). In this case the computational saving ranges from $\times 20000$ for $P_x = 200$ to $\times 4$ for $P_x = 2$. These savings are plotted over a continuous range of M_x in Fig. 4.2(d).

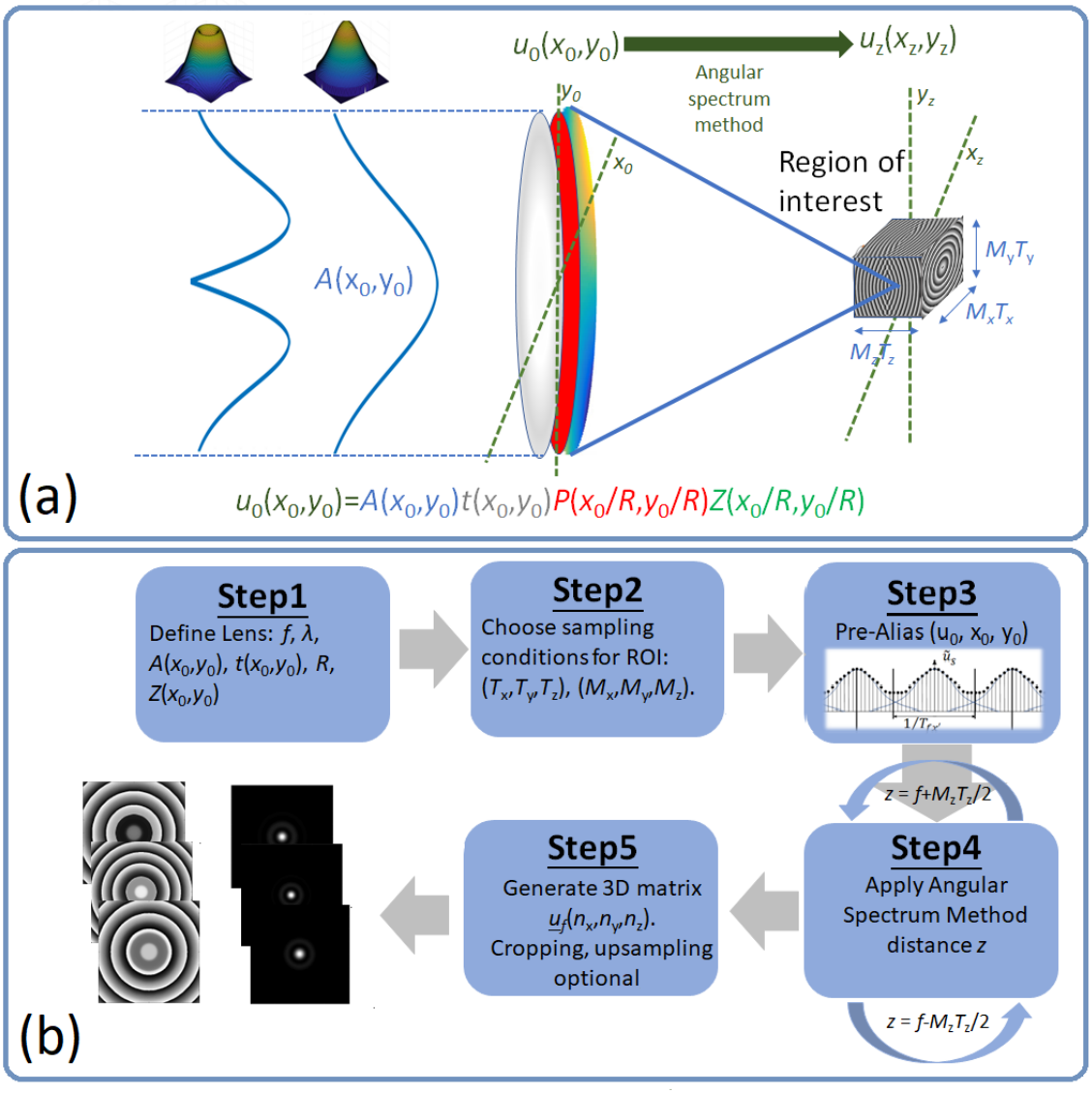

Figure 4.3: First paft of Fig. 4.4

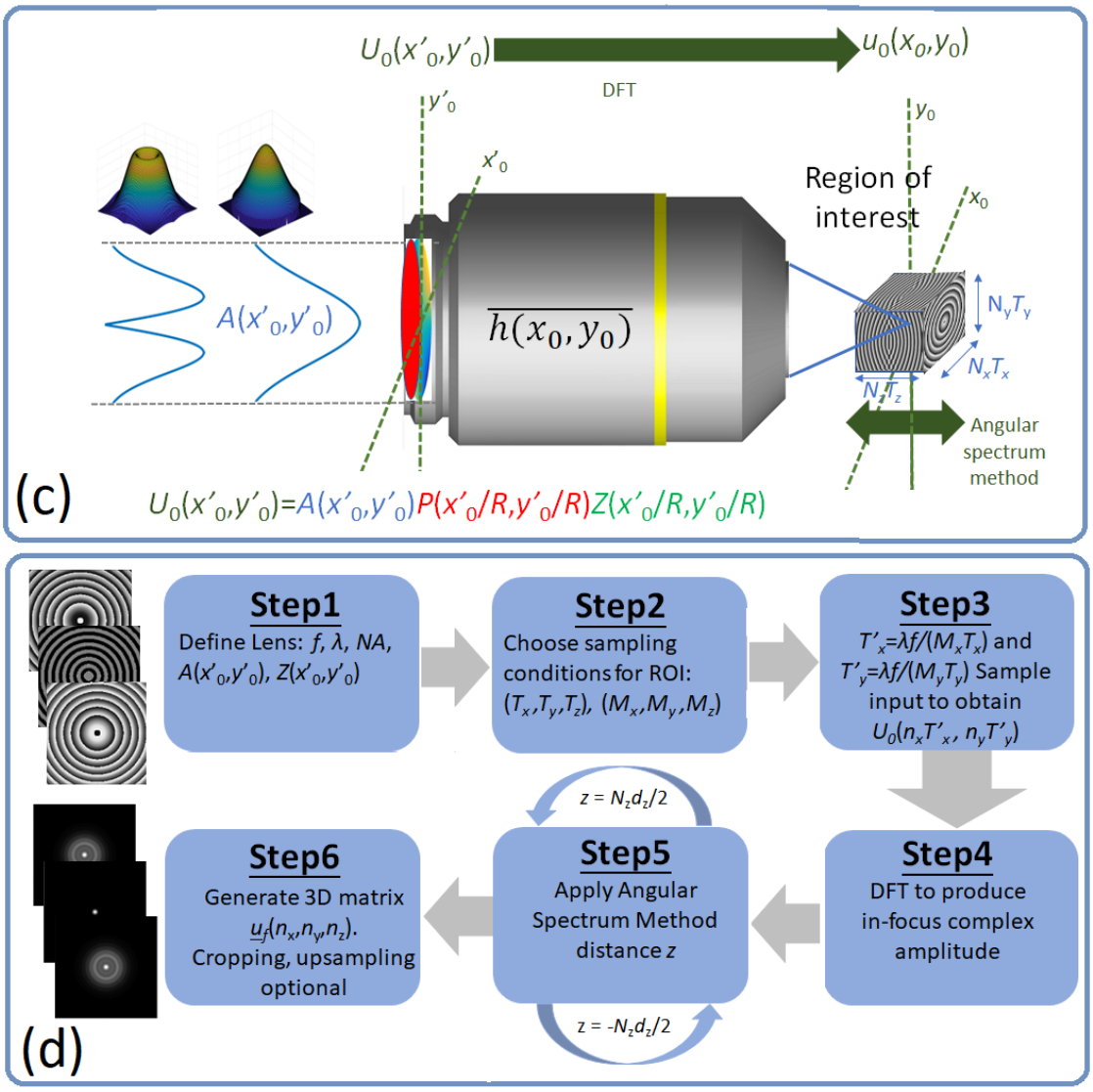

Figure 4.4: Illustration of the two algorithms developed in this chapter: (a) For the case of thin lens, the input to the algorithm is $u_0(x_0, y_0)$, which is the product of the complex thin lens transmission function t shown in grey, the laser profile A shown in blue, the lens pupil function of radius R shown in red, and the complex aberration term Z. The ASM is the key component in the algorithm and is iteratively applied to generate a 3D matrix of samples that represent the focused wavefield over a volume of interest; (b) The algorithm comprises five steps, which are detailed in the text. Note that Step 3 is a 'pre-aliasing' step whereby multiple copies of the input are shifted and added. This step is applied only once and serves to reduce down the number of sampled fed as input to the iteratively applied ASM; (c) For the case of ideal lens, the input to the algorithm is $U_0(x'_0, y'_0)$, which is the product of the laser profile A, the lens pupil function of radius R and the complex aberration term Z. A first DFT operation calculates the focused plane and following this the ASM is applied iteratively to generate a 3D matrix of samples that represent the focused wavefield over a volume of interest; (d) Algorithm 2 comprises six steps, which are detailed in the text.

4.3 Numerical Calculation of Focused Wavefields in the Paraxial Approximation

In this section, we outline two methods for calculating the complex amplitude of a focused wavefield in a three-dimensional volume which are illustrated in Fig. 4.3 and Fig. 4.4. Both methods use the ASM over a sequence of distances in order to generate a point cloud of complex values over a uniform three-dimensional grid. The first method is based on the thin lens approximation, whereby the lens can be described as a simple two-dimensional (2-D) phase-only function. The second method is based on modeling the effect of the lens as a simple pupil function in the Fourier domain for the case of coherent imaging.

4.3.1 The Thin Lens Case: Algorithm 1

The thin lens approximation [62] is commonly used within the paraxial approximation, to model a lens as a simple 2-D phase transformation, where the phase relates to the spatially variant delay imparted by the lens due to a varying thickness and/or refractive index. This approximation is based on the position of a ray that is incident on the lens in (x, y) being identical to the position of the ray as it leaves the lens; this requires negligible propagating of the ray within the lens. The phase transformation of the lens is given by:

$$t(x,y) = \exp\left[\, jk\phi(x,y)\right] P\left(\frac{x}{R}, \frac{y}{R}\right) \tag{4.10}$$

where the wave vector $k = 2\pi/\lambda$ and λ is the wavelength of the light; $\phi(x,y)$ represents the delay imparted by the lens at the position (x,y) in terms of a wave cycle. The aperture of the lens is defined by the pupil function $P(x,y)$, which is defined as unity inside the unit circle and zero elsewhere:

$$\begin{aligned} P(x,y) &= 1 \forall (x^2 + y^2) \leq 1 \\ P(x,y) &= 0 \forall (x^2 + y^2) > 1 \end{aligned} \tag{4.11}$$

Assuming, an incident plane wave of complex amplitude $A(x,y)$ is input to the plane of the lens, the resultant wavefield immediately after the lens is given by:

$$u_0(x_0, y_0) = A(x_0, y_0)\, t(x_0, y_0) \tag{4.12}$$

The wavefield u_0 will converge at the focal point of the lens, i.e. at a distance $\approx f$, where f is the focal length of the lens. The physical case is illustrated in Fig. 4.3 (a) and the corresponding algorithm to calculate the three dimensional complex field in the region of this convergence is illustrated in Fig. 4.3 (b) as a sequence of five steps, which are further defined below.

1. The first step is to define the wavelength λ and the continuous function $u_0(x_0, y_0)$; the user selects a lens phase delay function $\phi(x_0, y_0)$ as well as an input laser mode $A(x_0, y_0)$. In this selection, the radius of aperture, R, and the focal length,

f, must be defined. Finally, the user selects the weights w_n for the Zernike polynomials in the aberration function. Therefore, the input function is defined: $u_0(x_0, y_0) = A(x_0, y_0)P\left(\frac{x_0}{R}, \frac{y_0}{R}\right)\exp\left[jk\phi(x, y)\right]$

2. The second step is to define the size of the region of interest in terms of the sampling intervals, (T_x, T_y, T_z) and the number of samples to used in this region (M_x, M_y, M_z). As a maximum, the values of T_x and T_y should satisfy the Nyquist condition for $u_0(x_0, y_0)$; however, the user may opt for much smaller values than those imposed by the Nyquist limit. The value for T_z should also be selected to be sufficiently small to avoid aliasing in the z direction; this will depend on the convergence rate of wavefield. The values for (M_x, M_y, M_z) should be chosen to be large enough to fully capture the support of converged light in the region of interest.

3. The third step is to sample the input $u_0(x_0, y_0)$ with sampling intervals (T_x, T_y) and to pre-alias in order to generate the discrete function $u'_0(n_x, n_y)$ made up of $M_x \times M_y$ samples as defined in Equation 4.4. In total, the number of samples of u_0 that must be considered is given by $N_x \times N_y$, where these are the smallest integers that satisfy $N_x \geq 2R/T_x$ and $N_y \geq 2R/T_y$. However, it is not necessary to compute all of these values and to store these in memory; rather, $(2P_x + 1) \times (2P_y + 1)$ smaller batches of size $M_x \times M_y$ can be computed and superimposed as defined in Equation 4.4 and Equation 4.5. Some zero-padding of u_0 may be necessary if $(2P_x + 1) \times (2P_y + 1) \times M_x \times M_y > N_x \times N_y$; in such case N_x and N_y should be increased with zeros.

4. The fourth step is to apply the ASM to u'_0 repeatedly over a sequence of M_z distances given by $z_i = f + n_z T_z$ where the integer n_z takes the values: $n_z = -M_z/2 \rightarrow M_z/2 - 1$, and to store the resulting complex values in a three dimensional matrix $u_f(n_x, n_y, n_z)$ of size $(M_x \times M_y \times M_z)$. It should be noted that a large 3D matrix may present difficulties for storing in memory. In some cases, it may not be necessary to store the entire matrix. For example, in order to calculate the center slice for $y = 0$ one need only store a 2D matrix for which the n_y value is fixed at zero, and only a single 1D vector needs to be stored after each iteration of the ASM. In other cases where the full 3D volume is required and temporary memory is limited, completed portions of the matrix can be intermittently stored to permanent memory.

5. The fifth step is optional. This involves cropping a region of the 3-D matrix and upsampling in one or more dimensions. This may be applied if it is desirable to zoom in on a small area of interest. This can be implemented by calculating the 3D DFT of the cropped matrix using the FFT algorithm, zero-padding the result as desired and applying an inverse DFT.

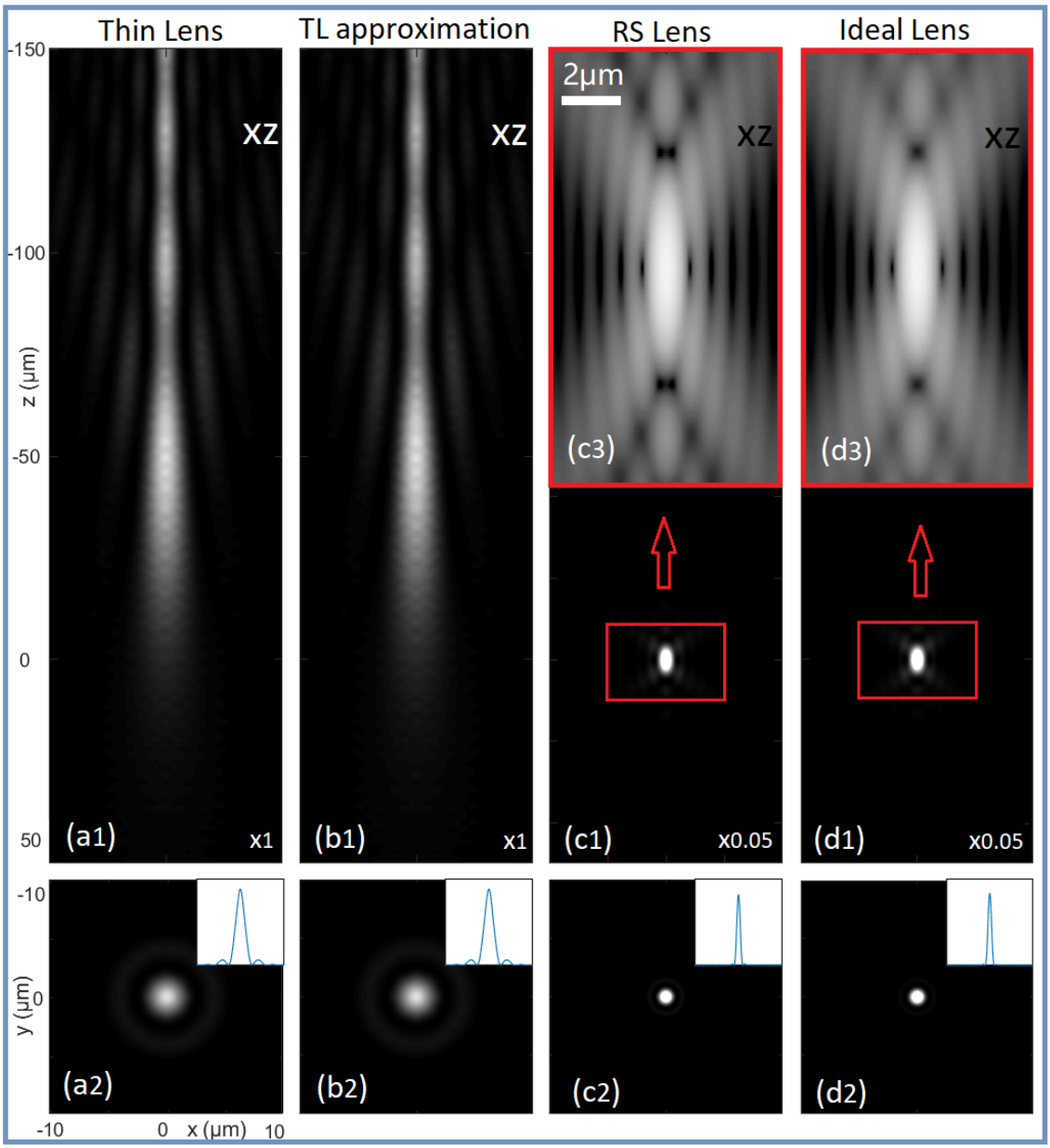

Figure 4.5: The intensity profiles of convergent light from four different lenses with the same numerical aperture and with same top-hat uniform illumination, $A = 1$. For the case of the thin lens defined in Equation 4.13, (a1) and (a2) are slices through the volume for $y = 0$ (i.e. the XZ plane) and $z = f - 49\mu$m, which was found to be the XY plane with the brightest spot. Also shown in (a2) is an inset image of the profile of this spot. In (b1) and (b2) the same figures are shown for the second thin lens with purely quadratic phase given in Equation 4.16. The profiles for these two cases are almost identical owing to the high accuracy in the approximation of obtaining t_2 from t_1. For the case of the third thin lens defined in Equation 4.17, (c1) and (c2) are slices through the volume for $y = 0$ and $z = f$; for this case, the light converges to a much smaller spot due to the aspheric profile of the lens function. Finally, the same slices are shown in (d1) and (d2) for the case of the ideal lens. Remarkably, the intensity profiles obtained using the RS lens and the ideal lens are almost identical. Note that the scale used for images (c1), (c2), (d1), and (d2) is 0.05 times that used for the other four images due to the significantly higher intensity values.

Some initial results from this algorithm are provided in Fig. 4.5. Here, the three-dimensional field is calculated for three different thin lens functions; for all cases, 'top-hat' illumination is assumed, $A(x,y)=1$, such that $u_0=t$, and for all cases the aperture of the lens was given by $R=2.5mm$ and the wavelength of light was defined to be $\lambda-1064$ nm. The first lens is defined as follows:

$$t_1(x,y)=\exp\left[jk(\Delta_0+\Delta(x,y)(n-1))\right]P\left(\frac{x}{R},\frac{y}{R}\right) \tag{4.13}$$

where $\Delta_0=\Delta_{01}+\Delta_{02}+\Delta_{03}$ is the maximum thickness of the lens, $\Delta(x,y)$ is the thickness at coordinate (x,y) and $n=1.515$ is the refractive index of the material (BK7), and $R=2.5mm$. The thickness function is defined to be:

$$\Delta(x,y)=\Delta_0-R_1\left(1-\sqrt{1-\frac{x^2+y^2}{{R_1}^2}}\right)+R_2\left(1-\sqrt{1-\frac{x^2+y^2}{{R_2}^2}}\right) \tag{4.14}$$

where $R_1=3mm$ and $R_2=-4mm$ are the radius of curvatures of the two lens surfaces. For such a lens, the focal length can be approximated to be given by:

$$\frac{1}{f}=(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right) \tag{4.15}$$

which for the values of R_1 and R_2 selected here is $f=-3.328$ mm. For this case the numerical aperture of the lens can be estimated as $NA=0.6$. For further details on this thin lens definition we refer the reader to Ref 62 Chapter 5. The parameters selected in Step 2 of the algorithm are as follows $N_x=N_y=50000$, $T_x=T_y=T_z=0.1\mu$m, $M_x=M_y=M_z=2000$. To save memory only (i) the centre slice corresponding to $y=0$ was stored producing an XZ image of $M_x\times M_z$ samples, and (ii) one XY image was stored at a single value of n_z containing $M_x\times M_y$ samples. Fig. 4.5 (a1) shows a cropped region of the XZ image with $\times10$ interpolation in both dimensions, and (a2) shows the XY image at distance $z=f-49\mu$m also with $\times10$ interpolation in both dimensions. The profile of the focused spot is shown in the inset image. The same set of results are shown in In Fig. 4.5 (b1) and (b2) for a second thin lens function with a simpler quadratic variation in phase:

$$t_2(x,y)=\exp\left[-\frac{j\pi}{\lambda f}(x^2+y^2)\right]P\left(\frac{x}{R},\frac{y}{R}\right) \tag{4.16}$$

We note that t_2 can be obtained from t_1 based on the paraxial approximation. The same values for f and r are used as for t_1 in the simulation. The results for t_1 and t_2 are almost identical, highlighting the validity of the paraxial approximation in this case. A third thin-lens function, which we name the Rayleigh-Sommerfeld Lens or RS Lens for short, was investigated where:

$$\begin{aligned} t_{RS}(x,y)&=\exp\left[-jkr_f(x,y)\right]P\left(\frac{x}{R},\frac{y}{R}\right)\\ r_f(x,y)&=\sqrt{x^2+y^2+f^2}\end{aligned} \tag{4.17}$$

This definition is borrowed from the Rayleigh Sommerfeld (RS) diffraction integral,

which relates a field $u_0(x, y)$ to the diffracted field at distance z as follows:

$$u_z(x_z, y_z) = -\frac{1}{2\pi} \times \int_{-\infty}^{\infty} u_0(x_0, y_0) \frac{\delta}{\delta z} \left[\frac{\exp(-jkr_z(x_z - x_0, y_z - y_0))}{r_z(x_z - x_0, y_z - y_0)} \right] \mathrm{d}x_0 \mathrm{d}y_0 \tag{4.18}$$

This integral can be rewritten as a convolution as follows:

$$\begin{aligned} u_z(x_z, y_z) &= u_0(x_0, y_0) * h_z(x_0, y_0) \\ h_z(x_0, y_0) &= -z \frac{\exp\left[jkr_z(x_0, y_0)\right]}{2\pi r_z^2(x_0, y_0)} \left[jk - \frac{1}{r_z(x_0, y_0)} \right] \end{aligned} \tag{4.19}$$

The complex term in square brackets has an approximately constant phase angle of $\pi/2$. Thus, a point source of light, represented by a Dirac-delta functional (see Appendix B,Fig. 1) will produce a diffraction pattern at distance f with a phase function given by the conjugate of t_{RS}. Such a lens function, we refer to as a Rayleigh-Sommerfeld (RS) lens going forward, can therefore be expected to produce a highly focused spot after propagating a distance $z = f$. Selecting $f = 3.3$ mm and choosing all of the same sampling parameters as for the previous two cases, the resulting intensity is shown in Fig. 4.5 (c1) and (c2). The spot in this case is significantly more tightly focused as a consequence more intense than for the previous cases, and is scaled $\times 0.05$ for comparison. The wide range of intensity allows only a small region around the focus using a linear scale; hence, an inset image shows the logarithm of the intensity to reveal the light pattern in the region of the focus in more detail. The inset image in Fig. 4.5 (c2) shows the profile of the focused spot, which is significantly narrower and $\times 0.05$ more intense that the corresponding profiles shown in Fig. 4.5 (a2) and (b2).

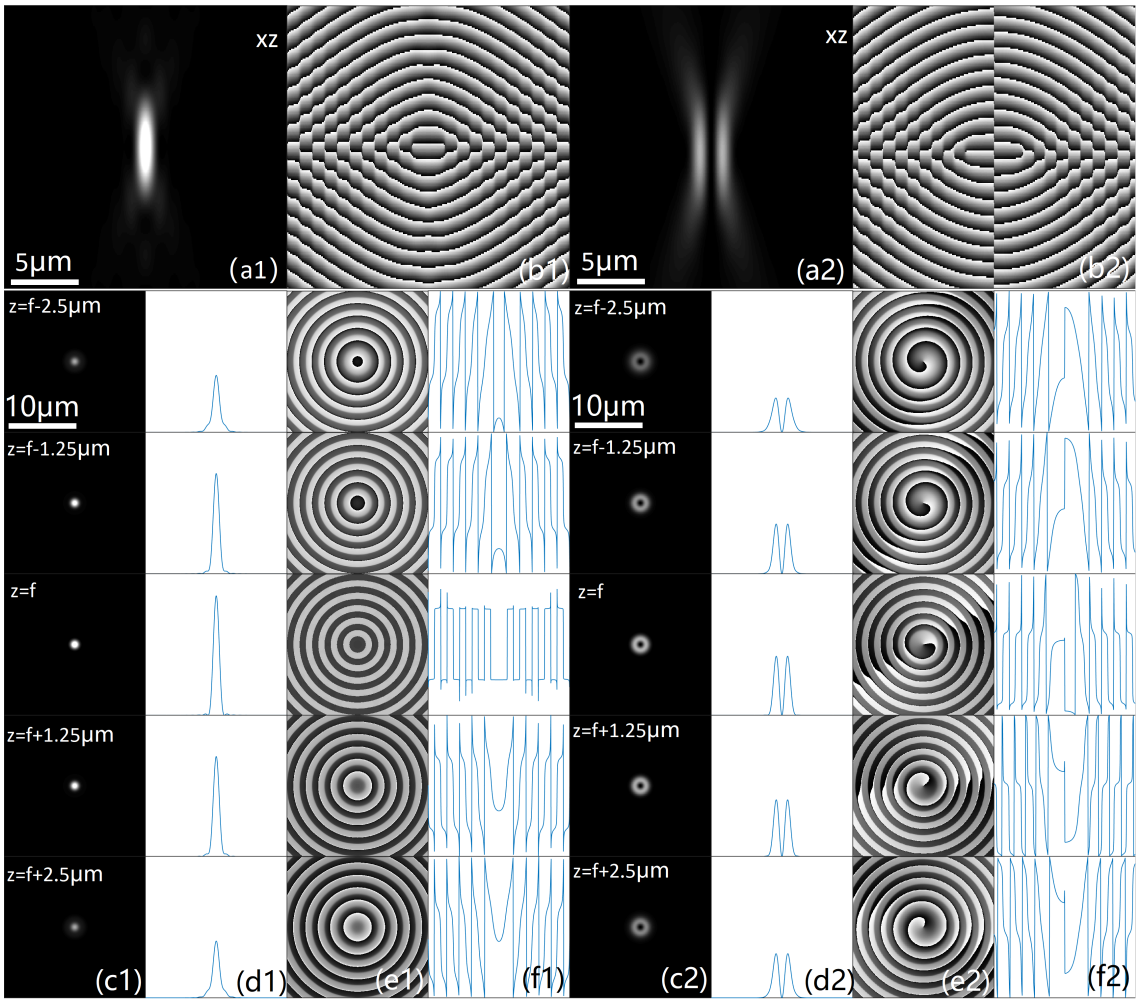

Figure 4.6: Phase and intensity of focused spot for the case of the Rayleigh Sommerfeld lens defined in Equation 4.17 for the case of both Gaussian and TEM_{01} laser spatial modes. All the images were calculated using Algorithm 1 for the thin lens case; (a1) and (b1) are the intensity and phase of the centre XZ slice for $y = 0$ for the case of a Gaussian beam; Video1 shows the XZ intensity/phase slice as a function of different values of y. (c1) shows the intensity of the XY plane at five different distances indicated in the lower left of the figure; (d1) shows the profile of the focused spot in the five planes. Note the use of different scaling parameters in both (c1) and (d1) as indicated in the upper right parts of the figure due to the wide range of values; (e1) and (e2) show the corresponding phase image and profiles. Video2 shows the XY intensity/phase slice as a function of different values of z. A similar set of results are shown in (a2)-(f2) for the case of a TEM_{01} laser mode. The characteristic doughnut shape of the focused spot is seen to expand with defocus, and the spiral phase pattern is seen to twist. Remarkably, the XZ-slice of the phase distribution is identical to that for the Gaussian case, except for a half wavelength phase shift for the right side of the image. Video3 and Video4 show the variation in the XZ and XY planes as for the Laguerre Gaussian case. The same results for the ideal lens calculated using Algorithm 2 are provided in Appendix B, Fig. 1

4.3.2 The Ideal Lens: Algorithm 2

Assuming an incident plane wave of complex amplitude $A(x, y)$ we define the effect of an ideal lens with focal length f and numerical aperture NA in terms of the optical Fourier transform: [62]

$$\begin{aligned} u_0(x_0, y_0) &= \mathrm{OFT}\{U_0(x_0', y_0')\}(x_0, y_0) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} U_0(x_0', y_0') \exp - \left[\frac{2\pi}{\lambda f}(x_0 x_0' + y_0 y_0')\right] \mathrm{d}x_0' \mathrm{d}y_0' \end{aligned} \tag{4.20}$$

where OFT is the operator for the optical Fourier transform mapping between spatial coordinates (x_0', y_0') and (x_0, y_0) at the focal planes of the lens and where:

$$U_0(x_0', y_0') = A(x_0', y_0') P\left(\frac{x_0'}{r\lambda f}, \frac{y_0'}{r\lambda f}\right) \tag{4.21}$$

Here, u_0 is the wavefield that is produced at the focal length of the lens. The pupil function of the lens, P (defined in Equation 4.11) has radius $r\lambda f$ where r is defined in terms of the numerical aperture as follows:

$$r = \frac{NA}{\lambda} \tag{4.22}$$

Setting, $A = 1$ for the case of an ideal infinite plane wave, the result is in the form of a Bessel function of the first kind [62]: $\pi r^2 J_1(2\pi r\rho)/(\pi r\rho)$ where ρ is the radial coordinate. The intensity of this pattern is known as the Airy pattern and the width of the central lobe is given by $1.22\lambda/NA$ which is the well-known limit of optical resolution for coherent imaging. Therefore, two ideal microscope objectives with the same NA but different focal lengths, f, will produce identical Airy patterns at their respective focal length for the same uniform plane wave illumination, but would have different input physical pupil radius given by $R = fNA$. Since the pupil plane of the microscope objective is typically close to the back aperture, we can take the radius to be equivalent to the physical radius of the lens back aperture.

The wavefield u_0 will converge at the focal point of the lens, and we proceed to define an algorithm to calculate the three-dimensional volume of light in this region. This algorithm is illustrated in Fig. 4.4 as a sequence of six steps, which are further defined below.

1. The first step is to define the wavelength λ, the focal length f, and numerical aperture NA of the lens. This will in turn define the input aperture to be given by a pupil function of radius $R = fNA$; the user also selects an input laser mode $A(x_0', y_0')$. Therefore, the input function is defined: $U_0(x_0', y_0') = A(x_0', y_0') P\left(\frac{x_0'}{R}, \frac{y_0'}{R}\right)$

2. The second step is to define the size of the region of interest in terms of the sampling intervals (T_x, T_y, T_z) as well as the number of samples to used in this region (M_x, M_y, M_z). Since $u_0(x_0, y_0)$ will be defined by $M_x \times M_y$ samples with uniform sampling intervals of T_x and T_y, the optical Fourier transform (OFT), $U_0(x_0', y_0')$

will be also defined by $M_x \times M_y$ samples but with sampling intervals given by $T'_x = \lambda f/(M_x T_x)$ and $T'_y = \lambda f/(M_y T_y)$. As for the previous algorithm, the sampling intervals (T_x, T_y, T_z) must satisfy the Nyquist criterion in all three dimensions. In this case, however, it is a little simpler to define the condition for x_0 and y_0: $M_x T'_x \geq R$ and $M_y T'_y \geq R$.

3. The third step is to sample the continuous input function $U_0(x'_0, y'_0)$ (this is the field passing through the back aperture of the MO) with sampling intervals (T'_x, T'_y) over an extent of $(M_x T'_x, M_y T'_y)$ such that $M_x \times M_y$ samples are obtained of the form: $U_0(n_x T'_x, n_y T'_y)$ where $n_x = -N_x/2 \rightarrow N_x/2-1$ and $n_y = -N_y/2 \rightarrow N_y/2-1$.

4. The DFT of U_0 is calculated to produce the discrete function $u_0(n_x T_x, n_y T_y)$ over an extent $(M_x T'_x, M_y T'_y)$. This is the in-focus image of the laser spot and will serve as input to the ASM algorithm in order to calculate complexity in the planes before and after.

5. Similar to the previous algorithm, the fifth step is to apply the ASM to u_0 repeatedly over a sequence of M_z distances given by $z_i = f + n_z T_z$ where the integer n_z takes the values: $n_z = -M_z/2 \rightarrow M_z/2-1$, and to store the resulting complex values in a three dimensional matrix $u_f(n_x, n_y, n_z)$ of size $(M_x \times M_y \times M_z)$. Similar difficulties may present in terms of storage as for the previous algorithm.

6. As for the previous case the final step is optional and involves cropping a region of the 3-D matrix and upsampling in one or more dimensions as desired.

An initial result from this algorithm is shown in Fig. 4.5 (d1-d3). Here, for Step 1 the following selections were made: As for the previous cases, 'top-hat' illumination was used, i.e. $A(x, y) = 1$, such that $U_0(x'_0, y'_0) = P(x'_0/fNA, y'_0/fNA)$, where f and NA were chosen to be 3.3mm and 0.6, respectively resulting in an aperture size of $R = 2.5\mu$m. As for the previous cases, the wavelength of light was defined to be $\lambda - 1064$ nm. The parameters selected in Step 2 of the algorithm are as follows $M_x = M_y = M_z = 2000$, $T_x = T_y = T_z = 0.1\mu$m. As for the previous examples, to save memory only one XZ slice and one XY were stored; cropped areas of both are shown in Fig. 4.5 (d1) and (d2) with $\times 10$ interpolation. The profile of the focused spot is shown in the inset image in Fig. 4.5 (d1). The spot in this case is similar to that for the case of the RS lens. The magnified image in Fig. 4.5 (d3) shows the logarithm of the intensity of the XZ plane in which it can be seen that there are striking similarities to the same image for the RS lens.

4.4 Including the laser spatial mode

In the discussion this far, top hat illumination has been assumed for which $A = 1$ and the laser power, P is given no consideration. In practice, the laser will have a particular spatial mode that defines its complex spatial profile and a total power, both of which should be taken into account. Many applications employ lasers with a TEM_{00} or TEM_{01} mode that are characterised by a Gaussian or more generally, by a Laguerre-Gaussian (LG) distribution. While 'top-hat illumination would produce the tightest

laser spot, most applications will require maximal use of the laser power and as such the laser cannot be widely expanded to achieve a uniform 'top-hat' profile across the lens aperture; it is common for the laser to be expanded such that the lens aperture apodizes the Gaussian/LG distribution at a particular radius in the distribution, which is deemed to provide the most acceptable compromise between the loss in laser power and and the enlargement of the spot size. Examples that apply this approach include optical trapping, confocal microscopy, and laser cutting. We examine two commonly used laser spatial modes here: TEM_{00} or TEM_{01}. The former is the most common laser mode and in its simplest form for which the laser has neglible expansion, the electric field is defined by a Gaussian distribution as follows:

$$A_{TEM_{00}}(x,y) = \pi^{-1/4}\sqrt{\frac{P}{\sigma}}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4.23}$$

where σ is the standard deviation, which relates to the width of the Gaussian function. The intensity of this field is given by:

$$|A_{TEM_{00}}|^2 = \frac{P}{\sqrt{2\pi}\sigma'}\exp\left(-\frac{x^2+y^2}{2\sigma'^2}\right) \tag{4.24}$$

where the standard deviations are related as follows: $\sigma = \sqrt{2}\sigma'$. The most commonly chosen cutoff point for the laser at the back aperture of a microscope objective is the $1/e^2$ width, which is the beam radius at which the intensity has dropped to $1/e^2$ or approximately 14 of its maximum value (by the same definition, this is the $1/e$ width for amplitude); see Chapter 17 in Ref 104 for further details. This radius is equal to $2\sigma'$ or equivalently $\sqrt{2}\sigma$. Therefore, for a lens aperture with radius R, a beam which is characterised with a $1/e^2$ width given by σ' is expanded by a factor M such that $2M\sigma' = 2\sigma'' = R$. For the RS lens that was simulated in the previous section (for which f = 3.3mm and R = 2.5mm), a Gaussian profile was applied at the input, i.e. $u_0(x_0,y_0) = t_{RS}(x_0,y_0)A_{TEM_{00}}(x_0,y_0)$ where σ was set equal to $R/\sqrt{2}$, where R is the radius of the lens. The parameters of the simulation using Algorithm 1 are the same as those defined in Section 4.3.1. The intensity and phase distributions of the field in the XZ plane is shown in Fig. 4.6 (a1) and (b1) over a size of 20μ m×20μm. The intensity and phase in the XY plane over an area of 10μ m×10μm are shown at five different distances in Fig. 4.6 (c1) and (e1) with profiles of these distributions given in Fig. 4.6 (d1) and (f1). The intensity images have been scaled by the factors given in Fig. 4.6 (d1) in order to preserve the image contrast.

The Gaussian distribution defined in Equation 4.24 is ideal and does not take into account expansion of the laser beam or the change in the phase distribution as a function of propagation. The TEM_{00} laser mode is more accurately described as a a special case of the more general Laguerre-Gaussian mode. The amplitude distribution of the Laguerre-Gaussian mode is uniquely defined by the beam waist width $w(0)$ and the Rayleigh range, z_R, as well as integers l (the azimuthal index) and p (the number of

radial nodes) as follows:

$$A_{LG_{pl}}(x,y) = \sqrt{\frac{2p!}{\pi(p+|l|)!}}\frac{1}{w(z)}\left[\frac{\sqrt{2(x^2+y^2)}}{w^2(z)}\right]^{|l|}\exp\left[\frac{-(x^2+y^2)}{w^2(z)}\right] \times L_p^{|l|}\left[\frac{2(x^2+y^2)}{w^2(z)}\right]\exp\left[jl\tan^{-1}\left(\frac{y}{x}\right)\right]\exp\left[\frac{ikz(x^2+y^2)}{2(z^2+z_R^2)}\right] \times \exp\left[-j(2p+|l|+1)\tan^{-1}\left(\frac{z}{z_R}\right)\right] \tag{4.25}$$

where the $1/e$ width of the Gaussian term is given by:

$$w(z) = w(0)\sqrt{\frac{z^2+z_R^2}{z_R^2}} \tag{4.26}$$

The argument of the last exponential term in Equation 4.25 is the Guoy phase. The term $L_p^{|l|}$ is the associated Laguerre-Gaussian polynomial obtained from Laguerre polynomials:

$$L_p^{|l|}[a] = (-1)^{|l|}\frac{\mathrm{d}^{|l|}}{\mathrm{d}a^{|l|}}L_{p+|l|}[a] \tag{4.27}$$

We note that for $l = p = 0$ and for the case of negligible beam expansion $z_R \to \infty$ the above expression reduces to the form of the Gaussian mode defined in Equation 4.24 where $w(0) = 2\sqrt{2}\sigma$. In order to normalise the amplitude $A_{LG_{pl}}$ such that the integral of the intensity distribution is equal to P, the laser power, the amplitude distribution can be scaled by a parameter α, where:

$$\alpha = \frac{\sqrt{P}}{\sqrt{\iint\left|A_{LG_{pl}}(x,y)\right|^2\mathrm{d}x\mathrm{d}y}} \tag{4.28}$$

Once again using the same RS lens, a LG profile was applied at the input, i.e. $u_0(x_0,y_0) = t_{RS}(x_0,y_0)_{LG_{01}}(x_0,y_0)$ where $w(0)$ was set equal to R and $z_R = w_0^2k/2$. Once again the parameters of the simulation using Algorithm 1 are the same as those defined in Section 4.3.1. The intensity and phase distributions of the field in the XZ plane is shown in Fig. 4.6 (a2) and (b2) over a size of 20µm×20µm. For this case, there exists a clear discontinuity in the phase at $(x_z, y_z) = (0,0)$ The intensity and phase in the XY plane over an area of 10µm×10µ are shown at five different distances in Fig. 4.6 (c2) and (e2) with profiles of these distributions given in in Fig. 4.6 (d2) and (f2). The intensity images have been scaled by the factors given in Fig. 4.6 (d1) relative to the in-focus Gaussian spot intensity in Fig. 4.6 (c1). The spiral phase shape of the distribution is clearly evident and appears to rotate with propagation distance.

Noll index [105]	$Z_n(x,y)$	Classical name
1	1	Constant
2	$2x$	Horizontal tilt
3	$2y$	Vertical tilt
4	$\sqrt{3}(2x^2+2y^2-1)$	Defocus
5	$\sqrt{6}(2xy)$	Oblique Astigmamatism
6	$\sqrt{6}(x^2-y^2)$	Vertical Astigmatism
7	$\sqrt{8}(3x^2y+3y^2-2y)$	Vertical Coma
8	$\sqrt{8}(3x^3+3xy-2x)$	Horizontal Coma
9	$\sqrt{8}(3x^2y-y^3)$	Vertical Trefoil
10	$\sqrt{8}(x^3-3xy^2)$	Oblique Trefoil
11	$\sqrt{5}(6x^4+12x^2y^2+6y^4\ldots$ $-6x^2-6y^2+1)$	Spherical
12	$\sqrt{10}(4x^4-3x^2+3y^2-4y^4)$	Vertical Secondary Astigmatism
13	$\sqrt{10}(8x^3y+8xy^3-6xy)$	Oblique Secondary Astigmatism
14	$\sqrt{10}(x^4-6x^2y^2+y^4)$	Vertical Quadrafoil
15	$\sqrt{10}(4x^3y-4xy^3)$	Oblique Quadrafoil

Table 4.1: Description of the first 15 Zernike polynomials, Z_n, in Catersian form. Here we have used the Noll index, commonly used in the Zemax classification.

4.5 Including lens aberrations

Aberration in a lens system is the failure of that system to produce an ideal image; this can be interpreted either in terms of geometrical ray optics, whereby all rays passing through a point in the input plane should also pass through a single point in the image plane, or in terms of wavefronts, whereby the spherical wavefronts emanating from a point in the input plane should be portions of spheres centred at a point in the image plane. Characterisation of aberrations can employ finite ray tracing through the optical system (see Ref 65) and wavefront aberration theory (see Ref 65 and Ref 106), which is often interpreted in terms of the Seidel aberrations. Both theories can be related using the characteristic function of Hamilton [65], but it should be noted that neither theory accounts for the effects of diffraction. In this chapter and the next chapter, we attempt to use non-linear ray tracing to account for the effects of aberration, including the effects of diffraction resulting from lens distortion.

The next step in our algorithm is to include the effect of lens aberration in the numerical model such that the lens can be distorted, and the resulting focal pattern can be examined in three dimensions. Rather than use the Seidel aberrations we employ the Zernike polynomials to account for lens aberration for both the thin lens case and the ideal lens case, albeit in different manners. We note that the Seidel aberrations could easily be substituted if desired. The Zernike polynomials are a set of orthogonal two-dimensional polynomials defined over the unit circle. These are particularly useful in describing the wavefront distortions that can exist for optical systems with

circular pupils due to the property that they form a complete and orthogonal basis set over the unit circle. Therefore, any given function defined over this unit circle, $W(x,y)$, can be represented as a weighted (infinite) sum of these polynomials as follows:

$$W(x,y) = \sum_{n=0}^{\inf} w_n Z_n(x,y) \tag{4.29}$$

where w_n represents the weights of the polynomials in the infinite summation and Z_n represent the 2D Zernike polynomials, where n is the Noll index [105]. The definition of the first fifteen polynomials, Z_n for $n:1\rightarrow 15$, are provided in Table 4.1 together with the descriptive name that is commonly used for this polynomial. For convenience, the polynomials in this table are defined as a function of x and y, and can be obtained by making appropriate substitutions into the corresponding table in Chapter 2, which is defined in terms of radius and angle as the classical approach.

When wavefront error exists in an optical system that contains aberration, a classical way to account for it is to describe its effect in terms of a phase delay distortion of the pupil function of the ideal lens system. An ideal diffraction-limited optical system can be described as having a frequency response or pupil function defined as the circle function, $P\left(x_0'/r\lambda f, y_0'/r\lambda f\right)$ as defined in Equations 4.20 and 4.21 in Section 4.3.2 where $P(x_0', y_0')$ is the unit circle over (x_0', y_0'). An aberrated optical system can be described in terms of the same pupil function but also with a spatially varying phase delay $kW(x_0', y_0')$ in the plane of the pupil function. Therefore, the pupil function, P' of an aberrated system is defined as follows:

$$P'(x_0', y_0') = P(x_0', y_0')Z(x_0', y_0') \tag{4.30}$$

where the phase delay term Z is defined as follows:

$$Z(x_0', y_0') = \exp\left[jkW(x_0', y_0')\right] \tag{4.31}$$

where W is defined in Equation 4.29. Thus, in the description of the Algorithm 2, which simulates an ideal lens system, Equation 4.21 is replaced with the following definition of the input function: [62]

$$U_0(x_0', y_0') = A(x_0', y_0')P\left(\frac{x_0'}{r\lambda f}, \frac{y_0'}{r\lambda f}\right)Z\left(\frac{x_0'}{r\lambda f}, \frac{y_0'}{r\lambda f}\right) \tag{4.32}$$

and the first step of the algorithm is rewritten as follows:

1. The first step is to define the wavelength λ, the focal length f and numerical aperture NA of the lens. This will in turn define the input aperture to be given by a pupil function of radius $R = fNA$; the user also selects an input laser mode $A(x_0', y_0')$. Finally, the user selects the weights w_n for the Zernike polynomials in the aberration function. Therefore, the input function is defined:
   $U_0(x_0', y_0') = A(x_0', y_0')P\left(\frac{x_0'}{R}, \frac{y_0'}{R}\right)Z\left(\frac{x_0'}{R}, \frac{y_0'}{R}\right)$

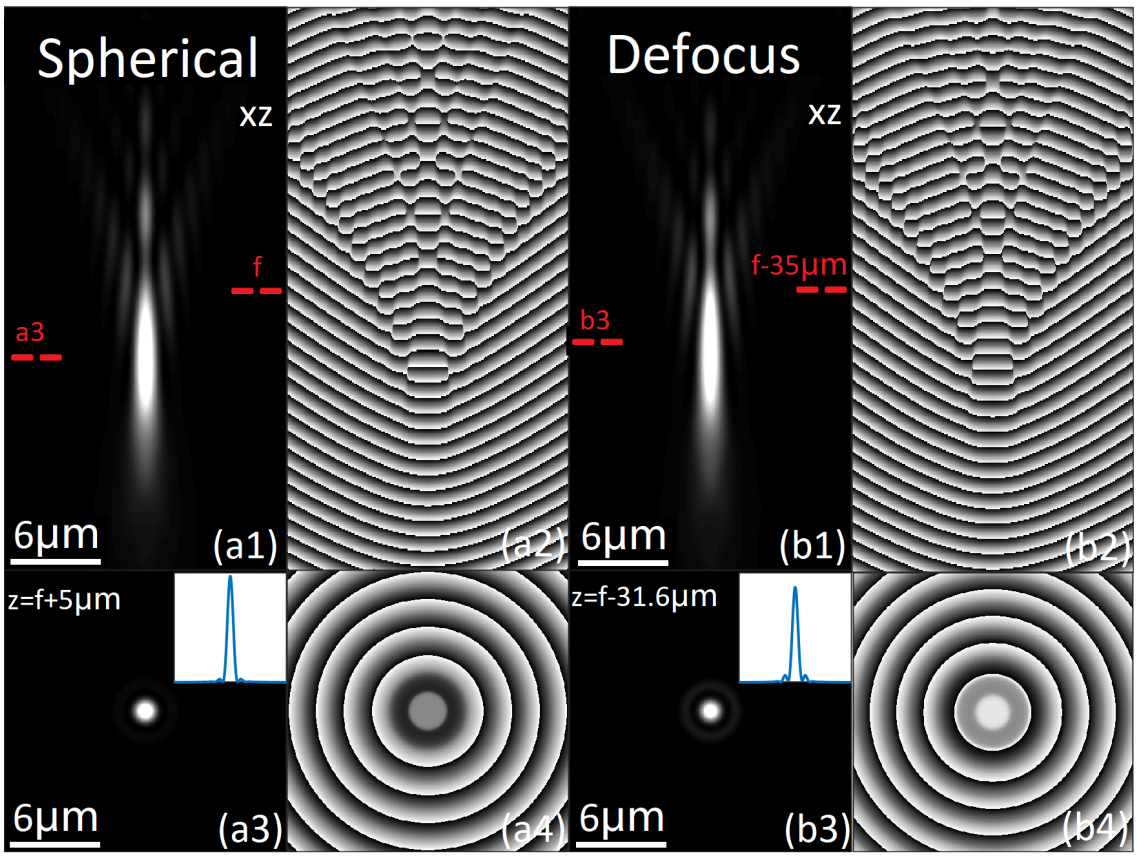

Figure 4.7: Intensity and phase distribution of focused spot for a Rayleigh Sommerfeld lens defined in Equation 4.17 for the case of Gaussian laser and with spherical ($w_{11} = 2 \times 10^{-7}$) and defocus aberrations ($w_4 = 2 \times 10^{-6}$), calculated using Algorithm 1. (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$. Note that the focal plane is indicated by the f symbol in the image and the brightest spot occurs at a different distance, $z = f + 5\mu m$. The intensity and phase of the XY plane at this distance are shown in (a3) and (a4). A similar set of results are shown for defocus aberration in (b1)-(b4). Note that for the case of the XZ plane shown in (b1) and (b2) the bright spot occurs at a large distance from the focus but the pattern of intensity and phase is remarkably similar to that for spherical aberration, albeit at a different distance. The same results for the ideal lens calculated using Algorithm 2 are provided in Appendix C,Fig.3

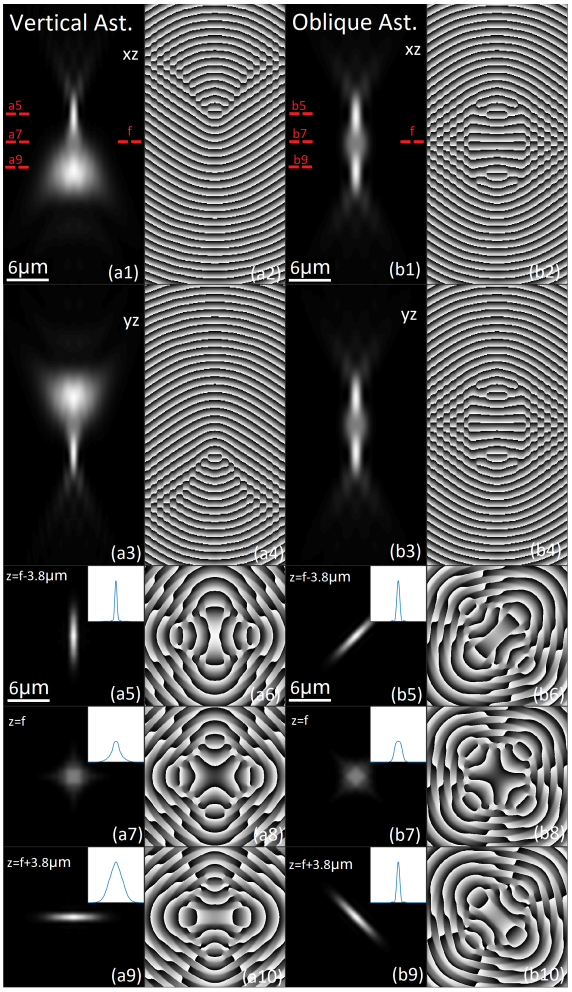

Figure 4.8: Intensity and phase distribution of focused spot for the RS lens defined in Equation 4.17 for a Gaussian laser and primary astigmatism, calculated using Algorithm 1. The case of vertical astigmatism is shown in figures (a1-a10) for $w_6 = 4 \times 10^{-7}$; (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$; the brightest spot appears at distance $f - 3.8\mu$m and the intensity and phase distribution at this distance in the XY plane is shown in (a5) and (a6); here, the light has converged along the X-axis but is distributed over a wider range along the Y-axis. (a3) and (a4) show the intensity and phase distributions of the YZ slice for $y = 0$; now the brightest spot appears at distance $f + 3.8\mu$m and the intensity and phase distribution at this distance in the XY plane is shown in (a9) and (a10); here, the light has converged along the Y-axis but is distributed over a wider range along the X-axis and the phase distribution is similar the pattern in (a6) rotated by $\pi/2$. The intensity and phase distributions at the focal plane are shown in (a7) and (a8). The same results are shown for the case of oblique astigmatism ($w_5 = 4 \times 10^{-7}$) in (b1-b10). Vertical and oblique astigmatism are related by a $\pi/4$ rotation of the 3D light volume; the intensity and phase distributions in the XZ (b1 and b2) and YZ (b3 and b4) are both identical and are symmetrical and the two focal points of light can easily be seen. The XY distributions shown for the three planes in (b5-b10) are equivalent to those shown for the vertical astigmatism case with $\pi/4$ rotation. Equivalent results for the ideal lens calculated using Algorithm 2 are provided in Appendix C, Fig. 4

The method above for the case of an ideal lens system, in which aberrations are described as a phase delay in the Fourier plane, or pupil function of the optical system is well established [62]. However, the thin lens case requires a different approach. Here we account for aberrations as direct distortions of the phase delay associated with the transmission function of the thin lens, $t(x,y)$. Therefore, in the description of the Algorithm 1, which simulates a thin lens system, Equation 4.33 is replaced with the following definition of the transmission function:

$$t(x,y) = \exp\left[jk\phi(x,y)\right] P\left(\frac{x}{R},\frac{y}{R}\right) Z\left(\frac{x}{R},\frac{y}{R}\right) \tag{4.33}$$

Thus, the phase delay imparted by the lens ϕ is added to the phase delay associated with each Zernike polynomial, $w_n Z_n$. The first step in Algorithm 1 is replaced with the following:

1. The first step is to define the wavelength λ and the continuous function $u_0(x_0,y_0)$; the user selects a lens phase delay function $\phi(x_0,y_0)$ as well as an input laser mode $A(x_0,y_0)$. In this selection, the radius of aperture, R, and the focal length, f, must be defined. Finally, the user selects the weights w_n for the Zernike polynomials in the aberration function. Therefore, the input function is defined: $u_0(x_0,y_0) = A(x_0,y_0)P\left(\frac{x_0}{R},\frac{y_0}{R}\right)\exp\left[jk\phi(x,y)\right]Z\left(\frac{x_0}{R},\frac{y_0}{R}\right)$

We note that the Rayleigh-Sommerfeld diffraction forumla has previously been used to theoretically investigate the impulse response from a thin planar lens, [107] which is similar in concept to the numerical approach applied here. In that paper, the authors showed that a thin quadratic lens produced an impulse response given by the integral of the pupil function multiplied by a phase function containing the Seidel aberrations. For our RS lens, we can expect no such aberrations (at least for on-axis imaging), since the thickness function of the lens has been modelled on the kernel of the Rayleigh Sommereld integral.

We proceed to examine the effect of aberration on the RS thin lens for the low-order Zernike polynomials using Algorithm 1. Equivalent results are shown in Appendix C for the ideal lens using Algorithm 2. For the RS lens with Gaussian profile that was simulated in the previous section (for which $f = 3.3$mm, $R = 2.5$mm, $\sigma = R/\sqrt{2}$), Algorithm 1 was implemented once again using an identical set of 3D sampling conditions as described earlier. This time, however, the aberration term was included, i.e. $u_0(x_0,y_0) = t_{RS}(x_0,y_0)A_{TEM_{00}}(x_0,y_0)Z(x_0/R,y_0/R)$ and the weights w_n were varied in order to examine the effect of specific cases of aberration on the 3D focal volume of light. The first case that was examined is that of spherical aberration for which the weight $w_{11} = 2\times10^{-7}$ and $w_n = 0\ \forall n \neq 11$. The intensity and phase distributions of the resulting field in the XZ plane is shown in Fig. 4.7 (a1) and (a2) over a size of 20µm×40µm. The effect of diffraction caused by the aberration on both the intensity and phase distribution is evident when compared with Fig. 4.6 (a1) and (b1). The correct focal plane of the lens (i.e. $z = f$) is indicated in the figure using the symbol 'f'. In addition to the diffraction there is also a clear shift in the axial position of the focus. The intensity and phase in the XY plane over an area of 10µm×10µm is shown at

the brightest point along the z-axis Fig 4.7 (a3) and (a4). The intensity and phase of this focused spot are similar in form to the unaberrated case shown in Fig. 4.6 (c1) and (d1). It should be noted, however, that increasing the value of w_{11} will result in further distortion of the spot as well as higher levels of diffraction and increased displacement of the focal point relative to the correct focal plane of the lens.

Also shown in the same figure is the case of defocus aberration for which the weight $w_4 = 2 \times 10^{-6}$ and $w_n = 0 \ \forall n \neq 4$. The intensity and phase distributions of the resulting field in the XZ plane is shown in Fig. 4.7 (b1) and (b2) over the same size as for the previous case. The effect of diffraction caused by the defocus aberration on both the intensity and phase distribution is remarkably similar in form to the previous results for the case of spherical aberration. However, there is a significant difference in the position of the focused spot along the Z-axis for both cases. Whereas the focused spot with spherical aberration occurs a few micrometers *after* the correct focal plane, $z = f$, it appears a few tens of micrometers *before* this plane for the case of defocus. The intensity and phase in the XY plane over an area of 10µm×10µm is shown at the brightest point along the z-axis in Fig. 4.7 (b3) and (b4). Once again this is similar to the equivalent result for the unaberrated but larger values of w_4 will result in further distortion and displacement.

The third aberration to be investigated (for the same RS lens with Gaussian beam) is primary astigmatism, which is commonly perceived as producing two focal points at different distances resulting from different lens curvatures along orthogonal axes. It should be noted that the polynomials Z_5 (vertical astigmatism) and Z_6 (oblique astigmatism) are similar, in that the latter is a $\pi/4$ rotation of the former. In Fig. 4.8(a1-a10) the results are shown for the case of vertical astigmatism with $w_5 = 4 \times 10^{-7}$ and $w_n = 0 \ \forall n \neq 5$. The intensity of the slice through the centre of the XZ plane and the YZ plane is shown in Fig. 4.8(a1) and (a3), respectively, with the corresponding phase images given in Fig. 4.8(a2) and (a4). The intensity and phase patterns of the YZ case appear to be identical to the patterns for the XZ case with an inversion of the Z-axis. It is clear that these two intensity images differ in terms of the distance at which the light is converging to a focus. For the XZ case, the light is seen to converge at a distance of $z = f - 3.8$µm, and the intensity and phase distribution in the XY plane at this distance are shown in Fig. 4.8(a5) and (a6) with the profile of the intensity along the X-axis shown in the small inset image; the light is obviously focused along the X-axis but has a wider support along the Y-axis. The opposite is true in the XY plane at a distance of $z = f + 3.8$µm, as demonstrated by the intensity and phase distribution shown in Fig. 4.8(a9) and (10). At the correct focal distance of $z = f$, half ways between the two 'focused spots,' the intensity pattern, shown in Fig. 4.8(a7) takes a diamond shape and the phase pattern shown in Fig. 4.8(a8) has the interesting property of each profile through the centre being equal to the inverse of the profile at a rotation of $\pi/2$. A similar set of results are shown in Fig. 4.8(b1-b10) for the case of oblique astigmatism with $w_6 = 4 \times 10^{-7}$ and $w_n = 0 \ \forall n \neq 6$. In this case, the intensity and phase distributions for the centre slices through the XZ and YZ planes are identical as shown in Fig. 4.8(b1-b4). The reason for this is that the two slices are passing though the elongated patterns at an angle of $\pi/4$. This becomes more apparent by observing the XY intensity and phase dis-

tributions at $z = f \pm 3.8\mu m$ shown in Fig. 4.8(b5), (b6), (b9), and (b10), which are clearly $\pi/4$ rotations of the equivalent images for the vertical astigmatism case. The intensity and phase distribution of the light at the correct focal plane shown in Fig. 4.8(b7) and (b8), is also $\pi/4$ rotation of the equivalent images for the vertical astigmatism case.

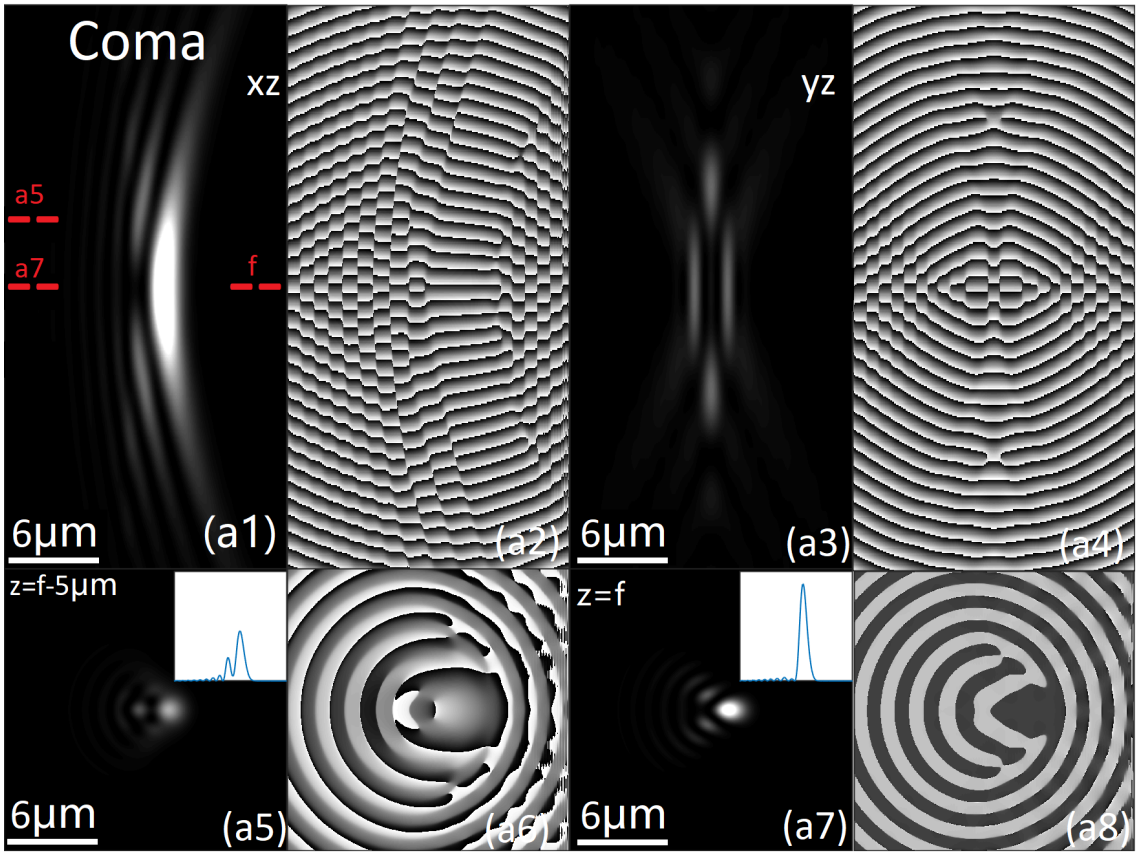

Figure 4.9: Intensity and phase distribution of focused spot for the RS lens defined in Equation 4.17 for a Gaussian laser and vertical coma ($w_7 = 3e^{-7}$), calculated using Algorithm 1. (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$, and the focused light clearly has a curved shape in this plane. The XY intensity and phase distributions are shown for two different distances: at $z = f - 5\mu m$ in (a5) and (a6) and at $z = f$ in (a7) and (a8) from which it can be seen the focused spot appears elongated along the X-axis and appears to move slightly as a function of distance z. Also shown is the intensity and phase distributions of the YZ slice for $x = 0$ in (a3) and (a4), which cuts through the curved spot. Equivalent results for the ideal lens calculated using Algorithm 2 are provided in Appendix C, Fig. 5

Comatic aberration was also applied to the RS lens using identical simulation parameters. This type of aberration is commonly encountered in imperfect lens designs and manifests as a comet like appearance particularly for off-axis points in the image. In Fig. 4.9 the intensity and phase patterns are shown for the case of coma with $w_8 = 3 \times 10^{-7}$ and $w_n = 0 \; \forall n \neq 8$. The intensity of the slice through the centre of the XZ plane and the YZ plane is shown in Fig. 4.9(a1) and (a3), respectively, with the corresponding phase images given in Fig. 4.9(a2) and (a4). The curvature of the focused spot along the Z-axis is evident both in terms of intensity and phase. The intensity and

phase distribution in the XY plane at the correct focal plane is shown in Fig. 4.9(a7) and (a8) with the profile of the intensity along the X-axis shown in the small inset image; the light is clearly displaced along the X-axis with the characteristic elongated distortion of the spot that is associated with coma. The intensity and phase images in the XY plane are also shown at a distance of $z = f - 5\mu m$ in Fig. 4.9(a5) and (a6). Here, it can be seen that the spot is further displaced from the centre of the X-axis due to the curvature of the focused light over the Z-axis. The results provided here relate to the 'Coma X,' see Table 4.1. An identical set of results can be obtained for 'Coma Y', for which $w_8 \neq 0$ except that the curvature of the spot takes place in the YZ plane and, therefore, the images for the XZ and YZ images are transposed.

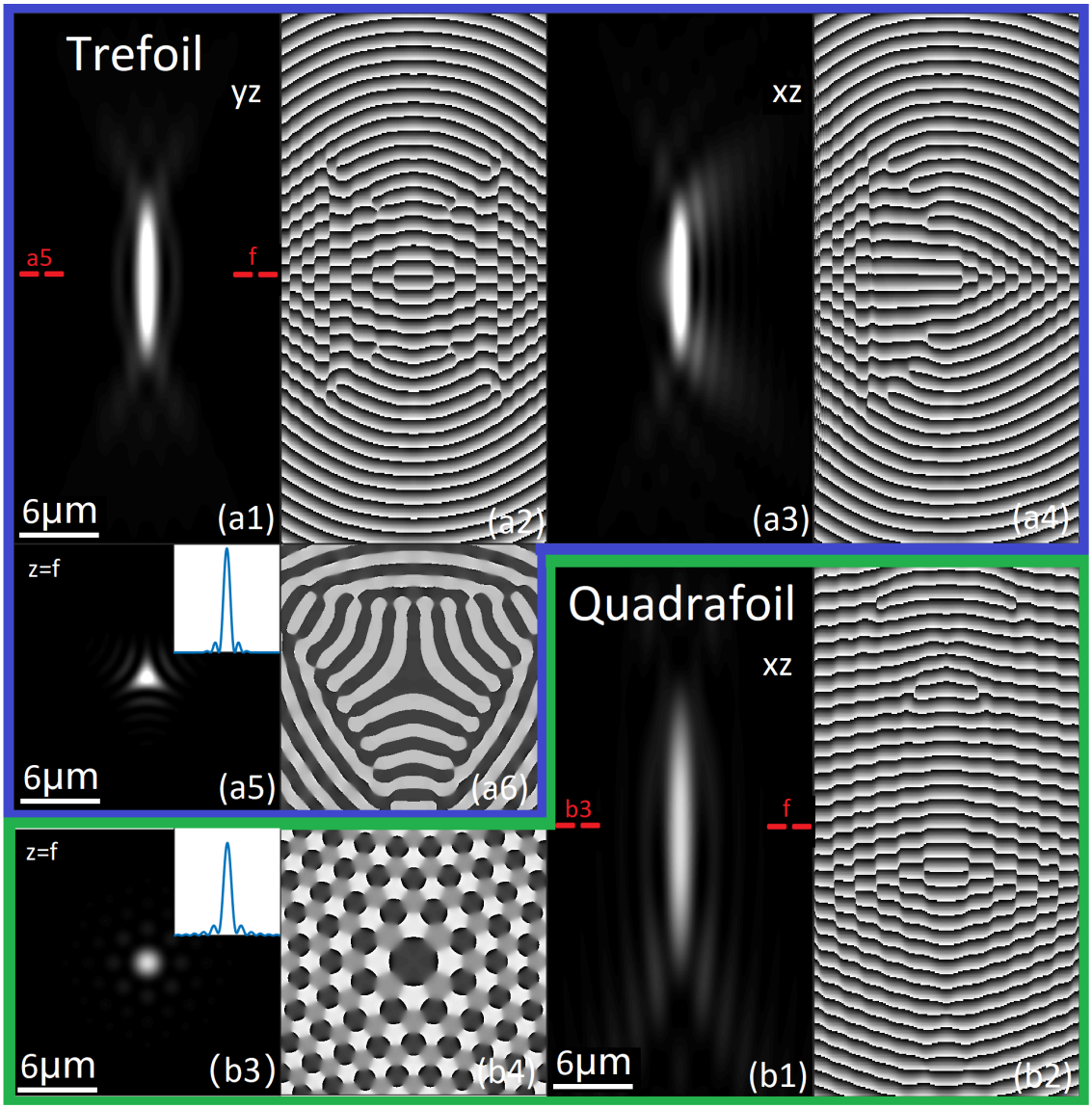

Figure 4.10: Intensity and phase distribution of focused spot for the case of a Gaussian beam incident on the RS lens with vertical trefoil aberration ($w_9 = 3 \times 10^{-7}$) shown in (a1-a6) and vertical quadrafoil aberration ($w_{14} = 1 \times 10^{-6}$) shown in (b1-b4), calculated using Algorithm 1; for the trefoil case (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$, and the focused light is symmetrical about the Z-axis in this plane; the intensity and phase distributions of the YZ slice for $x = 0$ is shown in (a3) and (a4), which is asymmetrical about the Z-axis. The intensity and phase distribution in the XY plane at the focal plane $z = f$ are shown in (a5) and (a6) in which the characteristic triangular shape of the focused light is seen. For the quadrafoil case (b1) and (22) show the intensity and phase distributions of the XZ slice for $y = 0$, and once again the focused light is symmetrical about the Z-axis but in this case there are two side lobes present around the focused spot; the intensity and phase distribution in the XY plane at the focal plane $z = f$ are shown in (b3) and (b4) in which the focused spot is surrounded by eight smaller spots. It is the two spots immediately to the left and right that correspond to the two side lobes in the XZ-slice. Equivalent results for the ideal lens calculated using Algorithm 2 are provided in AppendixC,Fig. 6

The final aberrations that were investigated with the RS lens are vertical trefoil

($w_9 = 3 \times 10^{-7}$) and vertical quadrafoil ($w_{14} = 1 \times 10^{-6}$) with resulting intensity and phase patterns shown in Fig. 4.10. Higher order aberrations such as trefoil and quadrafoil cannot easily be interpreted in terms of geometrical optics. For example trefoil can be thought of as a three axis astigmatism but this does result in three distinct focal points; however, triangular focal spots can be seen in the images of lens systems with this distortion. For the case of trefoil, the intensity and phase distributions in the XZ plane are shown in Fig. 4.10 (a1) and (a2). These patterns are symmetrical in both the X-axis and the Z-axis. However the intensity and phase in the YZ slice, shown in Fig. 4.10 (a1) and (a2), are not symmetrical in the Y-axis. This results from the triangular shape of the distribution in the XY plane, which is shown in Fig. 4.10 (a5) and (a6) and the fact that the Zernike polynomial Z_9 is periodic over rotations of $2\pi/3$. We note that the results presented here are for vertical trefoil, see Table 4.1. Similar results are obtained for oblique trefoil, which is effectively a $\pi/6$ rotation of vertical trefoil. Therefore, for the case of ($w_9 = 3 \times 10^{-7}$) we would observe a rotation of the XY distribution by $\pi/2$ or equivalently an inversion of the triangular pattern along the Y-axis. Also, the XZ and YZ distributions would transpose and the latter would be inverted along the X-axis. The results for quadrafoil are shown in Fig. 4.10 (b1-b4). The associated Zernike polynomial, Z_{14} is periodic over rotations of $\pi/2$ while the intensity and phase distributions of the focused spot in the XY plane ($z = f$), shown in Fig. 4.10 (b3) and (b4) are periodic over rotations of $\pi/4$; in the intensity pattern eight small spots can be seen positioned in an annular pattern around the focused spot. In Fig. 4.10 (b1) and (b2) the XZ distribution of the intensity and phase is shown. The presence of two of the aforementioned spots can be seen as side lobes around the centre spot. We note that the same pattern is observed at angles of $n\pi/4$ around the XY plane, where n is an integer. Not shown here is the opposing pattern at angles of $n\pi/4+\pi/8$ which contains only the centre spot without side lobes. We note that the results presented here are for vertical quadrafoil, see Table 4.1. Similar results are obtained for oblique quadrafoil, which is effectively a $\pi/8$ rotation of vertical quadrafoil. For that case the XZ-projection would not contain side lobes. A similar set of results for these various aberrations are shown in the Appendix C for the case of the ideal lens with a Gaussian beam, simulated using Algorithm 1 and using an identical set of parameters as detailed in Sections 4.3.2 and 4.4. The similarities between the intensity and phase patterns that are produced for the various aberrations are strikingly similar to those for the RS lens calculated using Algorithm 1, and using a fundamentally different approach to account for the aberrations, i.e. the latter accounts for the distortions as a variation in the thickness of the transmission function of the thin lens, while the former accounts for them in the pupil function or Fourier domain. There is one noticeable difference in that the effects of diffraction are less pronounced for the ideal lens case. This may result from the hard edge in the RS lens function or perhaps from the shift variant property of that particular lens.

4.6 Discussion

In this chapter, we have developed two algorithms for the numerical calculation of the three-dimensional complex amplitude of coherent wavefields that are focused by a lens. The first algorithm deals with the case of the thin lens approximation, while the second deals with an ideal lens system. In both cases, the angular spectrum method is applied in a loop in order to generate a stack of complex images over a regular grid around the focal region of the lens. However, the manner in which the ASM is applied differs considerably for both cases; for the thin lens, the ASM is applied directly to the lens transmission function but the difference in spatial support between the input plane (i.e the lens aperture) and the output plane (the focused spot) requires a 'pre-aliasing' step, recently proposed by other authors [13,95] and further developed in this chapter, which although numerically intensive, needs only to be applied once before the ASM is applied iteratively; for the ideal lens case, a single discrete Fourier transform is used to transform between the circular pupil function (in the Fourier domain) of the lens system to the focal plane of the lens, and then the ASM is applied iteratively to generate the stack of complex images immediately before and after this plane. A detailed list of steps is provided for both algorithms in Section 4.3 that includes the selection of the key lens properties as well as the optimal sampling conditions over the three spatial dimensions.

The thin lens functions that were initially selected for analysis included the case of the spherical lens for which the transmission function is directly related to the thickness of the glass. This was compared with the three-dimensional focus of an equivalent lens defined with a simple quadratic phase and the results were found to be the same, as expected. A third case was also investigated; here we proposed a transmission function defined by the kernel of the Rayleigh Sommerfeld diffraction integral. This aspherical 'lens' was simulated to produce a focused spot that was much smaller than the equivalent results for the spherical lenses. Indeed, the results were shown to be equivalent to those from an ideal lens with similar numerical aperture using the second algorithm.

In Section 4.4 both algorithms were augmented to account for the laser spatial mode for the case that the input to the lens is a temporally coherent laser input. In particular, TEM_{00} (Gaussian) and TEM_{01} (Laguerre-Gaussian) were investigated. Furthermore, in Section 4.5 both algorithms were augmented to account for aberrations in the lens and for both cases the well known Zernike polynomials were used to described the various forms of lens aberration. However, as for the case of the ASM, the manner in which the Zernike polynomials are applied differs considerably. The ideal lens case, the well established method of applying a phase delay to the pupil function (in the Fourier domain) was used, where this delay is a weighted sum of the Zernike polynomials. For the thin lens case, we proposed the application of a phase delay directly to the transmission function of the lens, where once again, this delay is a weighted sum of the Zernike polynomials. In the context of the thin lens we can explain this approach as a change in the thickness profile of the lens by an amount proportional to the Zernike polynomials. The intensity and phase profiles of the focused light for both algorithms

are remarkably similar for all of the lens aberrations investigated, namely spherical, defocus, vertical/oblique astigmatism, comatic, trefoil, and quadrafoil aberration.

This chapter is a useful reference for those who wish to simulate focused light over three-dimensions with lenses that have numerical aperture < 0.6 and can therefore be modelled using the paraxial approximation. It is particularly useful for those wishing to simulated focused laser beams by lenses with aberrations. However, this chapter represents the first part in a two part series. The second part, provided in Chapter 5, relates to calculating and visualising the flux of focused light fields, which has been investigated by several authors as an extension of the concept of geometrical ray optics; calculating the flux involves calculating the derivative of the phase of the wavefield as it propagates. A necessary first step is to calculate the complex amplitude of the light over three-dimensions with a sufficiently fine-grained sampling intervals such that the flux rays can be tracked accurately. Once these flux lines can be traced by the method developed in Chapter 5, these can be used to replace the geometrical rays in Chapter 3. The resultant flux-ray mode will follow in Chapter 5.

Chapter 5

Nonlinear Ray Tracing in Focused Fields.
Part 2: Tracking the Phase Derivative.

The work in this chapter has been prepared for a submission to a journal with the following title: Qin Yu, Bryan Hennelly, "Nonlinear Ray Tracing in Focused Fields, Part2:Tracking the Phase Derivative", To be submitted to Applied Optics with the abstract reproduced below:

In this two part chapter, we develop a method to trace the lines of flux through a three-dimensional wavefield by following a direction that is governed by the derivative of the phase at each point, a process that is best described as flux tracing but which we interchangeably name as 'non-linear ray tracing'. In the first part we focused on the high-speed calculation of focused three-dimensional complex wavefields in the paraxial approximation for TEM_{00} and TEM_{01} laser modes and in the presence of various lens aberrations. The algorithms developed in Chapter 4 are first used to generate the three-dimensional grid of samples of the complex wavefield in the focal region. In this second part, we focus on tracing a flux through this three-dimensional volume. This is achieved by first calculating the derivative of the phase (normal to the direction of propagation) throughout the three-dimensional volume, which is then used to direct a ray as it 'propagates' in a straight line between two consecutive planes within the volume. The origin of the ray can be chosen arbitrarily at any point and the ray can be then be traced through the volume with appropriate interpolation. Results are demonstrated for focused wavefields in the presence of aberrations, corresponding to the cases highlighted in Chapter 4. Some of the most interesting results relate to focused Laguerre-Gaussian beams, for which the rays are found to spiral at different rates of curvature, and for the cases of higher order aberrations such as Trefoil and Quadrafoil.

5.1 Introduction

Chapters 4 and 5 form a two-part chapter series; in Chapter 4 two algorithms were defined that can be used to compute the samples of the complex wavefield on a three-dimensional grid over the focal region of a lens. The differences in these algorithms relate to the physical model used to describe the lens, where one algorithm employs the thin lens approximation and the other models the ideal lens using a Fourier transform. Both algorithms allow for different laser spatial modes to be focused and different lens aberrations to be included. Either algorithm can be used as a basis for tracing the lines of flux using the methods described in this chapter. Understanding how these algorithms could be employed in fast time, and the sampling requirements over three dimension governed by the rules of Nyquist-Shannon interpolation is an important consideration in Chapter 4 since tracing the lines of flux will necessarily require accurate interpolation of the wavefield at arbitrary points within the the 3D grid of samples.

In wave optics, the angular spectrum, [8,62] which is the basis for the angular spectrum method [62] that is employed at the heart of both of the two algorithms developed in Chapter 4, represents the decomposition of a wavefront into a superposition of plane waves with different wavevectors (i.e., directions of propagation). Each plane wave in the superposition (which is defined elegantly using the Fourier transform) [61] has a different direction and amplitude, and collectively they form a complete representation of the original wavefront. The angular spectrum is related to the Eikonal function, [8,62,65] which describes the phase of the wavefront as a function of position and time. The Eikonal function can be used to define a "local wavevector" at each point in space, which describes the direction and magnitude of the local wavefront propagation. This local wavevector can be defined in terms of the local gradient of the phase of the wavefront [62]; and it is specifically this quantity that we measure for the wavefield in order to trace the direction of the Eikonal function through space. The resultant lines effectively trace the flux, i.e. the movement of the power that is the intensity of the diffraction pattern as is moves fluidly through space and we use the terms 'flux-lines' and 'non-linear ray tracing' interchangeably throughout the text. The use of the term 'non-linear ray tracing' is no doubt controversial and immediately forces a comparison with classical geometrical optics. The concept of the classical 'linear-ray' in geometrical optics is also related to the direction of propagation of the local wavevector at each point on the wavefront; the non-linear ray and the linear ray are one and the same so long as the gradient of the phase is well-described by the first two terms of its Taylor series expansion; in such case, the second derivatives of the phase and higher must be approximately zero.

Before proceeding with reading this chapter we encourage the reader to review Section 2.5. Specifically the reader should review how the angular spectrum can be physically interpreted in terms of plane waves that span all of space that are uniquely defined by a wavevector. They should also review the discussion of the Eikonal and the local wavevector, which naturally precipitates a discussion of the ray and the local spatial frequency. This forms a basis for a computational method to trace the flux lines, described in detail in this Chapter in Section 5.2. Since the method is based on se-

quentially propagating the ray over small steps, small errors could quickly accumulate. An important part of the discussion is an investigation of the accuracy of the proposed methods in terms of tracing along the correct path governed by the Eikonal function. This is followed by an array of results for various cases of laser spatial modes and lens aberrations.

Before proceeding to develop the flux-tracing algorithm, it is important to remind the reader of the context of this work in the overall thesis. As mentioned above Chapter 4 and 5 exists as a two-part series to develop this flux tracing algorithm that can track the paths of the flux lines through the propagating wavefield, effectively following continuous paths of power as they move through the three-dimensional focusing wavefield. Once the flux-tracing algorithm is developed it offers the possibility of replacing the rays in Chapter 3 (which propagate from the lens directly to the same point to focus) with the non-linear rays, which will have much different angles of incidence on the sphere and can even have a spin for the case of Laguerre-Gaussian lasers. This new 'flux-ray' model for predicting the forces on a trapped particle will follow in Chapter 6.

5.2 Tracing the flux

Before proceeding with this section, we encourage the reader to return to Chapter 2 and to review the concept of flux given in Section 2.5 and the concept of the phase derivative and its relationship to the geometrical ray. Specifically the local direction cosines that were introduced in Chapter 2 are the direction cosines of the ray through the (x, y) plane at each point. This definition of a ray, is consistent with the definition of a ray that was given in Chapter 2 in terms of the Eikonal function, and is the basis for the nonlinear ray tracing method proposed in this chapter.

5.2.1 Method

The crux of the proposed method is that we iteratively trace the flux between subsequent z-planes (separated by a distance h_z that is selected by the user) making use of the local spatial frequencies defined in Equation 2.36 to guide the ray angle at each plane. This is illustrated in Fig. 5.1(a) for which case 16 rays, with initial points uniformly arranged in a circle around a focusing Gaussian laser, have been traced through the focal plane. The method is illustrated in more detail in Fig. 5.1(b), were we highlight a single ray close to the focal plane. It is important to note that field $u_z(x, y)$ will have been sampled over a 3D grid with uniform sampling intervals T_X, T_y, and T_z as described in Chapter 4; however, the 3D positions at which the ray is sampled will not, in general, be coincident with the points on this grid. Therefore, interpolation is required in order to determine the complex amplitude of $u_z(x, y)$ at the desired points. For a given point on the ray (x_0, y_0, z_0), the field must be interpolated at four points, two of which are slightly displaced in x ($(x_0 + \delta_x, y_0, z_0)$ and $(x_0 - \delta_x, y_0, z_0)$) and two that are displaced in y ($(x_0, y_0 + \delta_y, z_0)$ and $(x_0, y_0 + \delta_y, z_0)$); this facilitates calculation of the angle cosines α_l and β_l according to Equation 2.36 and Equation 2.37. The ray

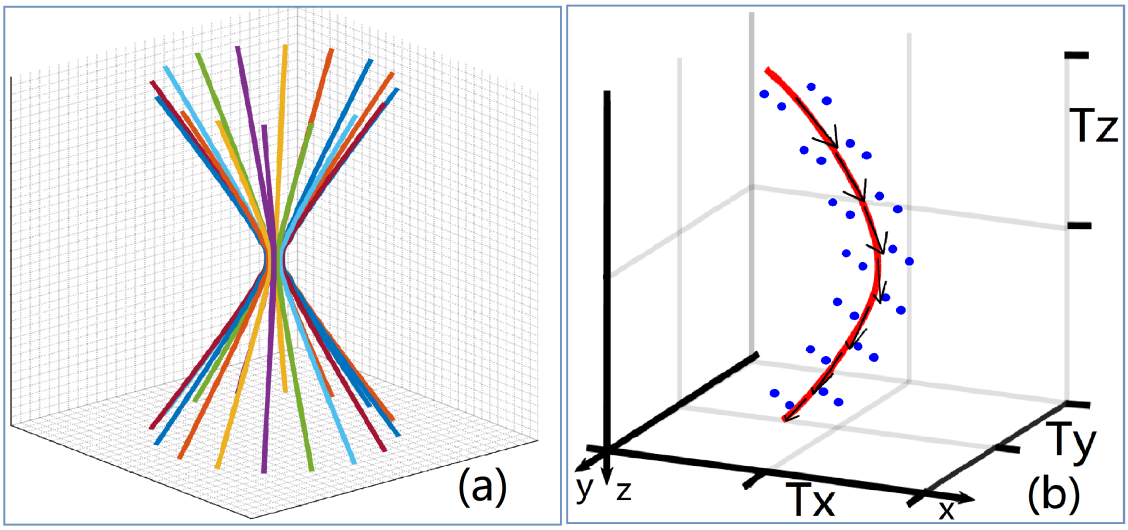

Figure 5.1: (a) Illustration of 16 rays uniformly arranged in a circle around a focusing Gaussian laser being traced through the focal plane being guided by the local spatial frequency of the field over subsequent z-planes; (b) Illustration of the method whereby the ray propagates in straight lines between the z-planes separated by a user-selected distance h_z. Note that the field must be interpolated at four points around the ray positions in order to calculate the phase derivatives, which are used to guide the ray to the next z-plane.

then traces a straight line from (x_0, y_0, z_0) to its new position $(x_0 + h_x, y_0 + h_y, z_0 + h_z)$, where h_x and h_y are defined in terms of the angle of the wave vector as follows:

$$\begin{aligned} h_x &= h_z \tan(\arccos(\alpha_l)) \\ h_y &= h_z \tan(\arccos(\beta_l)) \end{aligned} \tag{5.1}$$

where

$$\begin{aligned} \alpha_l &= \frac{\lambda}{2\pi} \frac{\phi(x_0 + \delta_x, y_0, z_0) - \phi(x_0 + \delta_x, y_0, z_0)}{2\delta_x} \\ \beta_l &= \frac{\lambda}{2\pi} \frac{\phi(x_0, y_0 - \delta_y, z_0) - \phi(x_0, y_0 + \delta_y, z_0)}{2\delta_y} \end{aligned} \tag{5.2}$$

Equation 5.1 can be rewritten more simply in terms of γ_l as follows:

$$\begin{aligned} h_x &= h_z \frac{\alpha_l}{\gamma_l} \\ h_y &= h_z \frac{\beta_l}{\gamma_l} \end{aligned} \tag{5.3}$$

where γ_l is calculated using Equation 5.2 and Equation 2.37. This method is then repeated at the new point resulting in another straight line being tracked to the next plane. The method, therefore, is to trace straight lines between subsequent z-planes in a piece-meal fashion. For a general disturbance, this method is accurate so long as the choice of h_z is small, which is discussed in more detail in the next section.

5.2.2 Accuracy

Following from the definition of the Eikonal, the ray must travel a path that is perpendicular to the phase front [8]. Therefore, the phase that it accumulates between two points along its path must be given by $\lambda/2\pi$ times the path length that was travelled between the two points. In general, this path will not be a straight line. However, if h_z is sufficiently small, then the path between the two points is in good agreement with linearity and this is the basis of the method proposed here for tracing the flux. Assuming linearity, we can define the phase accumulation between the two points as follows:

$$\phi_{z_0+h_x}(x_0+h_z, y_0+h_y) = \phi_{z_0}(x_0, y_0) + \frac{2\pi}{\lambda}\sqrt{{h_x}^2 + {h_y}^2 + {h_z}^2} \tag{5.4}$$

Substituting the values for h_x and h_z from Equation 5.3 produces the following relationship:

$$\phi_{z_0+h_x}(x_0+h_z, y_0+h_y) = \phi_{z_0}(x_0, y_0) + \frac{2\pi h_z}{\lambda\gamma_l} \tag{5.5}$$

This Equation is the model for the method proposed in this chapter. If the path of the Eikonal function is not linear between the two points, this will result in a phase error given by:

$$error = \phi_{z_0+h_x}(x_0+h_z, y_0+h_y) - \phi_{z_0}(x_0, y_0) - \frac{2\pi h_z}{\lambda\gamma_l} \tag{5.6}$$

Interestingly, the same two phase values can be related, without any approximation, by applying a three-dimensional Taylor series expansion to the continuous phase field, $\phi_z(x, y)$ as follows:

$$\begin{aligned}\phi_{z_0+h_z}(x_0+h_x, y_0+h_y) &= \phi_{z_0}(x_0, y_0) + \sum_{n=1}^{\infty}\frac{1}{n!} \\ &\times\left[h_x\frac{\delta\phi_z(x,y)}{\delta x}\Big|_{\substack{x=x_0\\y=y_0\\z=z_0}} + h_y\frac{\delta\phi_z(x,y)}{\delta y}\Big|_{\substack{x=x_0\\y=y_0\\z=z_0}} + h_z\frac{\delta\phi_z(x,y)}{\delta z}\Big|_{\substack{x=x_0\\y=y_0\\z=z_0}}\right]^n\end{aligned} \tag{5.7}$$

where the notation has the following meaning $[\delta\phi_z(x,y)/\delta x]^n = \delta^n\phi_z(x,y)/\delta x^n$. Therefore, the error function in Equation 5.6 can be rewritten in terms of the partial derivatives of the phase:

$$\begin{aligned}error &= -\frac{2\pi h_z}{\lambda\gamma_l} + \sum_{n=1}^{\infty}\frac{1}{n!}\left(\frac{h_z}{\gamma_l}\right)^n \\ &\times\left[\alpha_l\frac{\delta\phi_z(x,y)}{\delta x}\Big|_{\substack{x=x_0\\y=y_0\\z=z_0}} + \beta_l\frac{\delta\phi_z(x,y)}{\delta y}\Big|_{\substack{x=x_0\\y=y_0\\z=z_0}} + \gamma_l\frac{\delta\phi_z(x,y)}{\delta z}\Big|_{\substack{x=x_0\\y=y_0\\z=z_0}}\right]^n\end{aligned} \tag{5.8}$$

using the substiutions from Equation 2.37 this further reduces as follows:

$$error = \sum_{n=2}^{\infty} \frac{1}{n!} \left(\frac{h_z}{\gamma_l} \right)^n$$
$$[H] \times \left[\alpha_l \frac{\delta \phi_z(x,y)}{\delta x}_{\substack{x=x_0 \\ y=y_0 \\ z=z_0}} + \beta_l \frac{\delta \phi_z(x,y)}{\delta y}_{\substack{x=x_0 \\ y=y_0 \\ z=z_0}} + \gamma_l \frac{\delta \phi_z(x,y)}{\delta z}_{\substack{x=x_0 \\ y=y_0 \\ z=z_0}} \right]^n \tag{5.9}$$

This error function defines the phase error between two subsequent points that have been computed. If the second point were truly along the flux line defined by the Eikonal function, this error function would reduce to zero. If this error function is not close to zero, then the second point on the ray has not been selected correctly using the method defined in Equation 5.3. In this case, it could be expected that further errors will follow with subsequent ray traces and that the error would propagate resulting in an erroneous flux line.

Some interesting observations can be made from Equation 5.9. Firstly, for the case of a plane wave, such as that defined in Equation 2.25, the error is zero regardless of the value of h_z. This is due to the fact that all partial derivatives are zero for $n > 1$. This makes perfect sense since the plane wave is unidirectional and, therefore, any step size h_z will produce an error free value of h_x and h_y. A second observation is the dependence of the error on the term h_z^n where the smallest value of this term is $n = 2$, which will be applied to the second derivatives only. The error can be expected to be small if (i) the partial derivatives are small for $n > 1$ and/or (ii) h_z is small. In the results presented in the remainder of this chapter, we make no assumptions about the value of the higher order partial derivatives; instead, we make sure to use very small values of $h_z \ll 1/\Delta f_z$ (that are significantly smaller than the Nyquist interval) in order to guarantee accurate flux lines.

5.2.3 Algorithm

If only a single ray were to be considered, the 3D grid of samples that described $u_z(x,y)$ could be computed in the region of interest according to Algorithm 1 or Algorithm 2 described in Chapter 4 using the Nyquist sampling intervals for T_x, T_y, and T_z. It would then be possible to identify a starting position for the ray and to trace the ray through the 3D grid by performing 3D interpolation sequentially as the ray progresses in straight lines over steps of length h_z along the z axis. Here, we define an algorithm that is more straightforward, particularly for the case when multiple rays are traced in parallel. The algorithm defined below is based on simply selecting $T_z = h_z$ and applying 2D interpolation in the finely sampled z-planes. Although this method will often require more computation than the 3D interpolation method, it is simple to implement and produces reliable results. The steps of the algorithm are illustrated in Fig. 5.2 and are defined as follows:

1. The first step is to select the initial ray origins. A set of N points defined as (x_n, y_n, z_0) for $n : 1 \rightarrow N$ where each (x_n, y_n) is a unique point in the initial z-plane at arbitrary distance z_0. Values for δx and δy must also be selected relating to the discussion in Section 5.2.1. The value of these shifts should be small, but not so

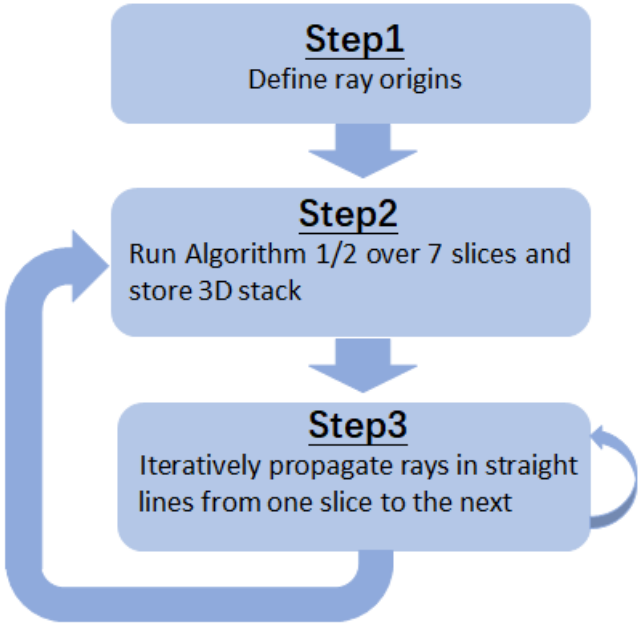

Figure 5.2: Illustration of the key steps in the non-linear ray tracing algorithm. The second step relates to running one of the two algorithms that were defined in Chapter 4 to generate a 3D grid of samples that is stored in memory. The third step is iteratively applied to propagate the rays in straight lines from one z-slice to the next using the derivative of the phase to define the ray angle. To reduce the memory load only a small number of z-slices are calculated and stored any one time.

small as to result in machine error; we have found the value of $1e^{-13}$ to provide excellent results.

2. The second step is to apply Algorithm 1 or 2. This includes selecting the properties of $u_z(x, y)$ including the aperture radius, R, and the Zernike phase term Z and laser mode A. This also includes the selection of the sampling intervals T_x, T_y, and T_z. Note T_z is chosen to be very small to ensure accuracy as described in the previous section. The range of distances defined by n_z and M_z in Algorithm 1/2 will be defined based on the initial ray positions. The result of Step 2 is that a 3D grid of samples of $u_z(x, y)$ is calculated over a region around the focal point of the lens. However, only seven slices are calculated and stored in order to reduce the memory load at any one time; the rays are propagated through these seven slices as described in the next step.

3. 2D Interpolation is applied in the region of each point (x_n, y_n, z_0) as described in Section 5.2.1. For each point, the local cosine angles, α_{ln} and β_{ln} are calculated according to Equation 5.2. Updated ray positions are given by $(x_n, y_n, z_0) \rightarrow r(x_n + h_x, y_n + h_y, z_0 + h_z)$, where a value for h_z is defined by the user; in all of the results presented in this chapter a value of $h_z = 10$ nm was selected. We proceed to the next z-slice and this step is repeated iteratively seven times until the ray propagates to the last z-slice. We then return to Step 2 and another seven z-slices are calculated. Steps 2 and 3 are applied repeatedly until the rays have completed their paths over the desired propagation distance.

5.3 Results for different laser spatial modes

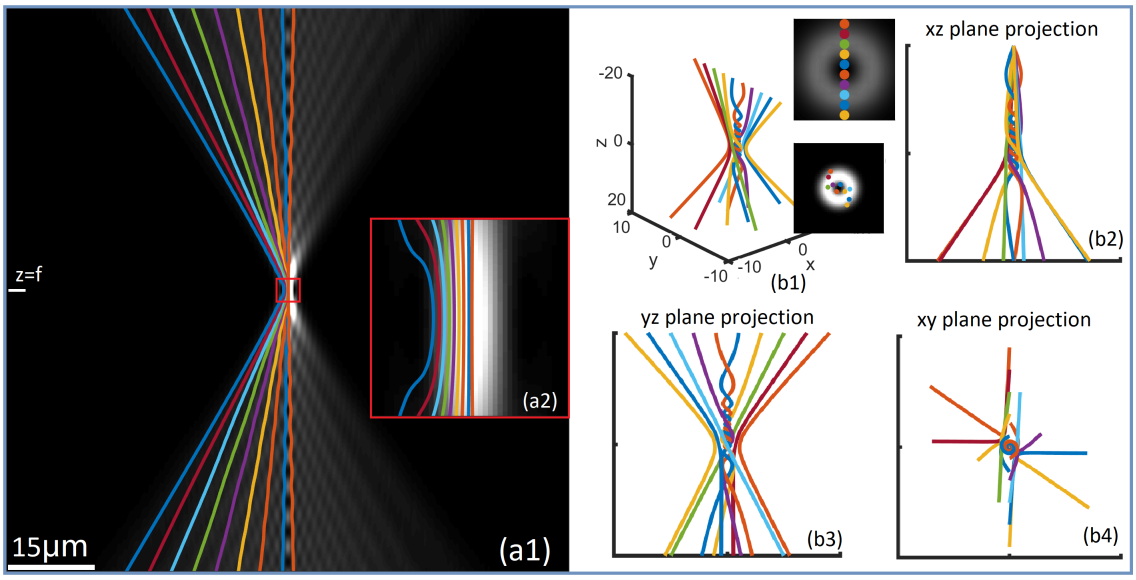

Figure 5.3: Nonlinear ray tracing for the same TEM_{00} TEM_{01} laser modes focused using the same Rayleigh-Sommerfeld lens with numerical aperture of 0.6 investigated in Fig. 4.6: (a1) TEM_{00} case with log scale of intensity in the background and focal region magnified in (a2); (b1) 3D image of rays spiralling for the TEM_{01} case. The two inset images show the initial ray positions at -20µm and at the focal plane; (a2)-(a4) show three different projects of the ray paths.

Geometrical ray tracing is typically applied without consideration of laser modes. In some cases, such as for example the calculation of optical trapping forces in the ray-optics regime, the rays are given weights according to the amplitude of the electromagnetic wavefield, i.e. relating to their position across the laser mode profile. In that case, the rays in the centre of a Laguerre-Gaussian trapping laser will have zero weight in the trapping force calculation, while those at the edge are stronger [21]. However, this approach does not in any way relate the laser mode to the direction of the rays, which must by definition travel along linear paths. In this section, we show the results of nonlinear ray tracing applied to focusing lasers with Gaussian and Laguerre-Gaussian spatial modes. These rays follow the lines of flux of the wavefront as it propagates. In Fig. 5.3 the results are shown for the same TEM_{00} TEM_{01} laser modes focused using the same Rayleigh-Sommerfeld lens investigated in Fig.4.6. All parameters used in Algorithm 1 are identical except for the reduced number of slices calculated in a given stack.

In Fig. 5.3 (a1) 9 rays are shown for the TEM_{00} case and all but one are on the left side of the image. The intensity is shown in the background on a log scale such that the effects of diffraction due to the apodization of the lens aperture are more clearly visible. Far from the focus, most of the rays appear to be linear and are similar to the classical geometrical rays. However, close to the focus, which is highlighted in the inset image (a2), the rays bend rapidly before straightening to vertical paths at the focus and then bending outwards again at an opposite angle to which they arrived at the focus. Along

the axis of propagation, some of the strongest effects of diffraction can be observed resulting in intermittent dark and bright regions and the rays close to this optical axis are seen to propagate around these dark regions. This feature is expected from the idea of flux.

In Fig. 5.3 (b1) 10 rays are shown for the TEM_{01} case. The positions of these rays are selected to be uniformly separated along the y-axis at a distance of -20μm before the focus as illustrated in the top inset image. The second inset image shows the positions of these rays at the focal plane. In all cases, the rays have moved away from the y-axis. The 3D paths that these rays have taken between the two planes can be seen in (b1) and the various 2D projections of these 3D paths are shown in Fig. 5.3 (b2)-(b4). In all cases, the rays are seen to spiral, with the rate of twisting depending on the distance to the optical axis and the distance to the focal plane. Most noticeably, the rays close to the centre twist rapidly as they approach the focal plane.

A deeper investigation of the ray paths for the Laguerre-Gaussian modes is presented in Fig. 5.4, once again for the same Rayleigh Sommerfeld lens and with numerical propagation implemented using Algorithm 1. All of the simulation parameters are the same as for the previous case except for the choice of l and p in the definition of the LG-modes L_p^l in Chapter 4. The p value varies across the columns from 0 to 2, while the l value varies over the rows from 1 to 3. For the first row, corresponding to $p = 0$, the intensity at the focal plane is shown for the three l values, as well as the intensity at a distance of -20μm before the focus; it can be seen that the size of the donut increases as a function of l. In each case of the three cases, eight ray origins were selected close to the outer rim of the donut in $z = -20\mu$m plane; these are shown as red dots in the intensity images. Another eight ray origins were selected close to the inner rim, which are shown as green dots. These rays were traced over a distance of 40μm, and are shown as red and black lines in the 3D plot. Although the donut converges as the beam focuses, it is interesting to note that the relative position of the rays within the donut remains unchanged in the intensity images at $z = -20\mu$m and at $z = -0\mu$m for all three cases. It is also interesting to note that the inner rays spiral rapidly, while the outer rays twist only slightly. The rate of angular velocity of the inner rays as a function of z is particularly striking with the rays appearing to undertake a full 2π revolution over a distance of z that is shorter than the Nyquist distance defined in Chapter 4; this is particularly noticeable for the $l = 3$ case. The rate of angular velocity of the rays is clearly dependent on the ray position relative to the centre of the donut together with the factor l; focusing on the inner rays, the rate of spin appears to approximately linearly increase as a function of l. This is consistent with the concept of angular momentum associated with Laguerre-Gaussian beams, which is directly proportional to the value of l; per photon, this momentum is given by hl where h is the reduced Planck constant [108].

A similar trend is observed for the second row, corresponding to $p = 1$. For all three cases, there exist two distinct donuts in the focal plane, although the size and shape of these donuts varies significantly as a function of l. The diffraction pattern at $z = -20\mu$m is more complex than for the previous case of $p = 0$. It it interesting to observe the movement of the rays as they approach the focal plane; most of the rays

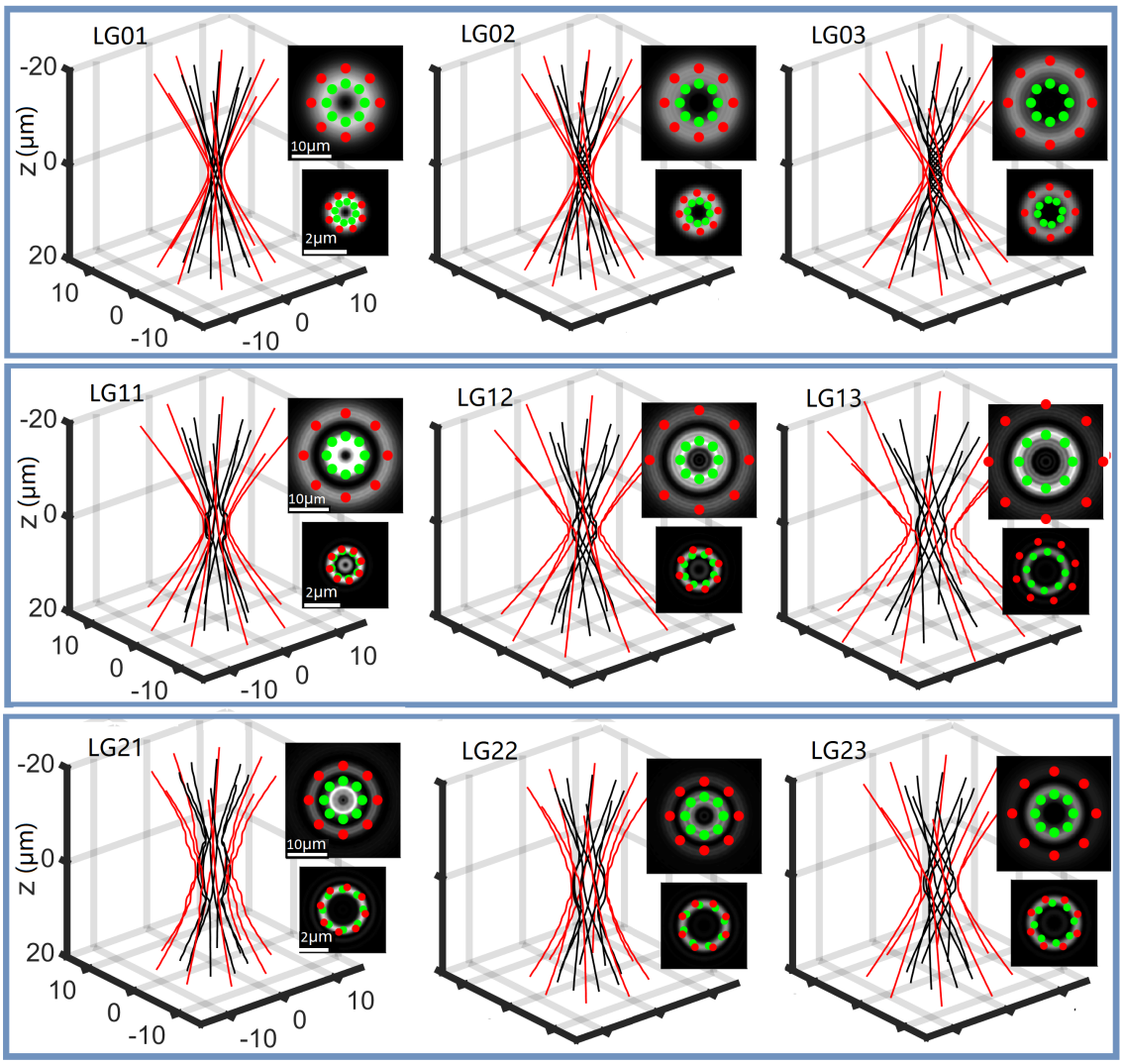

Figure 5.4: Nonlinear ray tracing for the different Laguerre-Gaussian modes focused using the same Rayleigh-Sommerfeld lens with numerical aperture of 0.6 investigated in Fig. 4.6: The values of l and p vary across rows and columns, respectively. Also shown in the figure are the intensity patterns at the focal plane and 20 µm above, which highlight the ray positions in these planes. It is interesting to note that the spin rate of the rays increases linearly as a function of l and also as a function of the position of the ray relative to the centre of the donut. An equivalent figure is shown in the Appendix C, Fig. 7 for the ideal lens (NA 0.6), computed using Algorithm 2.

diverge towards the outer donut, with many requiring a somewhat sudden and significant change in direction as they approach the focal plane in order to do so. Although, it appears that the spin rate of the rays is less than for the previous examples for $p = 0$, it must be the case that the aggregate of the spin over all of the rays (weighted according to the wavefield intensity) must be the same given that the angular momentum is related only to l and not to p. Similar trends are observed for the $p = 2$ case shown in the bottom row in the figure, although in this case the ray lines are less smooth, perhaps owing to higher levels of diffraction associated with more complex Laguerre polynomials. Similar results are shown for the ideal lens case using Algorithm 2 in the Appendix.

5.4 Results for different aberrations.

We begin this section with a brief discussion on the interpretation of lens aberration using the principles of geometrical optics. Low order aberrations are most easily understood in terms of geometrical optics and so we begin the discussion with an overview of those aberrations that can be interpreted in terms of geometrical rays that are distorted, by which we mean their path is deviated from the ideal path; [65] the ideal path refers to the path we might expect from an ideal lens system that focuses parallel rays to a single point in three dimensions at the focal plane of the lens. We note that interpretation in terms of geometrical optics is broadly limited to only some of the most commonly encountered low order aberrations; many more forms of higher order optical aberrations exist [8] that cannot easily be interpreted in terms of geometrical optics in the same way that is possible with lower order aberrations. Following this brief discussion, we will attempt to interpret both low and higher order aberrations in terms of non-linear ray tracing.

- Spherical aberration results in the rays that are converging from the lens coming to focus at different planes along the optical axis as illustrated in Fig. 5.5(a); for this case the 'focal length' for a given ray is related to its position of origin on the lens aperture. Rays that originate farther from the centre of the lens will come to focus at a plane some distance farther than the correct focal plane. Spherical aberration can significantly affect resolution and clarity across an image. Spherical aberration is an inherent property of perfectly spherical lenses, such as those described by the thin lens phase transmission functions t_1 and t_2 in Chapter 4.

- Defocus is a similar aberration in the sense that it also relates to a deviation from the correct focal plane. In this case, however, all of the rays come are understood to focus at a single plane, but not the correct plane. Defocus is illustrated in Fig. 5.5(b).

- Comatic aberration or 'coma' results from imperfect lens designs that cause off-axis point sources to appear distorted in the image, appearing to trail off like a comet tail. More specifically, coma is defined as a variation in the magnification over the lens pupil. Obliquely parallel rays will focus at different points in the

lens focal plane depending on the position of their origin in the lens aperture. The effect of comatic aberration is illustrated in Fig. 5.5(c).

- Astigmatism relates to an irregular lens curvature that is elliptical rather than spherical, which results in light coming to focus at two or more different distinct points, with obvious deleterious affects on the imaging properties of the lens. There are various classes of astigmatism depending on the shape of the irregularity. Vertical astigmatism, illustrated in Fig. 5.5(d), relates to different focal length in the x and y direction resulting in rays located along these two axes coming to focus at two different distances. Oblique astigmatism relates to a lens curvature whereby the slow and fast lens meridians are not at 90° or 180° as for the case of vertical astigmatism.

- Higher order aberrations such as trefoil and quadrafoil cannot easily be interpreted in terms of ray optics. For example trefoil can be thought of as a three axis astigmatism but this does result in three distinct focal points; however, triangular focal spots can be seen in the images of lens with this distortion. These types of higher order descriptions are better understood in terms of the phase delays that are imparted on the lens pupils using the Zernike polynomials. Specifically, when wavefront errors exist, we can imagine the exit pupil is illuminated by a perfect spherical wave, but that a phase-shifting plate exists in the aperture, thus deforming the wave front that leaves the pupil. In the discussion that follows, we will attempt to interpet these aberration in terms of the non-linear ray tracing proposed in this chapter.

The results of non-linear ray tracing applied to spherical aberration are presented in Fig. 5.6 (a1)-(a3) for the case of the Raleigh Sommerfold lens with numerical aperture of 0.6 with TEM_{00} illumination computed using Algorithm 1 as described in Section 4.5. Note that all of the parameters outlined in Chapter 4 are identical here except for the value of T_z and the use of three different weighting values, w, shown in the figures. Note that the third case (a3) corresponds to the result given in Fig 4.7. In each case, the log of the intensity is shown in the background on which 12 rays are overlaid with origins that are distributed along the x-axis in the z-plane that is 20μm before the focal plane. The ray origins are selected at positions that best highlight the effects of the aberration on a case-by-case basis. The focal plane, which we define as the plane with the maximum intensity, is seen to shift as a function of w, with a significant shift in offset occurring for the case $w = 5 \times 10^{-8}$. For the case with the lowest weight shown in (a1) only a slight offset in focus is observed and the results are similar to those for the aberration-free case presented in Fig. 5.4 with the innermost rays following paths that avoid regions of low intensity. The outermost rays appear to come to focus slightly before the focal plane. This effect is a little more pronounced for the second case in (a2) for which the weighing has been doubled $w = 1 \times 10^{-7}$. The third case, (a3), for which the weight is doubled again to $w = 2 \times 10^{-7}$ best exemplifies the effects of spherical aberrations. Here it can be seen the the outer rays focus before the inner rays, and the distance at which they focus depends on the position of the ray. This is consistent with

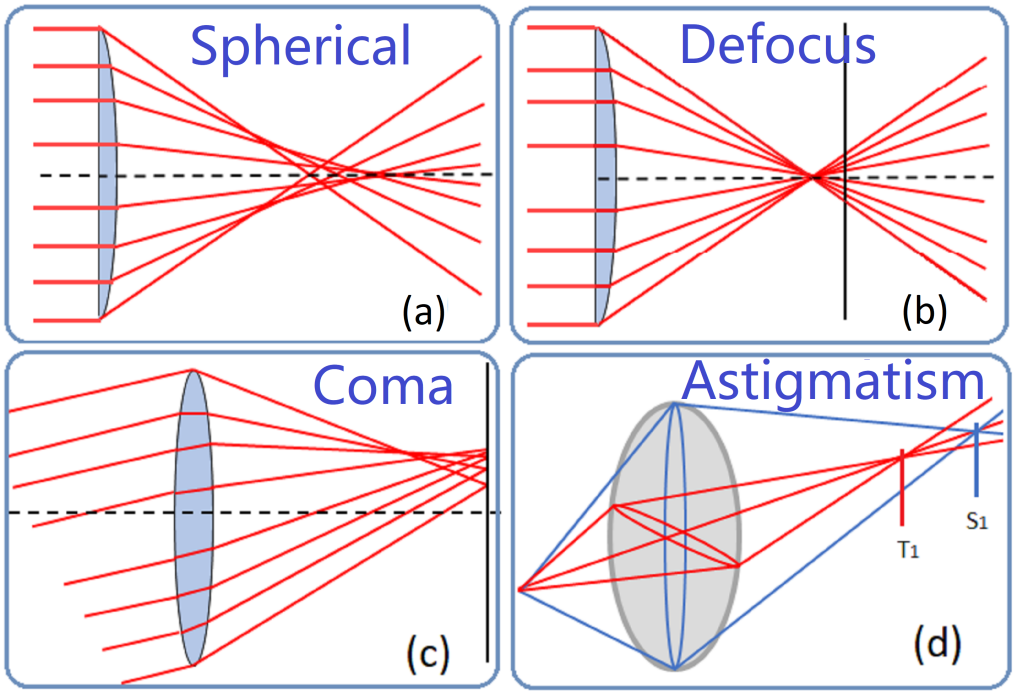

Figure 5.5: Four low order aberrations that can be interpreted in terms of geometrical optics: (a) Spherical aberration for which rays focus at different distances depending on their position in the lens aperture; (b) defocus aberration which causes the rays to focus at an incorrect focal plane; (c) comatic aberrations for which off-axis rays focus at different positions in the focal plane depending on their on the position in the lens aperture; (d) astigmatism aberration, which results in two or more focal points, most commonly for rays in the vertical and horizontal axes.

the conventional interpretation of spherical aberration in terms of geometrical rays that was discussed earlier.

A similar set of results are shown in Fig. 5.6 (b1)-(b3) for the case of defocus aberration. For the lower weight of $w = 5 \times 10^{-7}$, the result looks similar to the case of low spherical aberration in (a1) and the effects of the aberration appear to be minimal. However, the shift in the focal plane is significantly greater, which is consistent with the interpretation of defocus aberration in terms of geometrical optics described earlier. For the second case in (b2) with the value of w doubled, this shift in focus is also doubled. We also see in this case, that the outer most rays appear to converge at a plane slightly before the focal plane. This result is comparable to the second case of spherical aberration shown in (a2). The result of a further doubling of w is shown in (b3). As expected, the shift in focus is doubled again with respect to the previous case. However, in this case we can see that the outer rays follow a trend that is similar to the case of high spherical aberration in (a3). We can conclude that high levels of defocus and spherical aberration exhibit similar effects on the flux lines, with defocus resulting in greater shifting of the focal plane. It is also notable that the shift in focus is linearly dependent on w for the case of defocus aberration but this relationship appears to be non-linear for the case of spherical aberration, although the amount of defocus is much less for the latter. Lower levels of defocus aberration appear to shift the focal

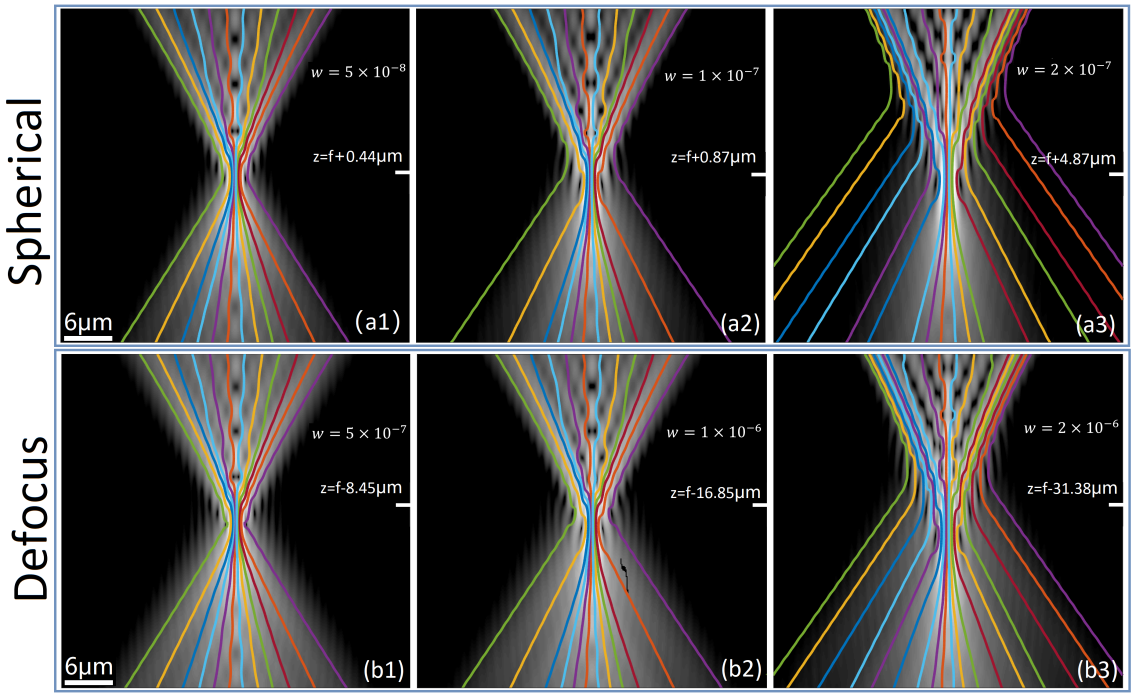

Figure 5.6: Non-linear ray tracing applied to different cases of spherical (a1-a3) and defocus (b1-b3) aberration on the ideal Rayleigh-Sommerfeld lens with numerical aperture of 0.6 and TEM_{00} illumination computed using Algorithm 1. This simulation is based on the identical method and parameters described in Fig 4.7. Three different weights, w, of spherical and defocus aberration are applied, shown in the figures. Note that the centre figures on both rows (a3) and (b3) correspond to the specific results given in Fig 4.7. Equivalent results are given in Appendix C, Fig. 8 for the ideal lens computed using Algorithm 2. Note that both algorithms model aberrations in fundamentally different ways and yet both produce remarkably similar results.

plane without exhibiting the properties associated with spherical aberration. We note that equivalent results relating to spherical and defocus aberration are presented in the Appendix C, Fig. 8 for the ideal lens computed using Algorithm 2. Note that both algorithms model aberrations in fundamentally different ways and yet both produce remarkably similar results.

The results of non-linear ray tracing applied to vertical astigmatism aberration are presented in Fig. 5.7 (a1)-(a12), again for the case of the Raleigh Sommerfeld lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 1. All of the parameters outlined in Chapter 4 are identical here (including the value of w used in the investigation of vertical astigmatism) except for the value of T_z. In the left-most column a 3D plot is shown that traces various sets of 10 or 12 rays from $z = -20\mu m$ before the focal plane to $z = 20\mu m$ after, and also shows two inset (log of) intensity images that correspond to the intensity at the plane $z = -20\mu m$ before the focal plane, and at the focal plane. The larger inset image has dimensions of $20\mu m \times 20\mu m$ and the smaller inset image has dimensions of $10\mu m \times 10\mu m$. The different rows correspond to a different choice in the ray origin positions in the plane at $z = -20\mu m$. The first row of images (a1)-(a4) relate to a uniform distribution of 10 ray origins along the x-axis.

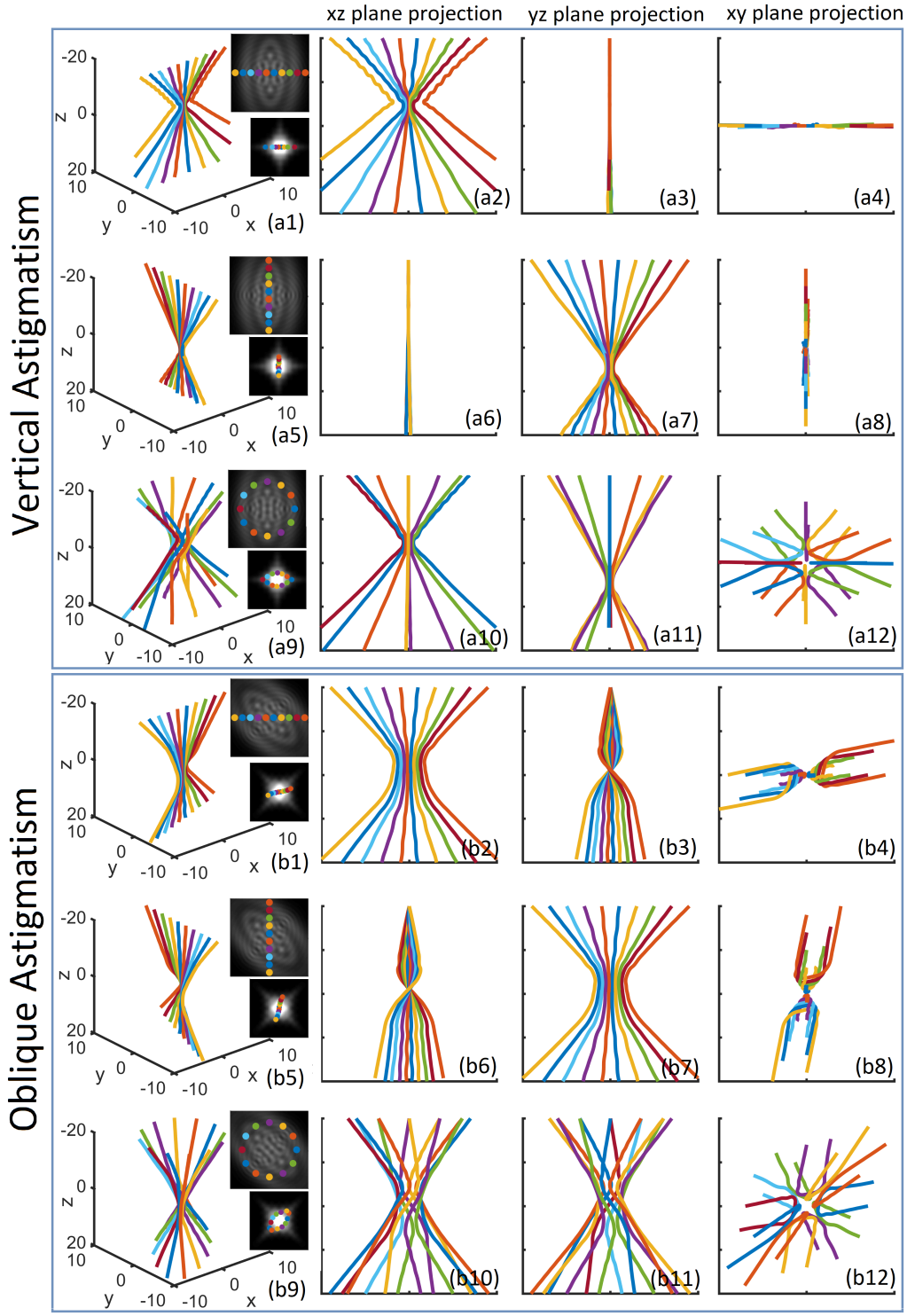

Figure 5.7: Non-linear ray tracing for vertical and oblique astigmatism for the Rayleigh-Sommerfeld lens; see text for details.

It is clear from the various projections, as well as from the two intensity images, that the rays do not deviate from the plane $y = 0$. It is also clear that the rays come to focus a number of micrometres before the expected focal plane. The second row of images (a5)-(a8) relate to a uniform distribution of 10 ray origins along the y-axis. In this case, the intensity images and the various projections indicate that the rays do not deviate from the plane $x = 0$ and come to focus a number of micrometres after the expected focal plane. These effects are entirely consistent with the conventional interpretation of vertical astigmatism in terms of geometrical rays that was discussed earlier. In the third row of images (a9)-(12) the effect of vertical astigmatism on non-axial ray origins is shown. In this case, 12 ray origins, are uniformly distributed around a circle in the plane at $z = -20\mu$m. It can be seen from the intensity images that the ray positions form an ellipse in the focal plane. From the various projections, it can be seen that those rays that are located on the x- and y-axes remain in the x- and y-planes, while rays that are not located on these axes appear to twist in three dimensions. Once again the two different focal planes can be observed in the xz and yz projections shown in (a10) and (a11).

A more detailed investigation of the effects of astigmatism, specifically relating to non-axial ray positions, can be provided by oblique astigmatism, which is shown Fig. 5.7 (b1)-(b12). It is important to emphasise that vertical and oblique astigmatism are related by a simple $\pi/4$ rotation of the wavefront aberration as described in the previous chapter. Therefore, sampling ray origins along the x- or y-axis for the case of oblique astigmatism is equivalent to sampling along a ± 45 degree line in the xy plane for the case of vertical astigmatism. These two cases are presented in figures (b1)-(b4) and (b5)-(b8), respectively. It is interesting to note from the two intensity images in (b1) that the rays that are sampled along the x-axis in the plane at $z = -20\mu$m are not found to be along the x-axis in the focal plane. Rather they are linearly distributed along some diagonal line in the xy plane. This explains why the rays never come to a tight focus in the xz plane in the projection shown in (b2). However, in (b3), it is clear that the rays do focus tightly in the yz-plane; first, the rays deviate away from the x-plane before turning back towards the focal point. These ray paths are further illuminated by the xy projection shown in (b4), which appears to show that the rays twist around by an amount that depends on the distance of the ray origin from the centre. The results of sampling the ray origins along the y-axis are shown in (b5)-(b8) and are found to be identical to the previous case but for opposite axes. The final row of images (b9)-(b12) relate to a circular sampling of 12 ray origins in the plane at $z = -20\mu$m. This result is identical to the corresponding result for vertical astigmatism, with a simple $\pi/4$ rotation of the xy plane. In this case, however, it is a little clearer to see the three-dimensional movement of the ray paths from the xz and yz projections in (b10) and (b11). We note that equivalent results relating to vertical and oblique astigmatism are presented in the Appendix C, Fig. 9 for the ideal lens computed using Algorithm 2. Note that both algorithms model aberrations in fundamentally different ways and yet once again, both produce remarkably similar results.

The results of non-linear ray tracing applied to coma aberration are presented in

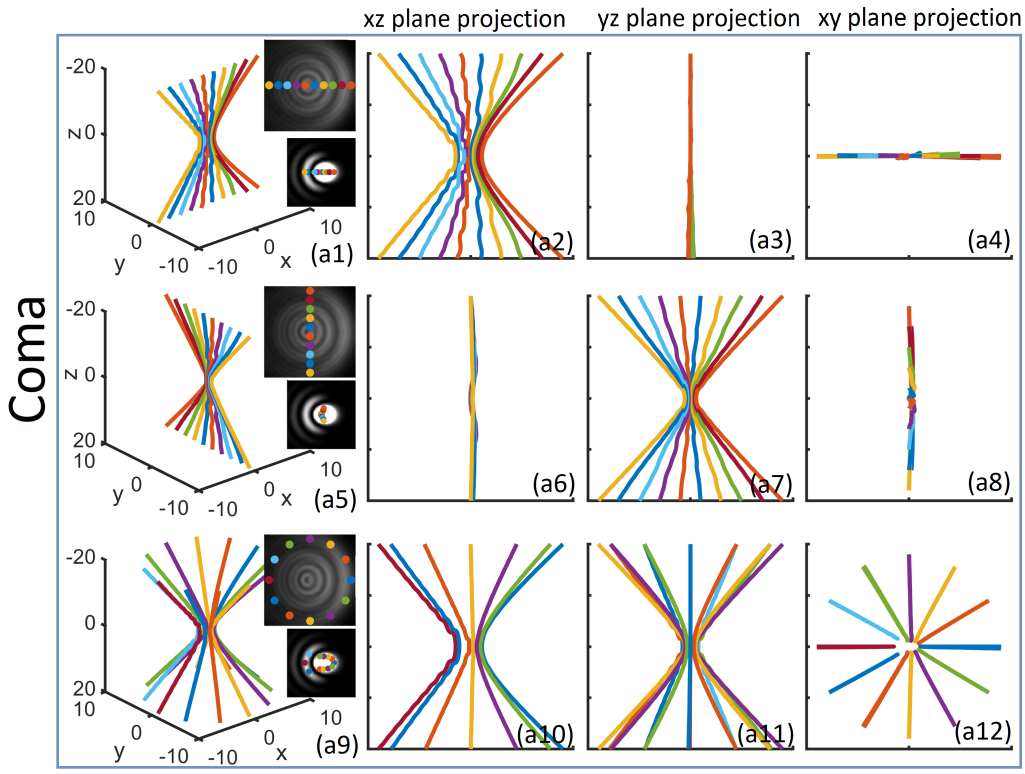

Figure 5.8: Non-linear ray tracing applied to coma aberration on the ideal Rayleigh-Sommerfeld lens with numerical aperture of 0.6 and TEM_{00} illumination computed using Algorithm 1 as described in Chapter 4. Various projections of different sets of 10 rays originating in the plane at $z = -20\mu m$: (a1)-(a4) ray origins uniformly distributed along the x-axis; (a5)-(a8) ray origins uniformly distributed along the y-axis; (a9)-(a12) ray origins uniformly distributed over a circle in xy plane. For all three cases two inset (log) intensity images are shown corresponding to the ray origin plane at $z = -20\mu m$ and the focal plane. The larger inset image has dimensions of $20\mu m \times 20\mu m$ and the smaller inset image has dimensions of $6\mu m \times 6\mu m$. Equivalent results are given in the Appendix C, Fig. 10 for the ideal lens computed using Algorithm 2.

Fig. 5.8 (a1)-(a12), again for the case of the Raleigh Sommerfeld lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 1. All of the parameters outlined in Chapter 4 are identical here (including the value of w used in the investigation of coma) except for the value of T_z. The larger inset image has dimensions of $20\mu m \times 20\mu m$ and the smaller inset image has dimensions of $10\mu m \times 10\mu m$. The first row of images (a1)-(a4) relates to the sampling of 10 ray origins along the x-axis. It can be seen from the various projections and the two intensity images, that the rays do not deviate from the $y = 0$ plane. It can also be seen that the rays do not form a tight focus, with the leftmost rays deviating towards the first crescent moon lobe. The second row of images (a5)-(a8) relates to the sampling of 10 ray origins along the y-axis. In this case, the intensity images and the xz and xy projections reveal a subtle protrusion of the rays towards the $-x$ direction and the rays come to a much tighter focus than the previous case as seen in (a7). The final row of images (b9)-(b12) relates to a circular sampling of 12 ray origins in the plane at $z = -20\mu m$. It is interesting to observe that those rays on the left side all deviate towards the first crescent moon lobe,

while other rays travel paths into the centre focal point. Once again, these observations are broadly consistent with the interpretation of coma aberration in geometrical optics, whereby the rays are interpreted as smearing in the direction of the coma.

While all of the previous results were for those low order aberrations that are commonly interpreted using classical geometrical optics, the final two sets of results are for the cases of Trefoil and Quadrafoil, which cannot be interpreted in terms of linear rays. Here we attempt to provide an interpretation in terms of the non-linear ray tracing method proposed in this chapter. The results of non-linear ray tracing applied to Trefoil aberration are presented in Fig. 5.9 (a1)-(a12), once again for the case of the Raleigh Sommerfeld lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 1. All of the parameters outlined in Chapter 4 are identical here (including the value of w used in the investigation for Trefoil) except for the value of T_z. The larger inset image has dimensions of 20μm×20μm and the smaller inset image has dimensions of 10μm×10μm. The first row of images (a1)-(a4) relate to the sampling of 10 ray origins along the x-axis. It can be seen from the intensity images in (a1) that these ray origins are along a straight line over the triangular shape that is the diffraction pattern at -20μm. It is interesting to note that the triangular pattern that appears in the in-focus plane has inverted with respect to the x-axis. It is also interesting to note that rays that originate on the right side are close to the flat side of the triangle at -20μm, but these rays end up on the right point of the in-focus triangle. Rays that originated on the left side close to the point of the triangle at -20μm end up in the left side of the in-focus triangle or in the nearly by diffraction lobe. It can be also be seen from the various projections and the two intensity images, that the rays remain in the $y = 0$ plane and that the rays do not form a tight focus, with the rays following various paths to the triangle focus or nearby lobes. The second row of images (a5)-(a8) relates to the sampling of 10 ray origins along the y-axis, which slice through the centre of the triangular diffraction pattern at -20μm. From the two intensity images in (a1) it can be seen that these rays end up distributed along a vertical arc through the centre of the inverted triangular focus pattern. From the various projections, the rays are clearly seen to bend with the degree of bending dependent on their origin position relative to the centre. The last row of image (a9)-(a12) relate to circular sampling of 12 ray origins in the plane at $z = -20$μm. From the two intensity images and the various projections, it can be seen that those rays originating from points close to the flat sides of the triangular diffraction pattern at $z = -20$μm end up propagating into the three corners of the in-focus triangle. The degree of ray bending is dependent on their position. This is most evident in (a12) where it can be seen that the ray originating at either at the centre of the flat sides, or at the corners, of the triangle at $z = -20$μm propagate through a single plane and to not bend as they propagate through the focus. Rays that are not at these points all bend as they propagate towards the outer lobes or the corners of the in-focus triangle.

The results of non-linear ray tracing applied to Quadrafoil aberration are presented in Fig. 5.9 (b1)-(b12), again for the case of the Raleigh Sommerfeld lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 1. All of the parameters outlined in Chapter 4 are identical here except here we use $w = 5 \times 10^{-6}$,

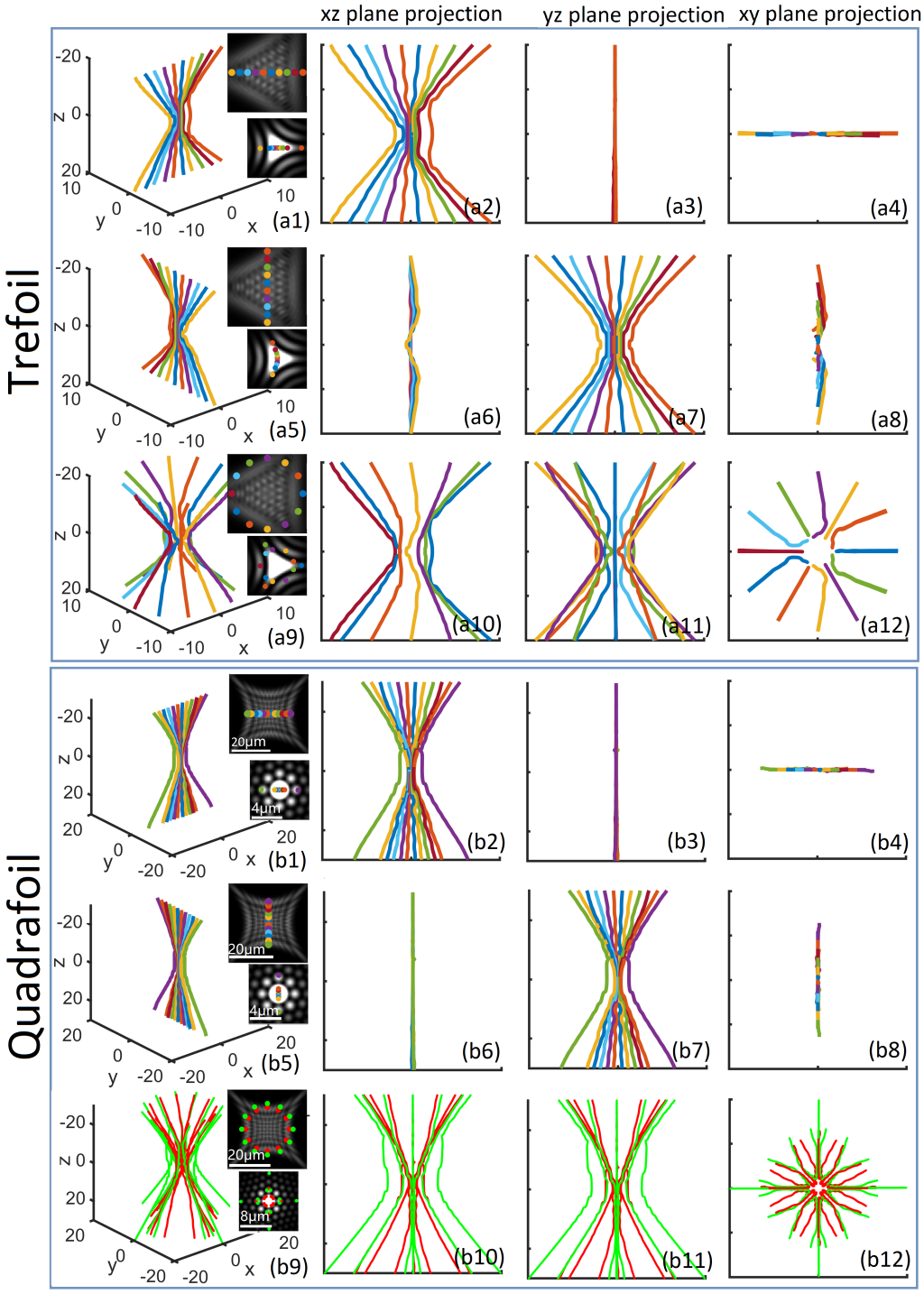

Figure 5.9: Non-linear ray tracing for Trefoil and Quadrafoil aberrations for the Rayleigh-Sommerfeld lens; see text for details.

which was was found to produce excellent results for non-linear ray tracing. The first row of images (b1)-(b4) relate to the sampling of 10 ray origins along the x-axis. It can be seen that all of these rays remain in the xz plane and most propagate to a tight focus. Those rays at the extreme edges of the diffraction pattern $z = -20\mu$m end up propagating into the left and right diffraction spots. The same effect is observed for the rays sampled over the y-axis as shown in figures (a5)-(a8). The final row of image (b9)-(b12) relates to a circular sampling of two sets of 12 ray origins in the plane at $z = -20\mu$m shown as red and green spots. It is very interesting to see that those rays on the $x = 0$ and $y = 0$ planes are not highly focused and all follow paths into the outer diffraction spots around the centre focus spot. Most inner rays (red spots) focus around the side of the center in-focus spot, while the outer rays (green spots) all follow paths into the surrounding diffraction spots.

5.5 Conclusion

In this chapter, we have proposed and demonstrated a new algorithm for tracing the flux lines through a focusing optical wavefield. This method builds on two algorithms that were developed in Chapter 4, which can be used to generate a 3D grid of samples that represent the discretized wavefield over three dimensions in the focal region of the lens. The difference between those two algorithms relate to the type of lens that is being simulated; the first algorithm simulates the effect of a thin lens and involves numerical propagation from the plane of the thin lens to the region around the focal plane; the second algorithm deals with the ideal lens, which can be modelled using an optical Fourier transform. These two algorithms can include the laser spatial mode, as well as the lens aberrations. Although the aberrations can be modelled using the well-known Zernike polynomials for both algorithms, the application of these polynomials has different physical interpretation for both cases.

Regardless of which algorithm is employed, the resulting 3D grid of samples serves as the staring point for the flux tracing method proposed in this chapter. Section 2.5.2 provides an overview of the Eikonal function, which can be used to relate the complex electromagnetic wavefront and the classical interpretation of geometrical ray optics. The Eikonal function can be used to define the flux lines, which trace the movement of the power over three dimensions as the wavefront propagates. In the same section we also review the concept of the derivative of the phase and explain how this is related to concept of a ray in the classical geometrical optics sense. We show that the the phase derivative can be used to define the angle at which the flux lines are moving at any instant of time. We proceed to develop an algorithm that traces the flux through a sequence of different propagation distances separated by a value h_z; more specifically, the phase derivative of the 'ray' is calculated at each plane to direct a linear ray trace between adjacent plane. Using the concept of the Eikonal, we prove that the accuracy of this algorithm is high so long as the step size h_z is small.

Having validated the method, the algorithm is applied to various cases of focusing different laser spatial modes and for lenses with various aberrations modelled using

the Zernike polynomials. The two different algorithms developed in Chapter 4 are applied to compute the 3D wavefield, relating to the the thin lens approximation and the ideal lens, and we find the results are very similar for both. The most striking results relate to the Laguerre-Gaussian laser spatial modes and to the effect of lens aberrations. For the former, the rays appear to spiral rapidly with a spin rate depending on the order of the LG mode and the ray position relative to the optical axis. Those closest to the centre spin most rapidly, and there appear to be a linear dependence with the Laguerre-Gaussian order, which is consistent with the existing interpretation of orbital angular momentum. It is very interesting to note the movement of the rays close to the focus donut, with many rays undertaking a rapid change of direction, and forming a bubble like shape around the focus.

Aberrations are a particular area of interest for non-linear ray tracing. Our results for the lower order aberrations, including spherical, defocus, astigmatism, and coma, were broadly consistent with the classical geometrical interpretation. Similarities between spherical and defocus aberration were drawn, with the latter exhibiting properties to the former for high levels of defocus. For the case of spherical aberration a small amount of defocus was also observed. Astigmatism was particularly interesting, and the results revealed the different focal planes for rays that were on the $x = 0$ and $y = 0$ planes. The results were also interesting for rays that originate at positions that are not on these plans. These rays appear to twist into one of the two focal points. The results for Coma, Trefoild, and Quadrafoil provided insights that cannot be gleaned from classical geometrical. The rays could be traced moving into the characteristic lobes and spots surrounding the centre spot. The rays could even be traced into the corners and sides of the triangular centre spot for the case of Trefoil. These results clearly demonstrate the power of the approach as an extension of classical geometrical ray optics. We believe the method may find future usefulness in lens design and aberration theory.

The next step is to replace the linear geometrical rays in Chapter 3, which form the basis of predicting the forces in an optical trap, with the non-linear rays that can be traced using the method presented in this chapter. This is the goal of Chapter 6.

Chapter 6

Force calculation using the flux model

6.1 Abstract

The core idea in this chapter is to augment the classical ray-optics approach to calculate the forces acting on the sphere developed by Ashkin (and further developed by us in Chapter 3), such that non-linear rays can replace linear rays in the calculation. Perhaps the greatest advantage of such an approach would be the capacity to model the orbital angular momentum imparted on the sphere using a Laguerre-Gaussian spatial mode laser. The flux model discussed in Chapter 5 is used to define the non-linear rays that intersect with the surface of the sphere and for each one of these rays, the scattering force, gradient force, and torque can be derived using the Equations defined in Chapter 3. Integration of these forces reveals the total three-dimensional force acting on the sphere as well as total rotational forces which can be decomposed into a 'vertical torque' and 'horizontal torque.' As well as investigating the single beam dielectric trap in the model of Ashkin, we additionally investigate the dual beam trap for all cases, which has the benefit of enhanced trapping forces.

6.2 Introduction

Building on the 3D numerical propagation algorithms developed in Chapter 4, we defined a method in Chapter 5 to trace the non-linear rays (flux lines) through the focal region of a lens with one particularly interesting use case being the spinning rays of a focused Laguerre-Gaussian mode laser. In this Chapter, we return full circle to the results of Chapter 3, in which we developed a method to calculate the axial and lateral trapping forces acting on a sphere in an optical trap as well as the absorption related torque that is imparted by the trap. In summary this method was based on deriving the (forward) scattering, (perpendicular) gradient, and (in-plane) rotational forces imparted by a single ray. We found that these forces relate not only to the incidence angle of the ray and the difference in refractive index, but also to the absorption coefficient of the sphere and its radius. To determine the overall axial and lateral trapping forces, we used the same approach as Ashkin, whereby the lens aperture is uniformly sampled and each position in that grid traces a straight line to the focus. The force imparted by

each of these rays is calculated and added together. However, we know from the previous chapter that the non-linear rays do not take such a simple path; two examples of this (i) the vertical direction of all of the rays in the region of the focus (see for example Fig. 5.4(a) which shows the focusing non-linear rays for a Gaussian laser with lens 0.6NA) and (ii) the spinning rays for a Laguerre-Gaussian laser (see Fig.5.4 (b) in Chapter 5 which shows the focusing non-linear spinning rays for a Laguerre-Gaussian laser with lens 0.6NA).

In this chapter, we employ the same approach as in Chapter 3 to calculate the axial, lateral, and rotational trapping forces, except that this time the rays that are incident on the sphere are not geometrical straight lines, they are the non-linear rays or flux-lines that can be traced using the methods of Chapter 4 and 5. The advantage of such an approach is that the effects of wave optics and diffraction can be accounted for in the model, and even more interestingly, the orbital angular momentum that can be transferred from Laguerre-Gaussian laser modes can now be accounted for using Ashkin's model. We name this new hybrid wave-optics/ray-optics model for estimating the trapping forces as the 'flux-ray model' and the computational steps that are used in the model are given in the next section. This is followed by a comparison of the flux-ray model and the extended model in predicting the trapping forces on spheres of various sizes and absorption. We investigate single beam traps as well as dual-beam traps (as described in Chapter 2) using both TEM_{00} and TEM_{01} illumintaion. For the latter case we demonstrate that the horizontal torque on the sphere, transferred via the orbital angular momentum of the Lagerre-Gaussian laser mode can be estimated using the flux-ray model but not by the other models.

One important point of note is that all of the lenses that are investigated in this chapter are ideal lens with numerical aperture of 0.6 or lower. The reason for this lies in the accuracy of the Angular Spectrum method which is known to be accurate only for numerical apertures less than 0.7 as discussed in Chapter 4. Since this algorithm is at the heart of the flux-ray method it prevents the investigation of higher NA lenses.

6.3 The flux-ray model for calculating optical trapping forces

In Fig. 6.1 the basic principle of the algorithm is illustrated. The sphere position is moved in the focal region of the lens. For example, in Fig. 6.1(a) three positions are shown along the z-axis. In each sphere position, the non-linear rays (calculated using the method described in Chapter 5) that are incident on the surface are subject to the extended ray optics force calculation for a single ray that was described in detail in Section 3.3 Chapter 3. This process is illustrated in Fig. 6.1(b). Each non-linear ray is taken to be linear as it intersects the surface and is, therefore, assumed to impart a scattering force, a gradient force, and a torque as described in Chapter 3. This method calculates the force imparted by the ray as a function of ray angle with respect to the surface normal, as well as the relative refractive indices of the sphere and surrounding medium, the polarisation of the laser, and the absorption coefficient of the sphere.

Another feature of the method is that the force imparted by the ray is also weighted by the intensity of the wavefield at the ray origin on the surface. The forces imparted from all of the non-linear rays over the surface are added together to produce the total forces acting on the sphere in that position. The entire process is repeated over a sequence of different sphere positions in order to produce a plot of the forces acting on the sphere as a function of its position relative to the laser focus, both laterally and axially. The details of the method are described in the algorithm below, including a description in Step 2 of how we 'uniformly' sample the ray positions on the surface of the sphere.

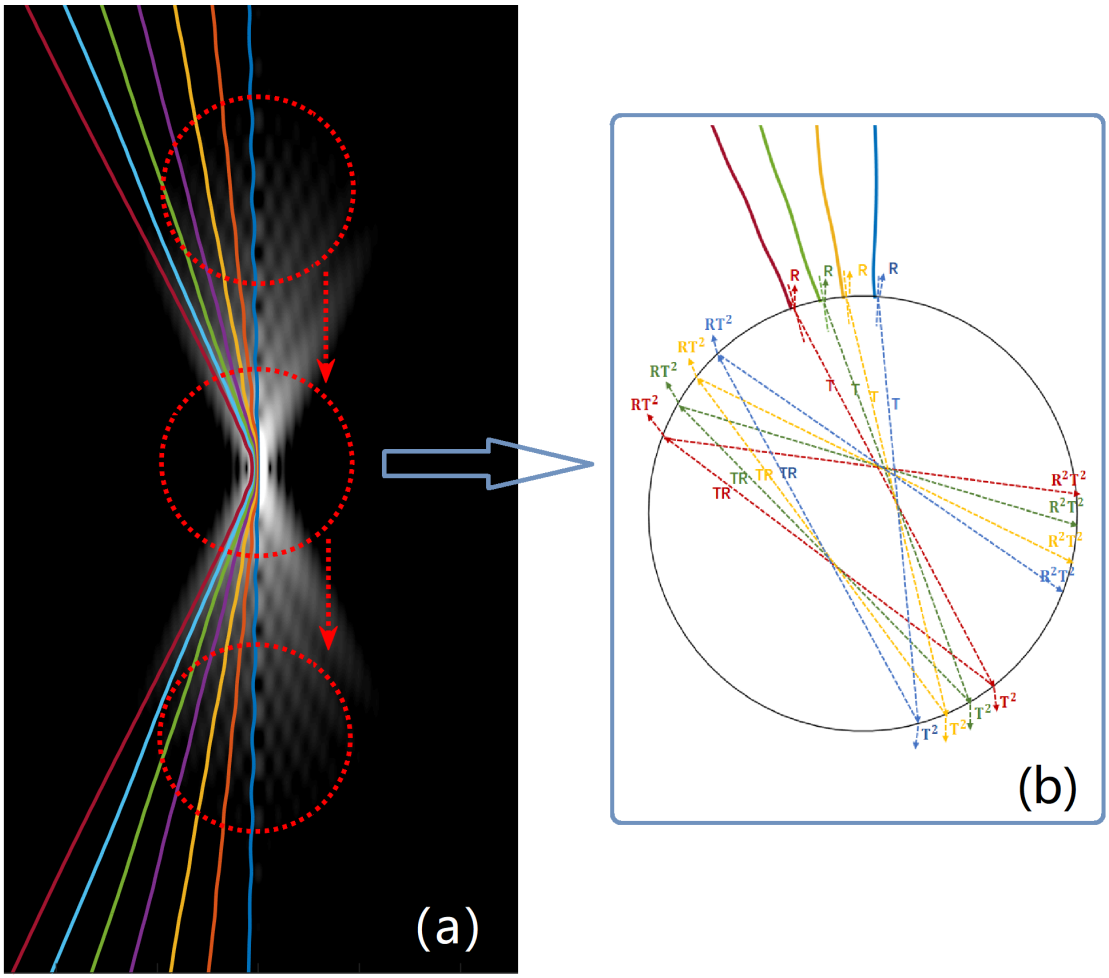

Figure 6.1: Illustration of the flux-ray model developed in this chapter for calculating the forces on a trapped particle. (a) The sphere position is moved along the optical axis, and at each position the non-linear rays (calculated using the method described in the previous chapter) that are included on the surface are sampled. (b) The forces imparted by each one of these rays is calculated according to the extended ray force method described in Chapter 3 and added together to give the total forces acting on the sphere, including the torque.

6.3.1 Algorithm

The algorithm is defined as a sequence of eight steps below, several of which are repeated iteratively as described in the text and flowchart provided at the end of the section.

1. **SELECTING THE PARAMETERS.** The first step is to define the various parameters necessary to implement Algorithm 1 or Algorithm 2 as described in Chapter 4. This includes selecting the properties of $u_z(x,y)$ including the focal length of the lens, f, the numerical aperture and aperture radius, R, and laser mode A. This also includes the selection of the sampling intervals T_x, T_y, and T_z. Other parameters that must be defined are the radius of the trapped particle r, the refractive index of the sphere and the medium n_1 and n_2, the absorption coefficient a, and the polarization state of the laser as described in Chapter 3. The set of discrete positions of the sphere centre (x_m, y_m, z_m) over which the forces are to be calculated must also be defined, where $m = 1 \rightarrow M$. Steps 2-7 will be implemented for each of these discrete positions.

2. **RANDOMLY SAMPLING THE SURFACE OF THE SPHERE.** The second step is to 'uniformly' sample the surface of the sphere to obtain a set of discrete points on the surface at which the non-linear rays that are focusing from the lens will be calculated using the algorithm defined in Chapter 5. For a given sphere centre position (x_m, y_m, z_m), a set of points on the surface must be identified for calculating the ray forces and these points must be evenly distributed over the surface; If this is not the case, the area on the surface that is more densely sampled will be taken to impart more force than an area that is sparsely sampled, even if the wavefront (and by extension the ray angles) is identical for both regions. Sampling the surface of a sphere uniformly is problematic: for example, using uniform angular steps to define the sample points results in a nonuniform distribution of the samples with respect to surface area. To overcome this problem we adopt a dense random sampling of the surface to define the ray positions. The discrete set of random points on the surface of the sphere are described as (x_{mn}, y_{mn}, z_{mn}) $n = 1 \rightarrow N$. This is implemented in Matlab using a random number generator to define random angles in the xy-plane and yz-planes. An illustration of a random selection of points on the surface of the sphere is given in Fig. 6.2 appearing as red dots in the figure. Note the overlay of the 3D samples of $u_z(x,y)$ calculated in the next step. These appear as blue dots and are sampled at a much lesser rate than the ray positions. Although sampled less densely than the ray positions on the surface, we note that the values of T_x, T_y, and T_z are significantly smaller than the spacing shown in the figure between the blue dots.

3. **CALCULATION OF THE 3D GRID OF $u_z(x,y)$ IN REGION OF THE SPHERE OVER LIMITED BAND.** The third step is to run Algorithm 1 or 2 as described in Chapter 4 using the parameters in Step 1. The range of distances of propagation over z will be conveniently bounded by the centre position $z_m - r \rightarrow z_m + r$. However,

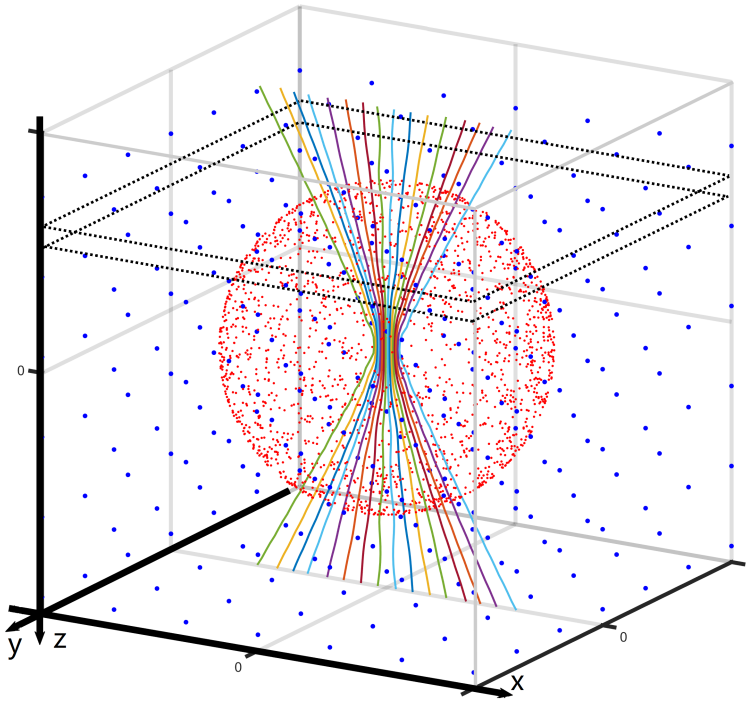

Figure 6.2: Illustration of the random sampling of the sphere surface shown in red dots and the (larger) 3D uniform sampling of $u_z(x,y)$ in blue dots (not to true scale). The complex field at the sampling positions on the surface can be obtained by interpolating the uniform 3D samples so long as u_z has been Nyquist sampled. To save memory this process is done over small bands highlighted by the dotted black lines. A single band contains only six z-slices.

to reduce memory load we break this range up into smaller bands and only six slices are calculated and stored at any one time; this is similar to the approach taken in the previous chapter for training flux lines over a range of propagation distances and is illustrated in Fig. 6.2 using the black dotted lines. Only those ray positions on the surface (x_{mn}, y_{mn}, z_{mn}) that lie within this band are subject to the next step. The result of this step is that we have a well sampled grid of points for $u_z(x,y)$ within the band of interest, as well as the desired ray position on the sphere surface within this band.

4. **3D INTERPOLATION TO OBTAIN VALUES OF $u_z(x,y)$ AT RAY POSITIONS ON SURFACE** 3D interpolation is applied to the grid of points for $u_z(x,y)$ within the band of interest in order to obtain the complex values of $u_z(x,y)$ at the ray positions on the surface; more accurately this interpolation provides the values of $u_z(x,y)$ at four points around the ray positions on the surface (2 points slightly shifted in x and 2 points slightly shifted in y as described in the previous chapter.) For each point on the surface, the local cosine angles of the ray α_{ln} and β_{ln} are calculated according to Equation 5.2. The process is illustrated in Fig. 6.3; a projection of the ray positions on the surface of the sphere is shown in Fig. 6.3 (a) where the ray positions are shown as red dots overlaid with the position of the samples of u_z shown as black dots distributed over six slices. In Fig. 6.3 (b) the

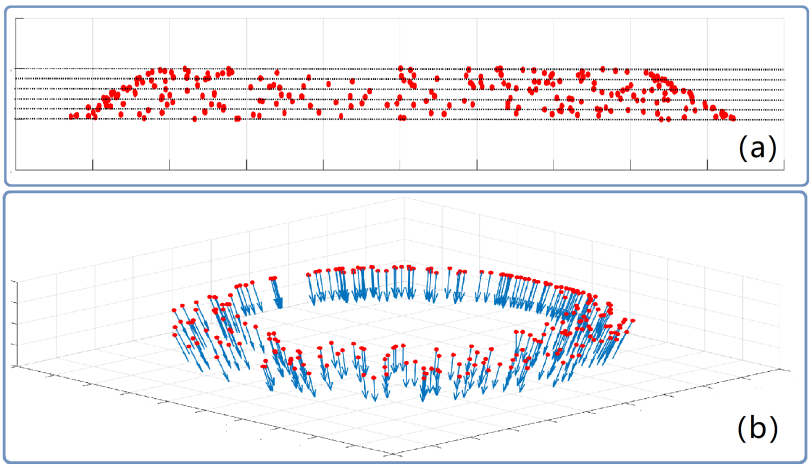

Figure 6.3: Illustration of the ray angles entering the sphere over a given band. (a) shows the random sampling on the surface over a vertical projection of the sphere. Also shown in the figure with black dots are the samples of the uniformly sampled u_z; (b) shows the rays associated with these positions. Ray angles are calculated using the methods described in Chapter 5. The complex values of the fields around the random points are calculated using 3D interpolation.

resultant ray angles are shown for these random positions.

5. **CALCULATE THE ANGLE OF INCIDENCE FOR EACH RAY** The fifth step is to calculate the angle of incidence θ_{mn} for each of the rays that were investigated in Step 4. Since we know the direction cosine $\vec{k} = (\alpha, \beta, \gamma)$ for the point of incidence (x_{mn}, y_{mn}, z_{mn}) for each ray, we calculate the normal vector $\vec{l} = (x_m - x_{mn}, y_m - y_{mn}, z_m - z_{mn})$ from the surface point to the sphere centre. The angle of incidence can then be calculated as follows: $\theta_{mn} = \arccos\left(\frac{\vec{k}\cdot\vec{l}}{|\vec{k}|\times|\vec{l}|}\right)$.

6. **CALCULATE THE FORCES OF EACH RAY** This step involves applying the previous equations developed in Chapter 3 that calculate the scattering and gradient forces, and the torque, for each ray. Specifically, Equations 3.8, 3.9 and 3.20 are applied using the values of θ_{mn} derived in the previous step as inputs. The trapping forces can be expressed as ratios relative to the input forces, Q_s and Q_g. To obtain the force in terms of Newtons, these values are scaled by the factor $\frac{n_1 P}{c}$ where P is given by the intensity of the field at (x_{mn}, y_{mn}, z_{mn}) scaled by the sampling area (the latter is defined as the total sphere area divided by the number of ray positions N). The scattering and gradient forces are calculated and decomposed into x-, y-, and z- projections such that the forces from multiple rays can be conveniently added together. The torque imparted by each ray is decomposed into a 'vertical torque' and a 'horizontal torque', which are perpendicular to each other for the same reason.

7. **REPEAT FOR ALL BANDS** This step involves repeating Steps 3 through 6 until all of the bands over the range of distances extended by the sphere diameter have been calculated and all of the resultant forces are integrated to give the total forces on the particle when it is centred at (x_m, y_m, z_m).

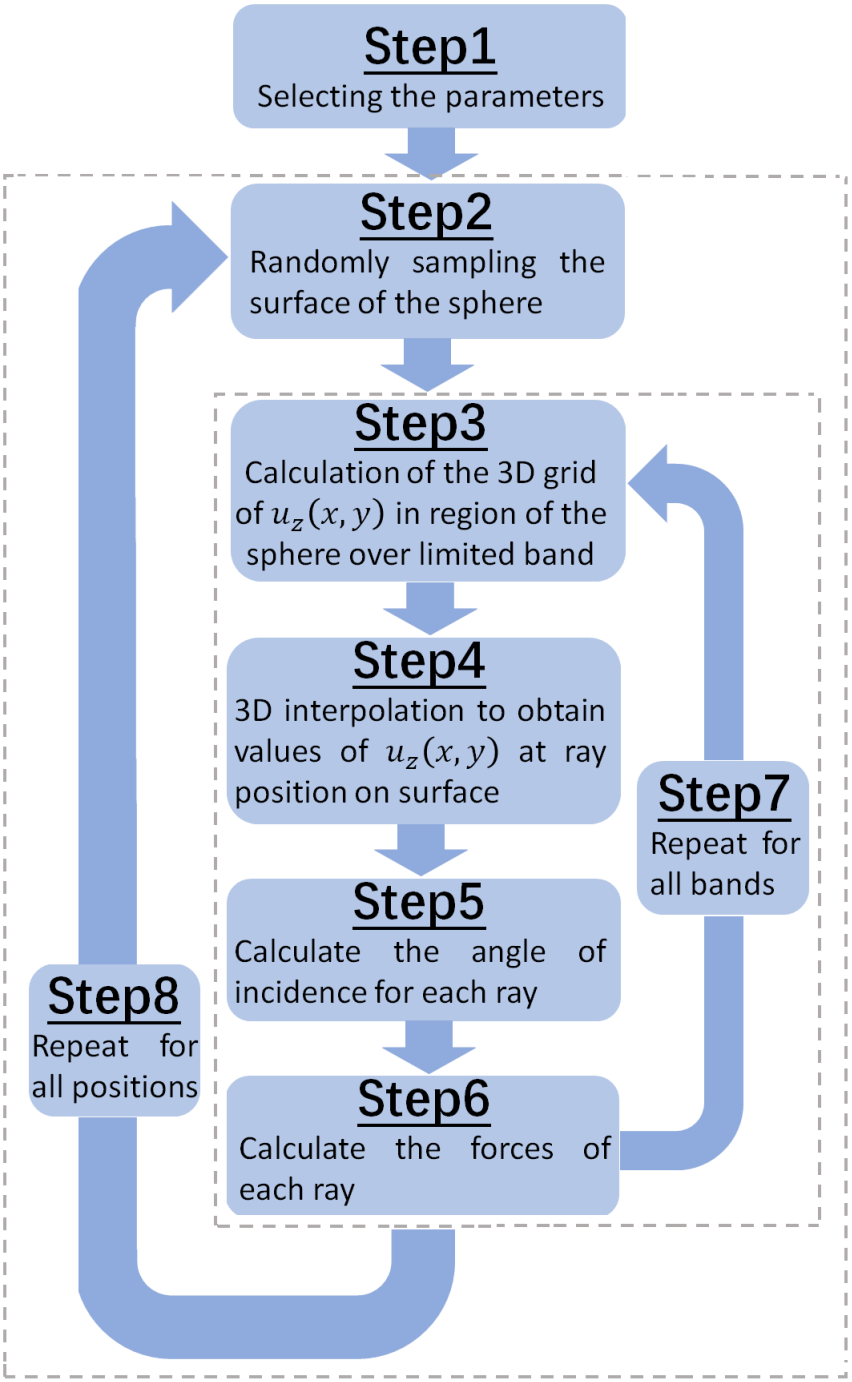

Figure 6.4: The various steps of the algorithm as described in section 6.3.1.

8. **REPEAT FOR ALL POSITIONS** Steps 2-7 are repeated for the full set of positions (x_m, y_m, z_m) where $m = 1 \rightarrow M$ as defined in the first step.

In the following section the results of this method are provided for several cases of different lenses with different numerical apertures, different laser spatial modes, and different sphere absorption coefficients.

6.4 Single Beam Traps Calculated using the Flux-Ray Model

In this section, several sets of results are presented for the axial, lateral, and rotational trapping forces calculated using the algorithm described in the previous section. In all cases the lens is an ideal lens that is numerically propagated using Algorithm 2 developed in Chapter 4. All of the parameters that were used in that simulation are identical to those given in Chapter 4 for the simulation of an ideal lens with no aberrations with numerical aperture of 0.6 and with either a TEM_{00} or TEM_{01} laser spatial mode. We limit the numerical aperture that is simulated to 0.6 since it in known that the Angular Spectrum method is accurate, which is employed by both Algorithm 1 and Algorithm 2, which are at the heart of the flux model, is only accurate for NA$<$ 0.7.

6.4.1 Gaussian Illumination

The trapping forces on three spheres with varying radius and absorption coefficients are shown in Figure 6.5 for the case of a TEM_{00} laser and an ideal lens of NA 0.6. The first set of results shows the forces on the particle calculated using Ashkin's method [21], as displayed in (a) for axial and (b) for lateral displacement; only a single result is shown here which is consistent for all three particles, i.e. Ashkin's model cannot account for absorption. A similar set of results is provided in Chapter 3 for NA of 1.3 for which case the gradient force is significantly stronger and the total force passes through zero much more closely to the centre indicating that the trapping position is further displaced from the centre relative to the NA of 1.3. The corresponding set of results are shown for the same two cases in (c) and (d) using the extended method developed in Chapter 3; similar results are also found in Chapter 3 for NA 1.3. It is interesting to see how the scattering forces increase for the particles with higher levels of absorption. The consequence of this is that the total forces increase such that these never reduce to a zero value indicating that for these two cases there is no trapping position - the trap will fail and the particle will simply be pushed away for all axial positions. The final two figures in (e) and (f) relate to the flux-model proposed in this chapter. There are several similarities and differences between the results from the extended model and the flux-model. One interesting point is the change in the values for torque from the for small lateral displacement. The torque can be decomposed into two components, 'vertical torque' and 'horizontal torque', which are perpendicular to each other. Owing to the symmetry property of the Gaussian beam, horizontal components cancel out for both models, leaving only the vertical component. It is also interesting to

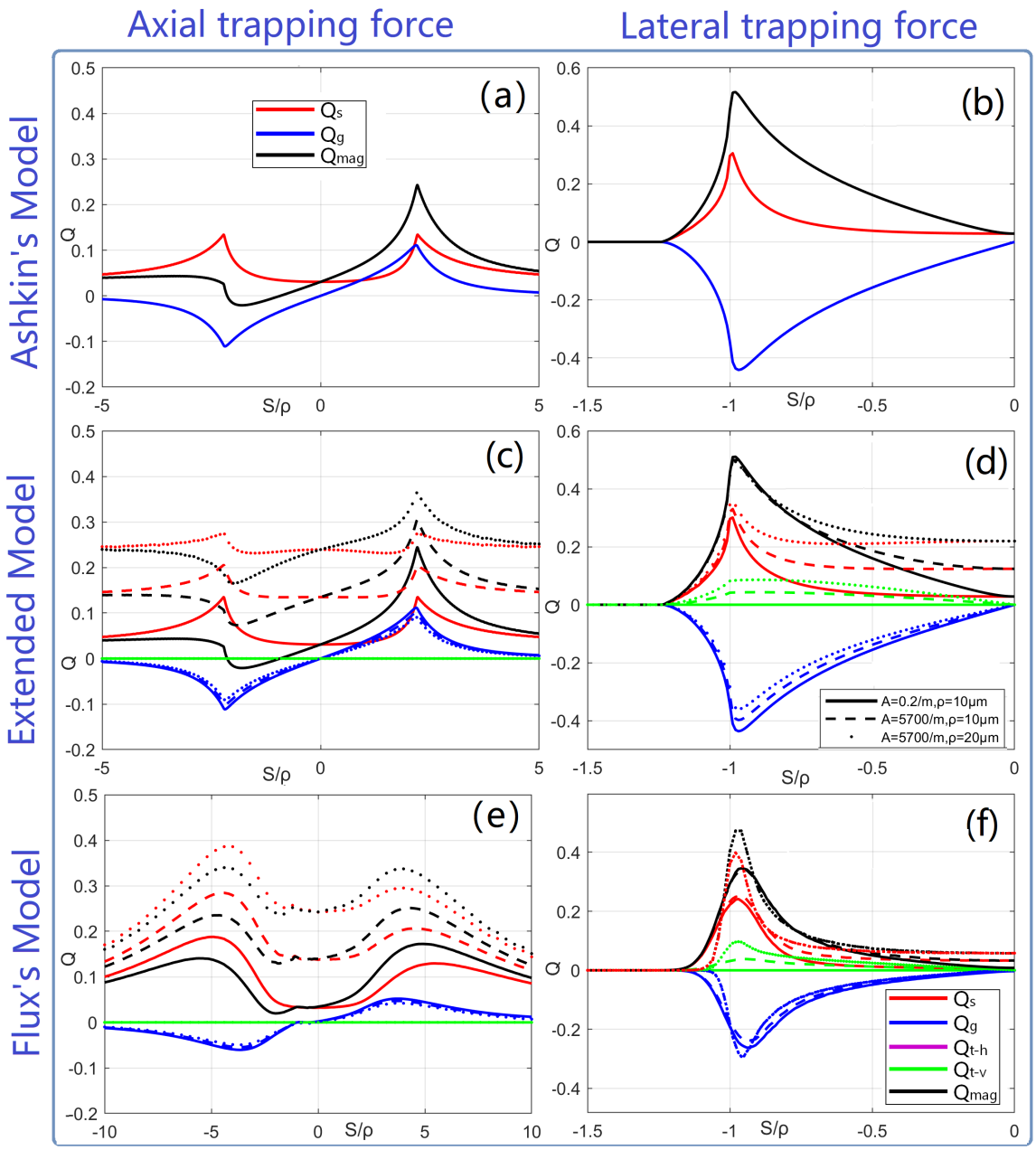

Figure 6.5: Prediction of the axial and lateral forces imparted by a laser focused from an ideal aberration-free lens with NA 0.6 and TEM_{00} laser mode with $1/e^2$ cut-off, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. All scattering and gradient forces are plotted as red and blue lines and the total force is given in black. Vertical torque is plotted in green. The three different spheres are identified using different dotted or dashed lines as defined in the legend. S/ρ is the relative value of distance and radius.

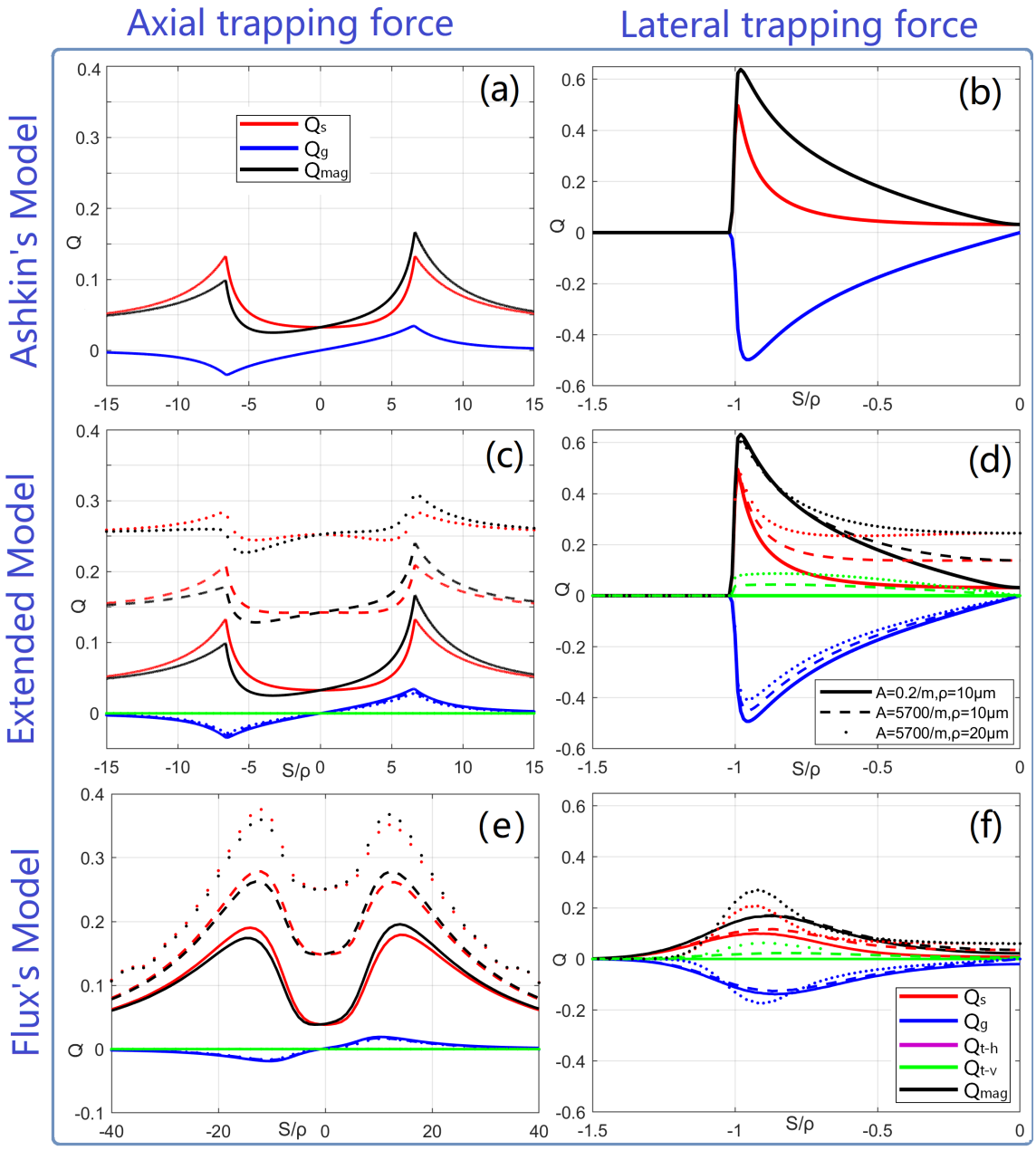

Figure 6.6: Prediction of the axial and lateral forces imparted by a laser focused from an ideal aberration free lens with NA 0.2 and TEM_{00} laser mode with $1/e^2$ cut-off, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. All scattering and gradient forces are plotted as red and blue lines and the total force is given in black. Vertical torque is plotted in green. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

note that both the scattering force and gradient force lose their symmetry. For both the scattering force and the gradient force, the left peak is 1.2 times stronger than the right peak. This phenomenon is caused by the different angles of incidence of the rays that are incident on the sphere using the two different models. In (f), the maximum amplitude of the lateral forces are approximately the same for all three models, while the decay of force is faster when the sphere moves along the y-axis of the flux model. Overall these results are broadly consistent across the various models.

A similar set of results is presented in Fig 6.6 . This time for the case of an ideal lens with numerical aperture of 0.2. All other parameters remain the same including the use of TEM_{00} laser. It can be seen from (a) that Ashkin's model predicts a similar scattering force for axial displacement as for the corresponding result in the previous figure for NA of 0.6. However, the gradient force is significantly weakened, resulting in a weaker total force that never reduces to a zero value, indicating that there is no natural trapping position. Interestingly, the trapping force for lateral displacement shown in (b) are very similar to the corresponding result for NA of 0.6 and are actually slightly stronger in amplitude. The extended model also predicts a similar set of results as for the NA 0.6 case. Indeed the three scattering forces shown in (c) are very similar to the corresponding results in the previous figure. However, owing to the weaker gradient force, the resultant total force will never reduce to zero and there can exist no optical trap in the axial dimension. The lateral trapping forces are almost identical to the previous result for the NA 0.6 case. The final two figures in (e) and (f) relate to the flux-model proposed in this chapter. Once again, there are several similarities and differences between the results from the extended model and the flux-model for the NA 0.2 case. In terms of maximum amplitudes of the axial forces for both models, the values are different but more significantly the peak values appear at different axial displacement. This is likely caused by the relative slow change of angle of incidence due to the low focus of 0.2NA; many of the non-linear rays will be approximately vertical over a long focal distance. As for (f), the forces are almost half in magnitude relative to those obtained using Ashkin's method [21] and the extended model. The forces also appear to be significantly more spread out for the flux model.

6.4.2 Laguerre-Gaussian Illumination

In this section, the same three spheres investigated in the previous section using TEM_{00} illumination, and the same two ideal lens numerical apertures 0.6 and 0.2 are investigated. However, the TEM_{00} laser is replaced with a TEM_{01}, which of course carries an orbital angular momentum that is known, through practical experiments, to be transferred to a trapped particle causing it to spin. The goal here is to show that the flux-ray model can in some way account for the transfer of a rotational force for these cases. The results of the trapping forces on the three particles are shown in Figure 6.5 for the lens NA of 0.6. The first row shows the forces on the particle calculated using Ashkin's method [21]: (a) as a function of axial displacement and (b) as a function of lateral displacement. It is seen that the gradient force and the total force for axial displacement is superior to the same result for TEM_{00} illumination. This is due

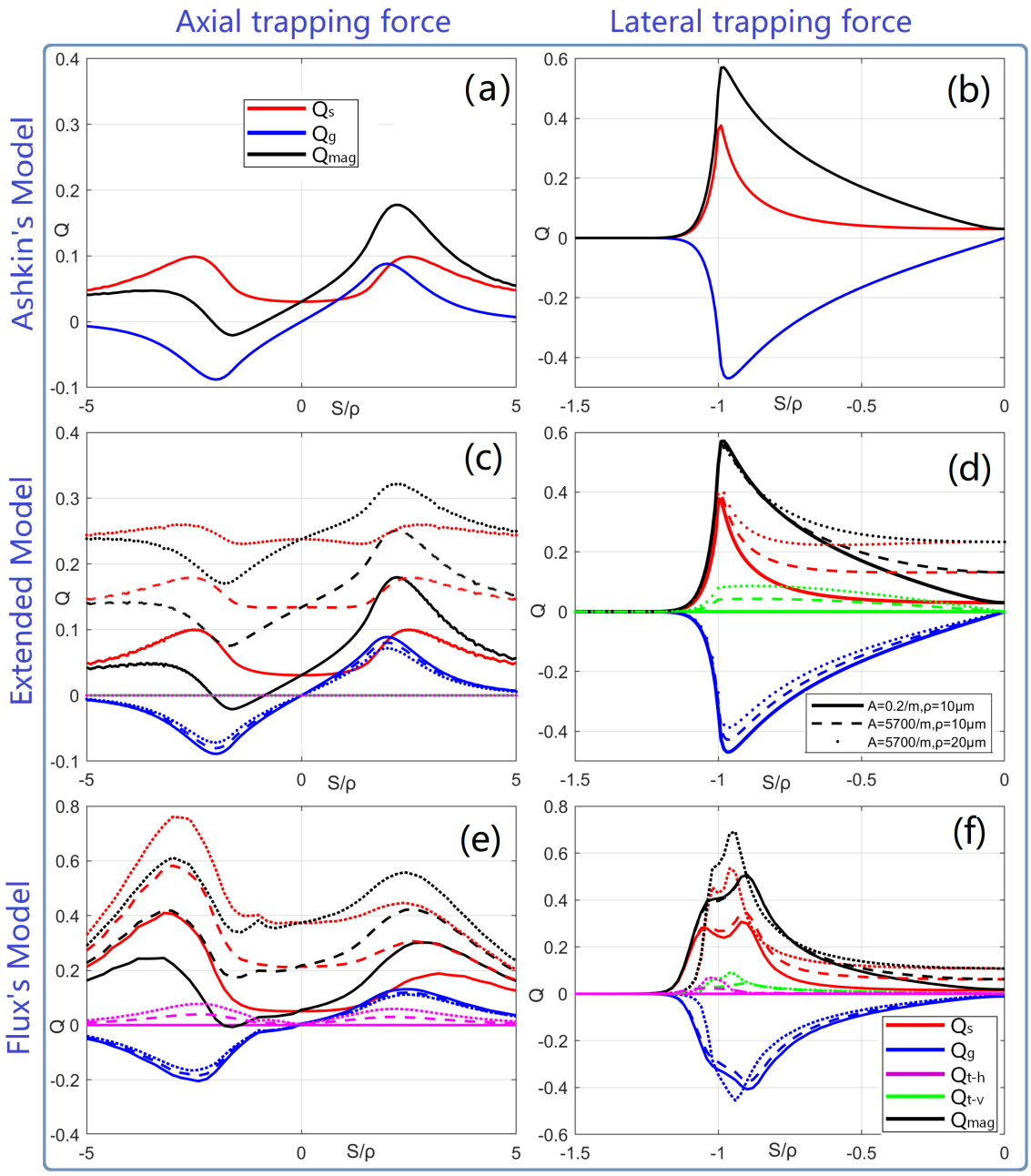

Figure 6.7: Prediction of the axial and lateral forces imparted by a laser focused from an ideal aberration-free lens with NA 0.6 and TEM_{01} laser mode with $1/e^2$ cut-off, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. All scattering and gradient forces are plotted as red and blue lines and the total force is given in black. Vertical torque is plotted in green. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

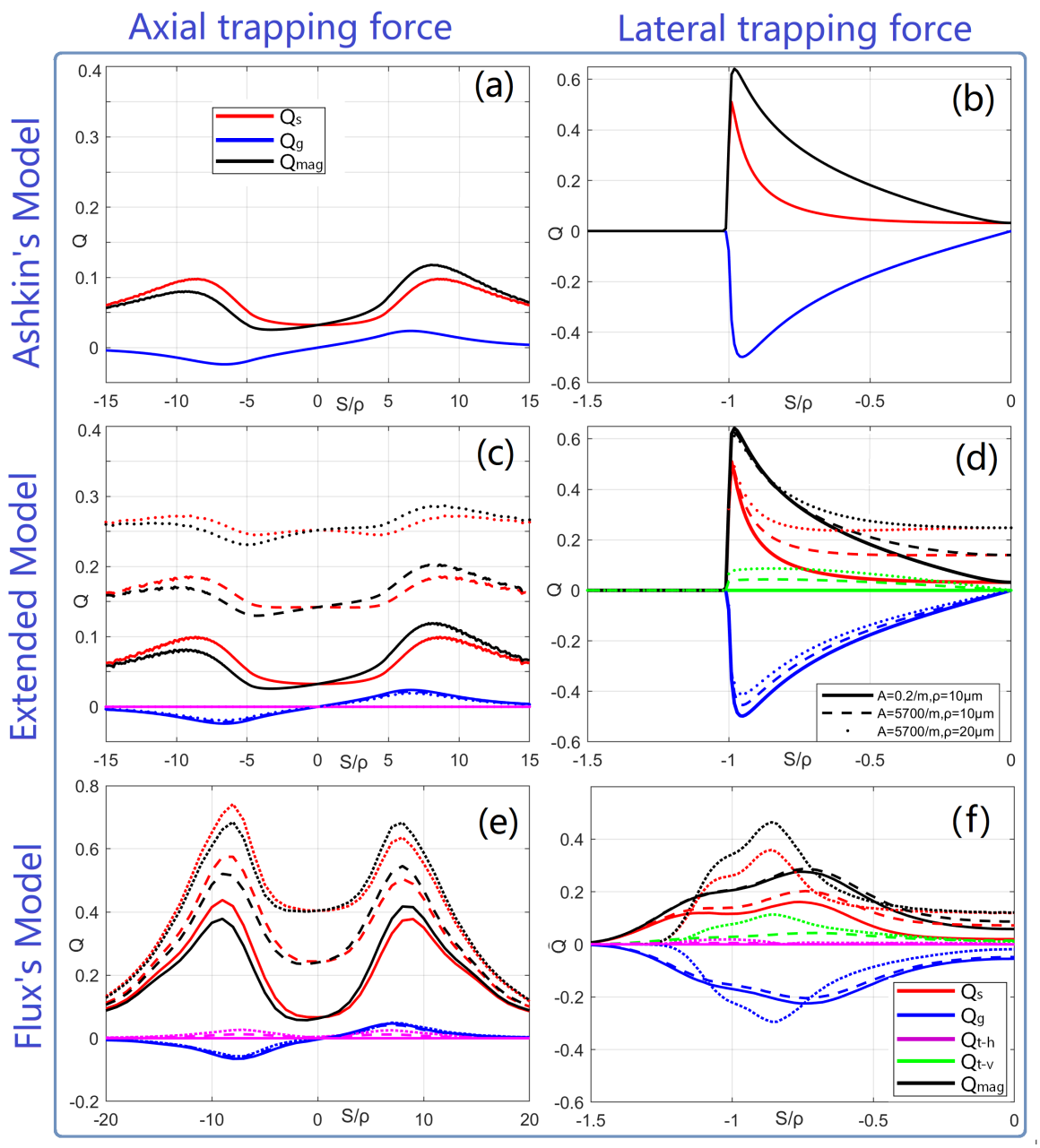

Figure 6.8: Prediction of the axial and lateral forces imparted by a laser focused from an ideal aberration-free lens with NA 0.2 and TEM_{01} laser mode with $1/e^2$ cut-off, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. All scattering and gradient forces are plotted as red and blue lines and the total force is given in black. Vertical torque is plotted in green. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

to the donut shape of the laser illumination; the on axis rays are given zero weighting while those rays propagating at high angles are weighted strongly as a function of the Laguerre polynomial. The result is a stronger trapping force and a zero crossing (the axial trapping position) closer to the centre focal point. The trapping forces for lateral displacement appear to be very similar when compared to the same result for TEM_{00} illumination. The corresponding set of results is shown for axial (c) and lateral (d) displacement calculated using the extended model and all cases are very similar to the same results for TEM_{00} illumination, except for a small difference in the gradient force for the axial displacement. Once again the vertical torque is seen to manifest for lateral displacements. It is interesting to note that neither the Ashkin nor the extended model can predict the presence of the horizontal torque imparted by the L-G orbital angular momentum. This is because all rays in these two models are modelled as travelling along straight lines from the lens to the focal point of the lens. The spinning of the rays cannot be modeled by either approaches, but can by the flux model. The final set of results is shown for the flux model in (e) and (f). The forces predicted for axial and lateral displacement of the three spheres are shown in (e) and (f), respectively. The trapping forces in (e) are highly asymmetrical as a function of axial displacement, and the gradient forces are somewhat weaker than the corresponding results for the extended model. Perhaps the most significant result is the emergence of a horizontal torque, shown in pink, which appears to show that a spin will be imparted to the particle. It can be seen that this rotational force increases as a function of sphere positional distance from the centre of focus. In (f), the forces acting on the particle as a function of lateral displacement are shown; there are several interesting observations from this figure. The overall shape is similar to that found the extended model. However, there can be two local maxima, rather than one, as the particle approaches a displacement of one radius. It can be reasoned that this is due to the shape of the donut pattern; the edge of the sphere passes over both the left and right sides of the donut causing a slight drop and rise again in force as it does so. Another point of note is the presence of both a horizontal and vertical rotational force, which reaches maximum values at approximately one radius displacement.

A similar set of results is presented in Fig. 6.8 this time for the case of an ideal lens with numerical aperture of 0.2. All other parameters remain the same including the use of TEM_{01} laser. It can be seen from (a) that Ashkin's model predicts a similar scattering force for axial displacement as for the corresponding result in the previous figure for NA of 0.6. However, the gradient force is significantly weakened, resulting in a weaker total force that never reduces to a zero value, indicating that there is no natural trapping position. Interestingly, the trapping force for lateral displacement shown in (b) is almost identical to the corresponding result for NA of 0.6. The extended model also predicts a similar set of results as for the NA 0.6 case although in this case, the total forces for all three particles will never reduce to zero and there can exist no optical trap in the axial dimension. The lateral trapping forces, however, are once again almost identical to the previous result for the NA 0.6 case. The final two figures in (e) and (f) relate to the flux-model proposed in this chapter. Once again, there are several similarities and differences between the results from the extended model and the

flux-model for the NA 0.2 case. The flux model predicts much stronger forces than the other two models; however, it also predicts no zero crossing points. The total forces in (e) is much more symmetrical in relation to axial displacement about the focal point than the corresponding result for NA of 0.6. Also, the horizontal torque can be seen and has similar form to the corresponding result for NA of 0.6, albeit the strength of this force is much weaker. As for (f), the forces are almost half in magnitude relative to those obtained using Ashkin's method [21] and the extended model. The forces also appear to be significantly more spread out for the flux model.

6.5 Dual Beam Traps Calculated using the Flux-Ray Model

In this section, we model the forces imparted by dual beam optical trap. The concept for such a trapping mechanism was introduced in Chapter 2. The principle is that a sphere is trapped at the focal point of two opposing lasers propagating in opposite directions each focused by an independent lens, i.e. the particle is effectively 'sandwiched' between two opposing optical traps. The reason for such an approach is the natural cancellation of the two scattering forces from each of the two lasers and the summing of the two gradient forces from both of which has a zero crossing point. The result is that lenses with lower numerical apertures, such as for example 0.6 or even as low as 0.2, can be used to produce stable optical traps in the axial direction. As we have seen in the results in the previous section, all of the various models predict that axial trapping cannot be provided by a single beam optical trap from a 0.2NA lens with either TEM_{00} of TEM_{01} illumination; the results for 0.6NA were also not promising with only weak traps possible. An additional benefit of the dual beam low NA approach is the much wider field of view that is afforded by lenses with lower numerical aperture. This facilitates the use of spatial light modulators to implement holographic optical tweezers that can be use to generate multiple optical traps simultaneously that can move dynamically over a wide range. The facility to image the scene over a wide field enables the user to target different particles and move them dynamically.

A related area of research is known as mirror trapping, [35, 40–42] which was also introduced in Chapter 2, whereby an SLM is illuminated with a TEM_{00} laser and then expanded to the back aperture of a low numerical aperture microscope objective such as a 0.2NA chromatic. Two distinct patterns are displayed on the SLM: The first is a flat phase, the result of which will be a standard laser focus at the traditional focal plane of the MO (this is the single beam trap we have analysed already for 0.2NA TEM_{00}), and the second is a slowly varying 2D chip, similar to the defocus Zernike polynomial, which results in a second laser focus at some plane (approx 100-200μm) after the traditional focal plane. A mirror is used to reflect this second laser focus such that it is approximately coincident with the first. The net result of this approach is that a dual beam optical trap is effectively achieved using only a single low numerical aperture objective. Although it is possible for us to simulate such a system by applying Algorithm 2 from Chapter 4 to the phase pattern on the SLM, we do not do so here, electing only

to investigate the simpler dual beam approach.

6.5.1 Gaussian Illumination

In this section, results are provided for the forces acting on a sphere from a dual beam trap for the case of TEM_{00} illumination. The trapping forces on the same three spheres with varying radius and absorption coefficients (that were investigated in the previous section for a single beam trap) are shown in Figure 6.9 for the case of a TEM_{00} laser and an ideal lens of NA 0.6. In this case, we do not include the various scattering forces and gradient forces in the figures, which would be difficult to include in a single figure. Instead we show only the total force acting on the sphere when the total forces from both laser focus have been added together. The method for calculating the axial forces for all three models in parts (a), (c), and (e) is simply to add the total force/displacement distribution for the single beam 0.6NA trap in Fig. 6.5 together with the same force/displacement distribution having been rotated about the origin by π radians. The scattering forces effectively cancel, while the gradient forces double. In all cases, in order to provide comparable results with the previous single beam cases, the laser power is halved, such that the total laser power is equal for the single and dual beam simulations. The method for calculating the lateral forces for all three models in parts (b), (d), and (f) is to apply the exact same approach; once again the scattering forces will cancel and the gradient forces will double.

The first set of results shows the forces on the particle calculated using Ashkin's method [21], as displayed in (a) for axial and (b) for lateral displacement; as before, only a single result is shown here which is consistent for all three particles. It can be seen that the axial trapping force for two lenses with 0.6NA is comparable to the axial trapping force shown in Chapter 3 for a single beam trap with a lens of 1.3NA. Unlike the single beam 0.6NA case, the total force passes through zero much more closely with strong positive and negative forces acting for positive and negative axial displacements indicating a strong axial trapping position close to the focal point. The lateral force stiffness shown in (b) are superior to the equivalent result for a single beam trap, particularly for the higher absorbing spheres, due to the cancellation of the scattering forces. The corresponding set of results are shown for the same two cases in (c) and (d) using the extended method developed in Chapter 3 and using the same methods described above to generate the dual beam trapping forces using the single beam forces. It is interesting to see just a slight change in the trapping forces due to higher levels of absorption and the trapping position is the same for all the spheres. The final two figures in (e) and (f) relate to the flux-model proposed in this chapter. Although weaker, the trap stiffness for the axial cases are similar and the overall shape of the forces are similar. It is likely that the lower trapping forces that are in general predicted by the flux-model are caused by the different angles of incidence of the rays that are incident on the sphere using the two different models. One interesting point is the change in the values for torque from the lateral displacement. The 'vertical torque' has effectively vanished because the two opposing vertical torques from the two lasers will cancel each other.

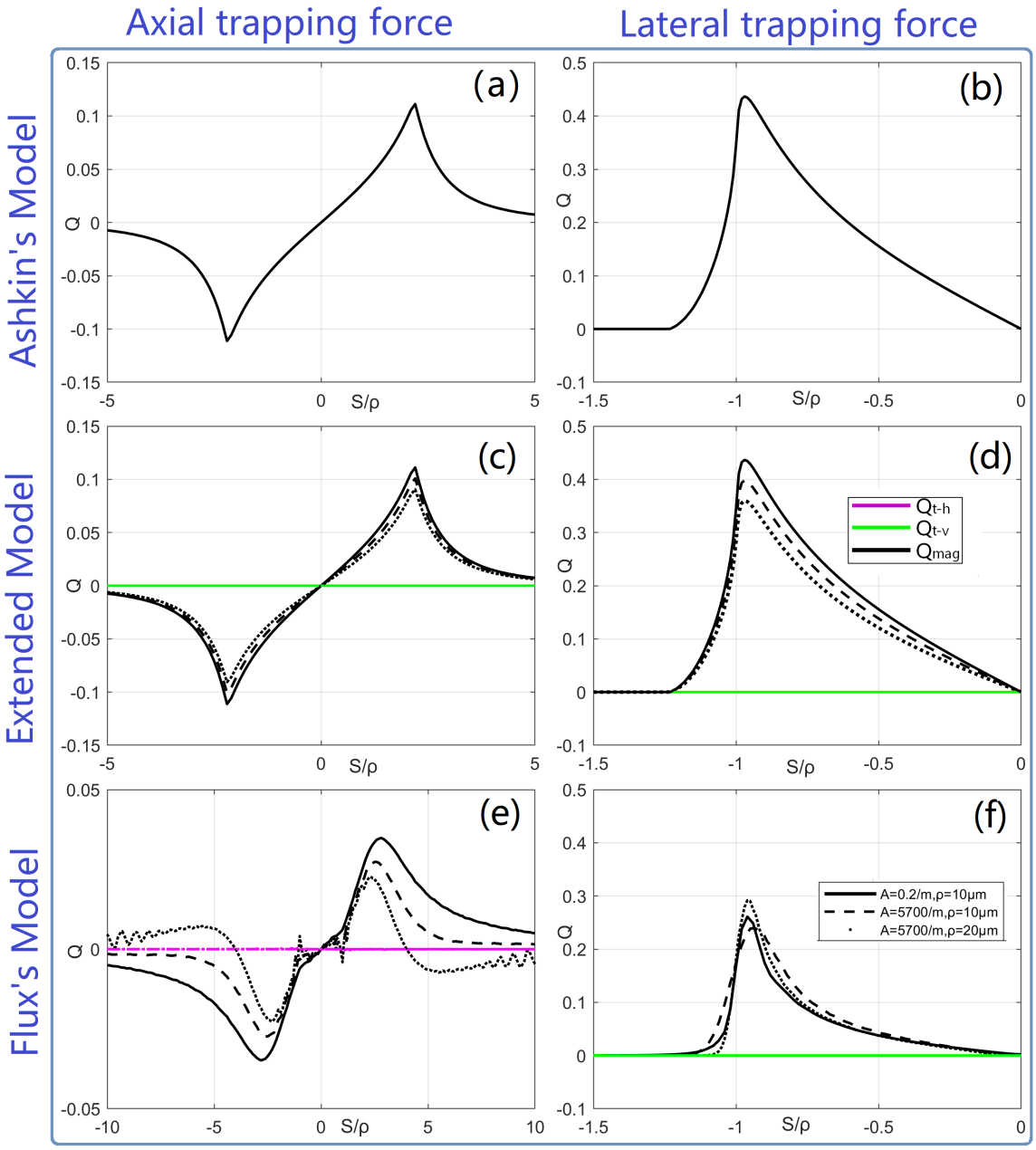

Figure 6.9: Prediction of the total axial force and the horizontal/vertical torques imparted by a dual beam optical trap from two opposed (ideal aberration free) lenses with NA 0.6 and TEM_{00} laser mode with $1/e^2$ cut-off passing through the back apertures of both lenses, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts slightly different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. The total force is given in black. Vertical torque is plotted in green and horizontal torque is given in pink. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

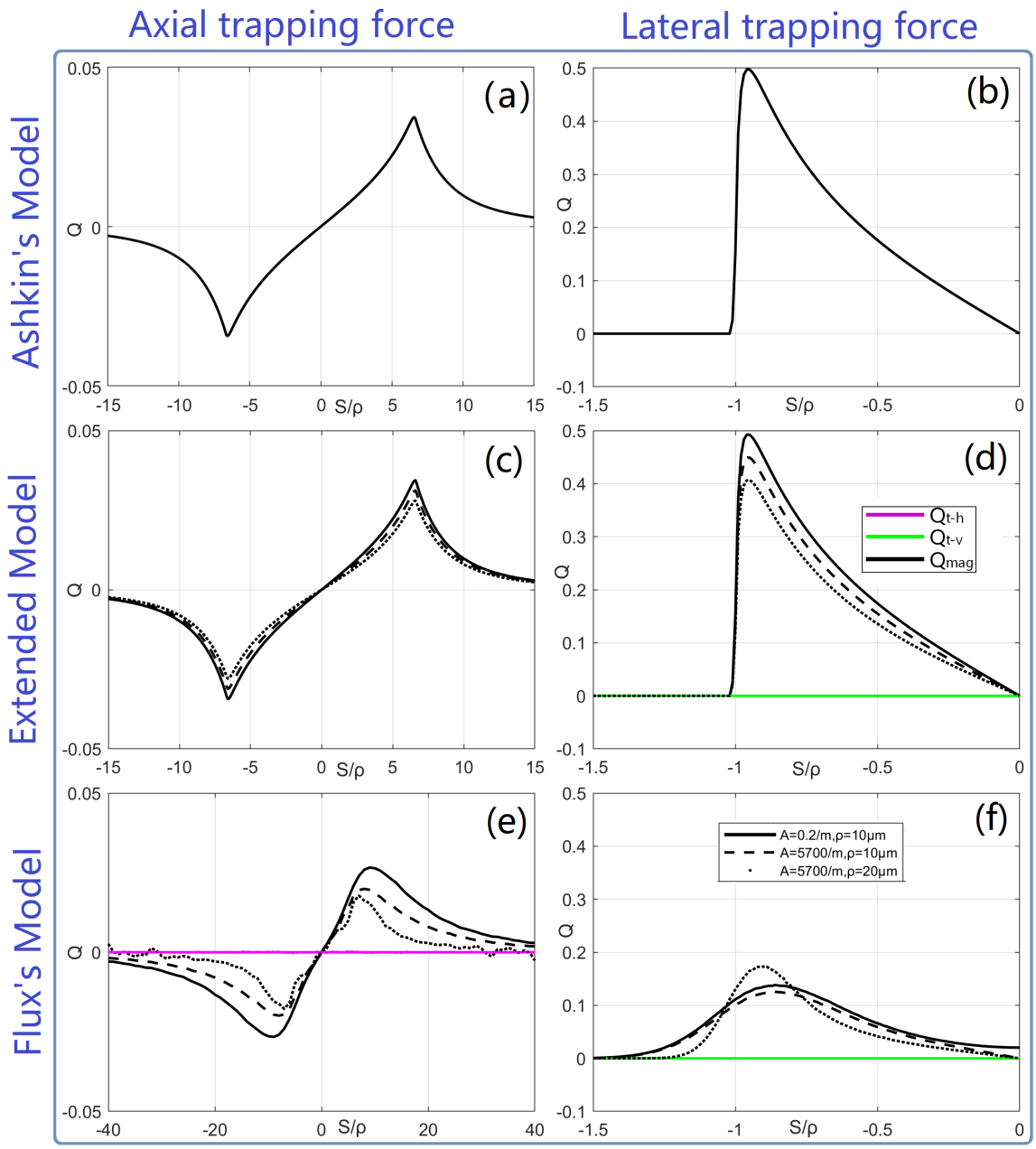

Figure 6.10: Prediction of the total axial force and the horizontal/vertical torques imparted by a dual beam optical trap from two opposed (ideal aberration free) lenses with NA 0.2 and TEM_{00} laser mode with $1/e^2$ cut-off passing through the back apertures of both lenses, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts slightly different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. The total force is given in black. Vertical torque is plotted in green and horizontal torque is given in pink. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

A similar set of results is presented in Fig. 6.10. This time for the case of two opposing ideal lens with numerical aperture of 0.2. All other parameters remain the same including the use of TEM_{00} lasers. The first point of note is that all three models predict excellent axial trapping forces in (a), (c), and (e) with strong optical traps positioned at the focal point. A second point is that the overall shape of all of the forces with respect to displacement is very similar as for the 0.6NA case. Both Ahskin's model and the Extended model predict total axial forces that are approximately half of what there were for the 0.6NA case, while the flux model predicts similar values to the 0.6NA case. Conversely, both Ahskin's model and the Extended model predict slightly higher total lateral forces than for the 0.6NA case, while the flux model predicts lower values for the 0.6NA case. Once again, the vertical torque is seen to be zero.

6.5.2 Laguerre-Gaussian Illumination

The same results are repeated in this section, where the TEM_{00} illumination is replaced with TEM_{01} illumination. The results for two opposing 0.6NA lenses are shown in Fig. 6.11 and for two opposing 0.2NA lenses in Fig. 6.12. The best axial force trap strength and stiffness across all of the results are found for the 0.6NA case in Fig. 6.11 (a), (c), and (e) and all three models predict similar force values. The lateral forces are also found to be similar in (b), (d), and (f). As for the TEM_{00} case all of the vertical torque has been cancelled. However, the horizontal torques have added together from both lasers resulting in the distribution shown in (e), which appears to change spin direction as a function of displacement. Although all of the forces are weaker for the corresponding cases in Fig. 6.12, the overall shape of the force/displacement distributions are very similar except for the lateral forces predicted by the flux models which appear to be much more spread out.

6.6 Conclusion

In this chapter, we provide a method that can estimate the amount of orbital angular momentum transferred to a particle from a Laguerre-Gussian trapping laser in the ray optics regime. We believe that this is the first time this has been attempted. The traditional approach for estimating the transfer of OAM is to argue that the total orbital angular momentum in a Laguerre Gaussian beam is given by hl (where h is Planks constant and l relates to the number of 2π phase shifts that occur in a rotation around the profile of the spatial mode) and that some portion of this is transferred to the sphere in an amount that is absorption related. However, such a vague description cannot take into account the numerical aperture of the lens. The flux-ray model presented in this chapter can achieve this and demonstrates that the spin that is imparted on the sphere is related the NA of the lens.

It is important to stress that we have not provided any experimental validation of the model and comprehensive testing is required. This would not be trivial. The simplest ways to experimentally evaluate trapping forces relate to lateral trapping forces

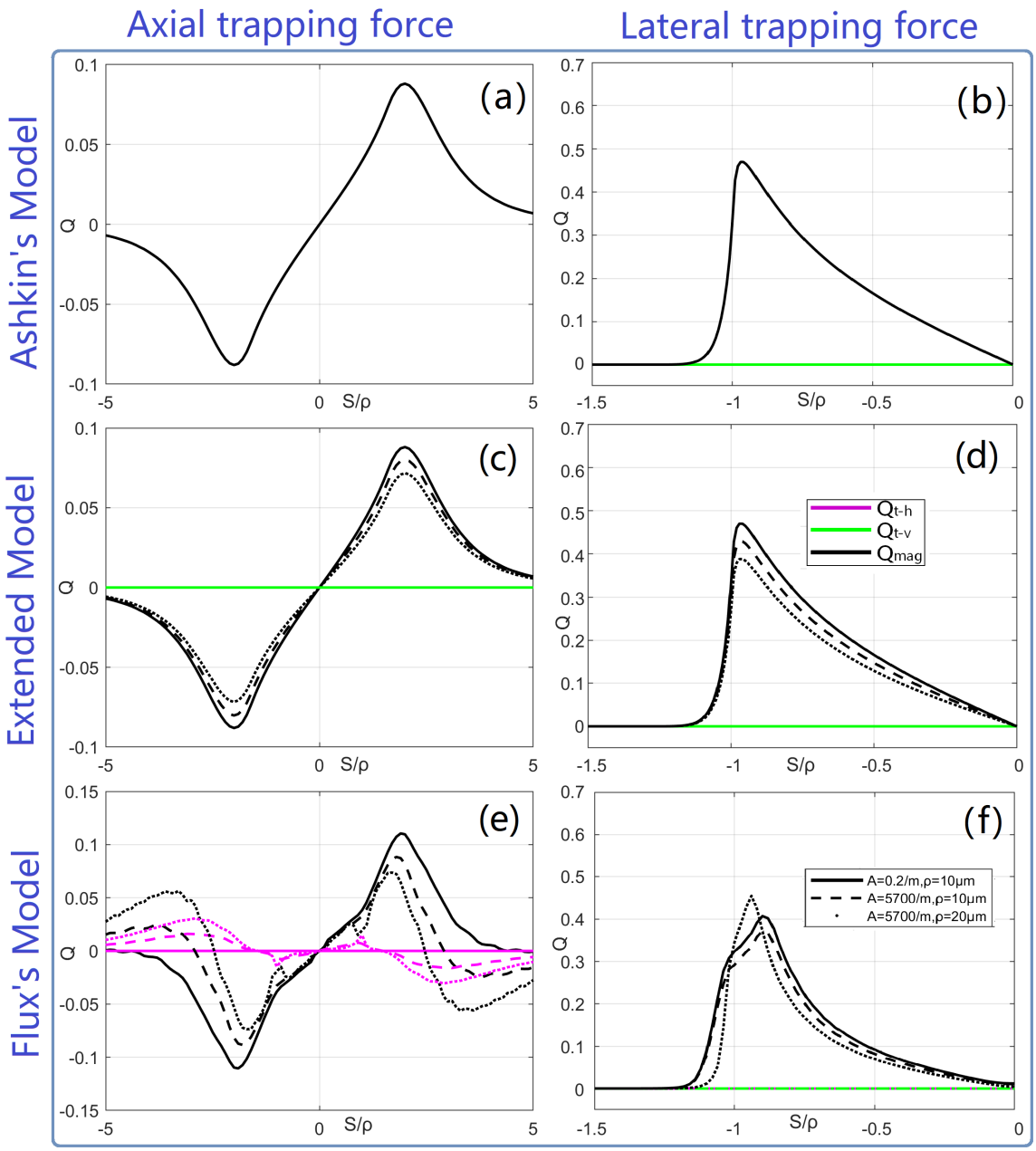

Figure 6.11: Prediction of the total axial force and the horizontal/vertical torques imparted by a dual beam optical trap from two opposed (ideal aberration free) lenses with NA 0.6 and TEM_{01} laser mode with $1/e^2$ cut-off passing through the back apertures of both lenses, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts slightly different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. The total force is given in black. Vertical torque is plotted in green and horizontal torque is given in pink. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

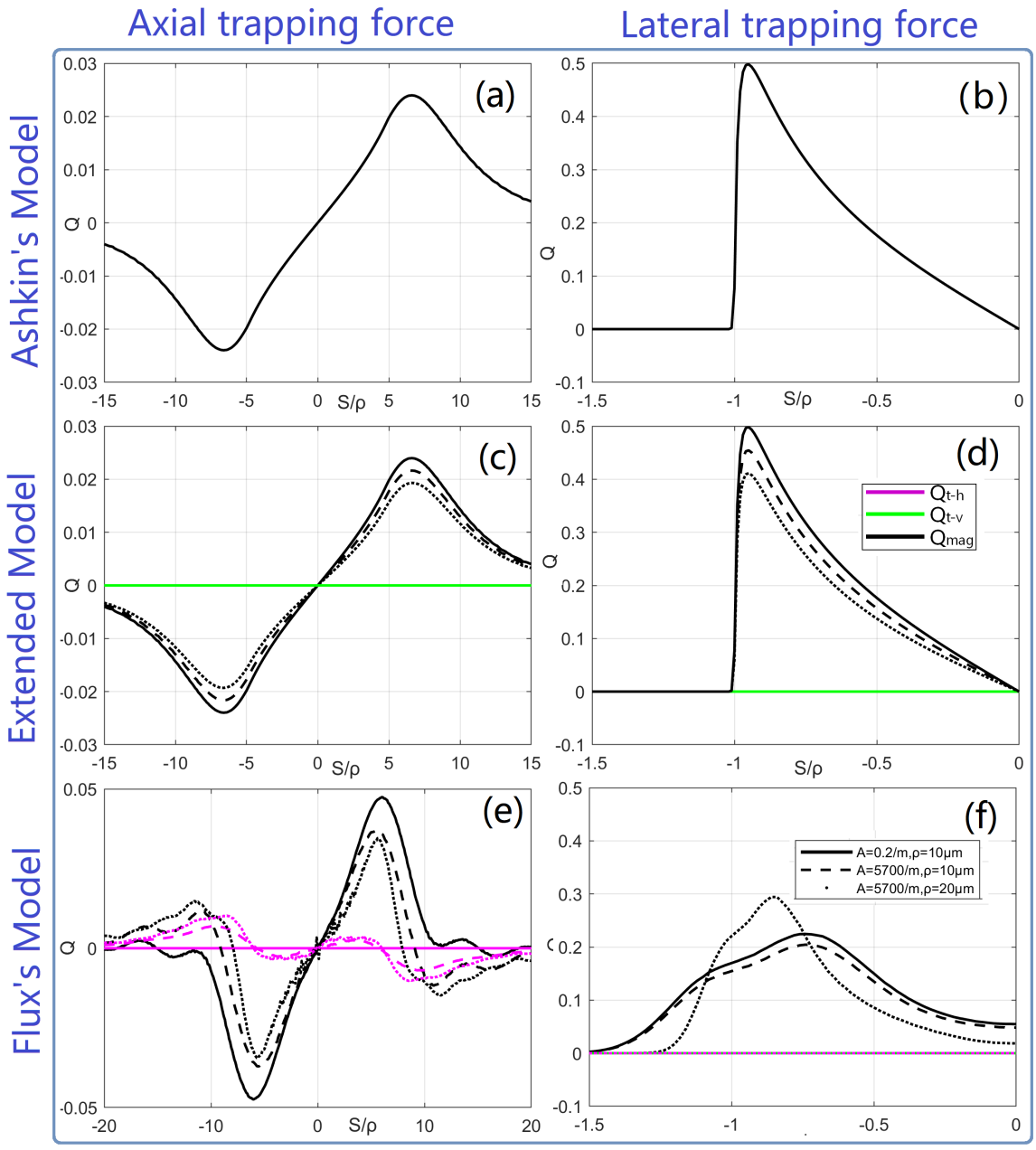

Figure 6.12: Prediction of the total axial force and the horizontal/vertical torques imparted by a dual beam optical trap from two opposed (ideal aberration free) lenses with NA 0.2 and TEM_{01} laser mode with $1/e^2$ cut-off passing through the back apertures of both lenses, acting on three different spheres of various size and absorption modelled using: (a),(b) Ashkin's original ray-optics model which predicts the same results for all three spheres; (c),(d) the extended ray-optics model developed in Chapter 3 which predicts slightly different results for all three cases, including a vertical torque (about the $y = 0$ axis) that emerges with lateral displacement; (e)-(f) the flux-ray model developed in this chapter. The total force is given in black. Vertical torque is plotted in green and horizontal torque is given in pink. The three different spheres are identified using different dotted or dashed lines as defined in the legend.

whereby a particle is placed in a micro-fluidic channel and is subject to the force associated with a controllable flow. The displacement of the particle (as a function of flow rate) can be measured either using a camera or more accurately using a position sensor based on four photo-diodes each forming a quadrant in a circular detector. To validate the flux-ray model we would need to conduct a flow experiment to see if the shape of the lateral force/displacement is more similar to the results for this model compared with Ashkin's results and this would have to be done for a dual beam or mirror trap since the model can only predict forces for NA < 0.7. Ideally, the vertical and rotational forces could be observed as a function of axial or lateral displacements.

Chapter 7

Conclusion and Future Work

In Chapter 3, we proposed an extension of Ashkin's model for deriving the the forces imparted on a dielectric sphere by as single ray. The only conceptual change that we introduce is the inclusion of a parameter to account for partial absorption of the rays as it passes through the particle as a function of the materials absorption coefficient. Based on this extended model for a single ray, we calculate the axial and lateral trapping forces for a number of different spheres with various radius and absorption coefficients. The extended model predicts that while the trap strength and stiffness are only slightly changed, the trapping position will change significantly as a function of absorption. This manifests from a significant increase in the scattering force for small axial displacements. The most striking result is the emergence of a strong rotational force which increases a function of lateral displacement and absorption. For the particle with low or negligible absorption coefficient, this rotational force reduces to zero and the extended model provides identical results to Ashkin. Although the rotational forces imparted by an optical spanner (an optical trap with a Laguerre-Gaussian (L-G) laser) has been experimentally observed and has been lined to absorption, Chapter 3 offers the first attempt to predict a rotational force caused from a single ray caused by refraction, and for a TEM_{00} spatial mode.

In Chapters 4 and 5 we deviate from the discussion of optical traps as we attempt to develop a computational method that can rapidly trace the flux lines through the laser focus. This will enable us to return to the equation of Chapter 3 and replace the rays with the flux lines. In Chapter 4 two algorithms are proposed for the numerical calculation of the three-dimensional complex amplitude of coherent wavefields that are focused by a lens. The first algorithm deals with the case of the thin lens approximation, while the second deals with an ideal lens system. In both cases, the angular spectrum method is applied in a loop in order to generate a stack of complex images over a regular grid around the focal region of the lens. However, the manner in which the ASM is applied differs considerably for both cases; for the thin lens, the ASM is applied directly to the lens transmission function but the difference in spatial support between the input plane (i.e the lens aperture) and the output plane (the focused spot) requires a 'pre-aliasing' step, recently proposed by other authors [13, 95] and further developed in this chapter, which although numerically intensive, needs only to be applied once before the ASM is applied iteratively; for the ideal lens case, a single dis-

crete Fourier transform is used to transform between the circular pupil function (in the Fourier domain) of the lens system to the focal plane of the lens, and then the ASM is applied iteratively to generate the stack of complex images immediately before and after this plane. A detailed list of steps is provided for both algorithms in Section 4.3 that includes the selection of the key lens properties as well as the optimal sampling conditions over the three spatial dimensions. Both algorithms were augmented to account for the laser spatial mode for the case that the input to the lens is a temporally coherent laser input. In particular, TEM_{00} (Gaussian) and TEM_{01} (Laguerre-Gaussian) were investigated. Furthermore, both algorithms were augmented to account for aberrations in the lens and for both cases the well known Zernike polynomials were used to described the various forms of lens aberration.

Regardless of which of the algorithms is employed from Chapter 4, the resulting 3D grid of samples serves as the staring point for the flux tracing method proposed in Chapter 5. The Eikonal function, can be used to relate the complex electromagnetic wavefront and the classical interpretation of geometrical ray optics. The Eikonal function can be used to define the flux lines, which trace the movement of the power over three dimensions as the wavefront propagates and can also be related to the derivative of the phase and the concept of a ray in the classical geometrical optics. Based on these concepts, we develop an algorithm in Chapter 5 that traces the flux through a sequence of different propagation distances separated by a value h_z; more specifically, the phase derivative of the 'ray' is calculated at each plane to direct a linear ray trace between adjacent plane. Using the concept of the Eikonal, we prove that the accuracy of this algorithm is high so long as the step size h_z is small. Having validated the method, the algorithm is applied to various cases of focusing different laser spatial modes and for lenses with various aberrations modelled using the Zernike polynomials. The most striking results relate to the Laguerre-Gaussian laser spatial modes and to the effect of lens aberrations. For the former, the rays appear to spiral rapidly with a spin rate depending on the order of the LG mode and the ray position relative to the optical axis. Those closest to the centre spin most rapidly, and there appear to be a linear dependence with the Laguerre-Gaussian order, which is consistent with the existing interpretation of orbital angular momentum. It is very interesting to note the movement of the rays close to the focus donut, with many rays undertaking a rapid change of direction, and forming a bubble like shape around the focus. The results for Coma, Trefoild, and Quadrafoil aberration provided insights that cannot be gleaned from classical geometrical. The rays could be traced moving into the characteristic lobes and spots surrounding the centre spot. The rays could even be traced into the corners and sides of the triangular centre spot for the case of Trefoil. These results clearly demonstrate the power of the approach as an extension of classical geometrical ray optics. We believe the method may find future usefulness in lens design and aberration theory.

Building on the results in Chapters 3, 4, and 5, in Chapter 6, we provide a method that can estimate the amount of orbital angular momentum transferred to a particle from a Laguerre-Gaussian trapping laser in the ray optics regime. We believe that this is the first time this has been attempted. The traditional approach for estimating the

transfer of OAM is to argue that the total orbital angular momentum in a Laguerre-Gaussian beam is given by hl (where h is Planks constant and l relates to the number of 2π phase shifts that occur in a rotation around the profile of the spatial mode) and that some portion of this is transferred to the sphere in an amount that is absorption related. However, such a vague description cannot take into account the numerical aperture of the lens. The flux-ray model presented in Chapter 6 can achieve this and demonstrates that the spin that is imparted on the sphere is related to the NA of the lens.

It is important to stress that we have not provided any experimental validation of the models presented in Chapter 3 and Chapter 6 and comprehensive testing is required. This would not be trivial but some ideas are put forth in Chapter 6 that might be able to achieve this. Even in the absence of experimental validation, we believe the methods developed in this thesis are valuable; in particular, the result in Chapter 3 that predicts rotational force as a function of absorption, with only a slight extension of Askins model and the algorithms developed in Chapters 4 and 5 for sampling the 3D focal volume of a lens and tracing the lines of flux must be considered to be useful tools; this is evidenced by the extraordinary results for the study of aberrations provided in Chapter 5.

Extension of the model to numerical aperture higher than 0.7 is not straightforward in Chapter6. One approach to avoid the use of the Angular Spectrum method in Algorith 2 for the ideal lens might be to apply a simple defocus chirp before applying the discrete Fourier transform and to move the z distance in this way. The argument for taking such an approach is that this method (displaying a defocus term) is used with spatial light modulators imaged to the back aperture of the MO, in order to move the laser focus in the z direction for dynamic holographic tweezers. This approach could yield results for NA up to 1.3 which would be easier to experimentally validate using a single-beam optical trap.

7.1 Proposed experimental set-up for validation

In this section, we describe a possible experimental set-up that may confirm the predictions presented in this paper, shown in Fig. 7.1 (a). A microscope comprising a lamp, a condenser lens, a microscope objective, a tube lens and a camera enables imaging of a spherical particle immersed in water in a chamber. The base of the chamber is a glass coverslip that is transmissive to the lamp and the laser wavelength while the top of the chamber is a dichroic mirror, which transmits the lamps but reflects the laser and facilitates an optical 'mirror trap' [35, 40, 109]. The Spatial Light Modulators (SLM) modulate the wavefront of the trapping laser, which is then demagnified by a two-lens imaging system before reflecting from dichroic beamsplitters into the back aperture of the MO. By displaying a phase pattern given by the angle of the complex sum of four quadratic (and linear) phase functions, the SLMs can produce different slightly diverging (and tilted) beams, which come to focus at different focal distances and positions in the region of the focal plane of the MO. The three-dimensional posi-

tions of these four focal points can be electronically controlled by varying the phase pattern on the SLM such that two beams come to focus before reflecting from the mirror while the other two come to focus after. In this way, a stable counter-propagating laser mirror trap can be generated by two of the beams, which are highlighted as foci 1 and 2 in Fig. 7.1 (b). The other two beams will be given a lateral displacement of $\pm d$ that is electronically controlled by adding a linear phase term to the quadratic phase corresponding to the given laser focus; there are highlighted as foci 3 and 4 in Fig. 7.1 (b).

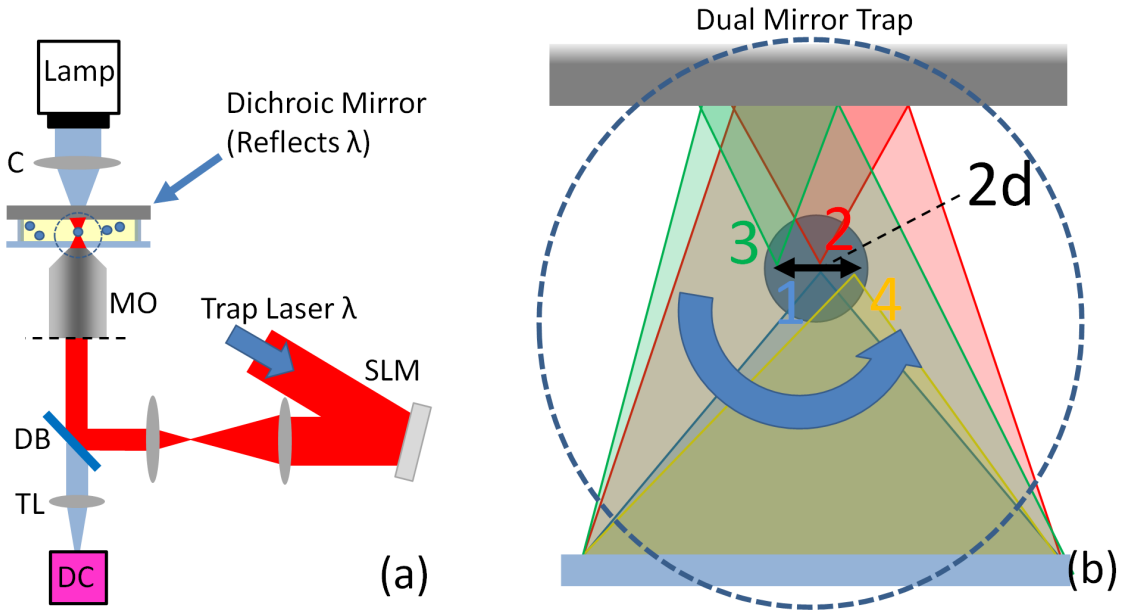

Figure 7.1: (a) Proposed set-up to demonstrate the theoretical prediction presented in this paper comprising a trapping laser with wavelength λ which will be partially absorbed by the trapped particle. Further description is provided in the text. Condenser Lens (C); Tube Lens (TL); Microscope Objective (MO); Digital Camera (DC); Dichroic Beamsplitter (DB); Spatial Light Modulator (SLM); (b) Illustration of the four laser foci generated by the SLM. Two mirror traps are used to trap and rotate the particle. Rotation is produced by the two foci from the laser that are displaced by distance d from the centre, while the other tow foci maintain a stable mirror trap.

7.2 Potential applications

One property of the proposed experimental set-up in the previous section is the capability to controllably rotate the trapped sphere (possessing the desirable properties of absorption) around the x or $y-$axis or any diagonal axis in the xy plane. In addition, by applying spiral phase patterns (associated with Laguerre-Gausian modes) to the SLM patterns relating to foci 1 and 2 in the propsoed experimental set-up, it is possible to rotate the particle around the z-axis making use of the transfer of orbital angular momentum. It may therefore be possible to manipulate the three dimensional position of a trapped particle as well as the rotation of the particle around any axis.

Vertical rotation of trapped particles has recently been employed for optical diffraction tomography [110] making use of two tweezers acting on an asymmetrically shaped biological specimen but such a approach would fail on a spherically shaped particle. However, application of absorption/refraction for particle rotation may be problematic in the field of biology due to the aforementioned heating effects.

The results in this paper may be of more interest to the growing area of research in employing novel materials as handles for the optical trapping and manipulation of tiny systems. This area of research focuses on unconventional materials that can offer characteristics significantly different from those obtained with the traditional dielectric polystyrene or silica beads. [111–116] One common feature in this body of work is the use of trapped materials with high absorption coefficients. In Ref. 112 the authors observe a direct relationship between the absorption coefficient and the radial force that pushes the particle toward the optical axis. However, for larger particles they attribute these forces refraction/reflections/transmission within the particle but no model has previously been offered to account for these forces. We believe that the results in this paper will contribute to the understanding and interpretation of the results observed in these papers and could potentially lead to applications that exploit the effect of absorption to induce controllable rotation of handles in optical trapping.

Chapter 8

Appendix

A Appendix A: Closed form solutions for Ashkin's scattering and gradient forces for a single ray

We begin with the two infinite summations reproduced here in Equations 1 and 2:

$$F_z = \frac{n_1 P}{c}\left[1 - \cos(\pi + 2\theta) - \sum_{n=0}^{\infty} T^2 R^n \cos(2(\theta - r) + n(\pi - 2r))\right] \tag{1}$$

$$F_y = \frac{n_1 P}{c}\left[0 - \sin(\pi + 2\theta) - \sum_{n=0}^{\infty} T^2 R^n \sin(2(\theta - r) + n(\pi - 2r))\right] \tag{2}$$

We apply the following variable substitutions to the summation appearing in Equation 1: $A = 2(\theta - r)$,and $B = \pi - 2r$, which enables the following simplification:

$$\begin{aligned}
\sum_{n=0}^{\infty} T^2 R^n \cos(A + nB) &= T^2 \sum_{n=0}^{\infty} R^n \Re(e^{i(A+nB)}) \\
&= T^2 \sum_{n=0}^{\infty} R^n \Re(e^{iA} e^{inB}) \\
&= T^2 \Re\left\{e^{iA} \sum_{n=0}^{\infty} R^n e^{inB}\right\} \\
&= T^2 \Re\left\{e^{iA} \frac{1}{[1 - Re^{iB}]}\right\} \\
&= T^2 \Re\left\{\frac{\cos A + i \sin A}{1 - R[\cos B + i \sin B]}\right\}
\end{aligned} \tag{3}$$

where $\Re\{\}$ denotes the real part. Multiplying above and below by the conjugate of the denominator provides further simplification:

$$\begin{aligned}
\sum_{n=0}^{\infty} T^2 R^n \cos(A + nB) &= T^2 \Re\left\{\frac{(\cos A + i \sin A)(1 - R\cos B + iR\sin B)}{(1 - R\cos B - iR\sin B)(1 - R\cos B + iR\sin B)}\right\} \\
&= T^2 \Re\left\{\frac{(\cos A(1 - R\cos B) - R\sin A \sin B + i \sin A(1 - R\cos B) + iR\sin B\cos A}{(1 - R\cos B)^2 + (R\sin B)^2}\right\}
\end{aligned} \tag{4}$$

Taking only the real part:

$$\begin{aligned}\sum_{n=0}^{\infty} T^2 R^n \cos(A+nB) &= T^2\left[\frac{\cos A(1-R\cos B)-R\sin A\sin B}{(1-R\cos B)^2+(R\sin B)^2}\right] \\ &= T^2\left[\frac{\cos A - R\cos(A-B)}{1+R^2-2R\cos B}\right]\end{aligned} \tag{5}$$

and finally, undoing the earlier variable substitution:

$$\begin{aligned}\sum_{n=0}^{\infty} T^2 R^n \cos(A+nB) &= T^2\left[\frac{\cos(2\theta-2r)-R\cos(2\theta-\pi)}{1+R^2-2R\cos(\pi-2r)}\right] \\ &= T^2\left[\frac{\cos(2\theta-2r)+R\cos(2\theta)}{1+R^2+2R\cos(2r)}\right]\end{aligned} \tag{6}$$

Therefore, Equation 1 can be rewritten as follows:

$$F_z = \frac{n_1 P}{c}\left[1+R\cos 2\theta - T^2\frac{\cos(2\theta-2r)+R\cos(2\theta)}{1+R^2+2R\cos(2r)}\right] \tag{7}$$

An identical approach can be taken for Equation 2 where in this case the imaginary part is taken in place of the real providing the simplification:

$$F_y = \frac{n_1 P}{c}\left[R\sin 2\theta - \frac{T^2\sin(2\theta-2r)+R\sin(2\theta)}{1+R^2+2R\cos 2r}\right] \tag{8}$$

B Appendix B: Derivation of sampling theorem for Angular Spectrum Method from first principles.

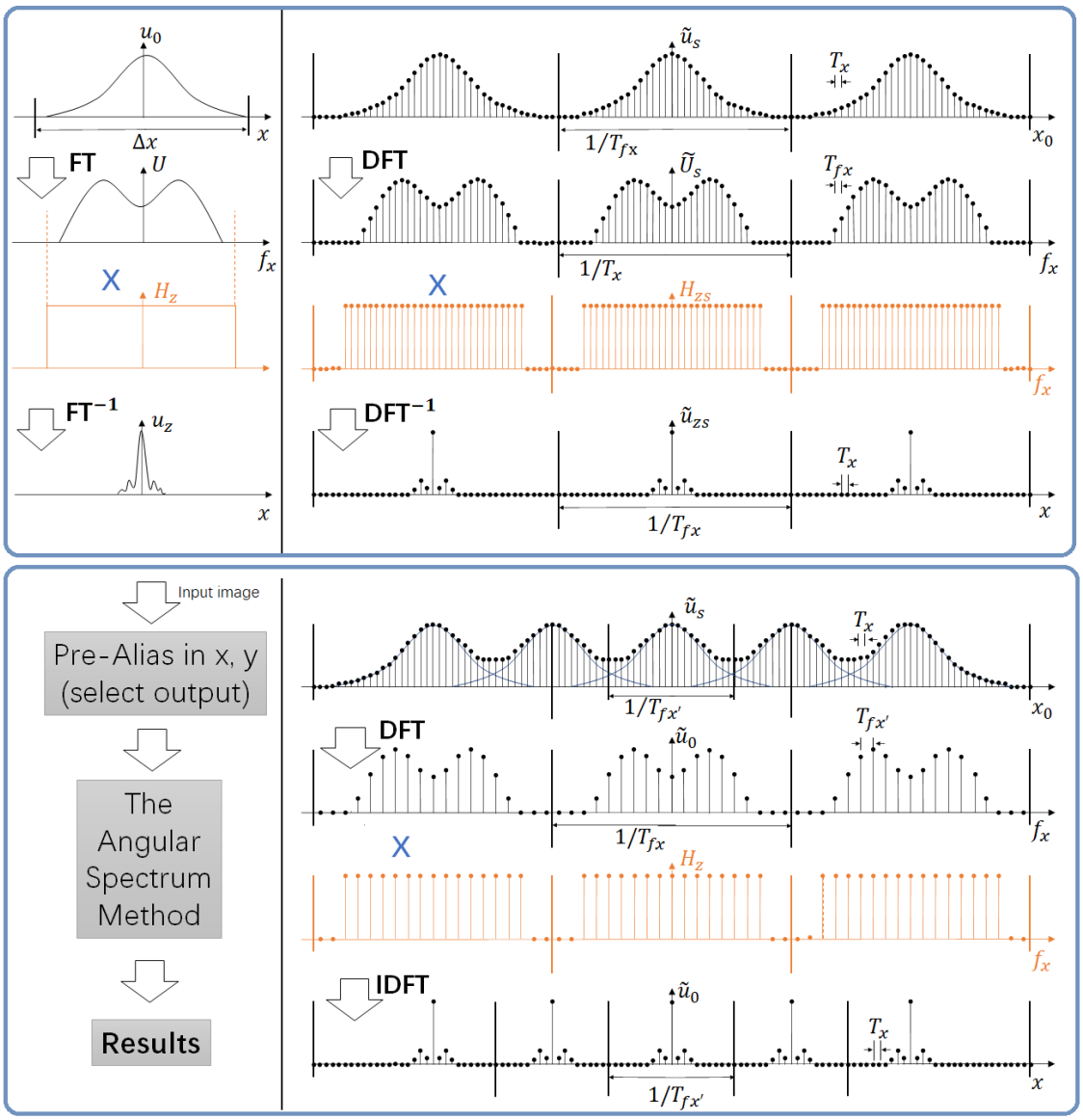

Figure 1: Illustration of the effects of sampling in both the space and spatial frequency domains.

We begin this section by defining some well known relationships. Convolution is defined as follows:

$$g(x) * h(x) = \int g(x')h(x-x')dx' \tag{9}$$

where the $*$ denotes the convolution operation between the two functions $g(x)$ and $h(x)$. The Dirac-delta functional is defined in terms if a limit as follow:

$$\delta(x) = \lim_{N\to\infty} N^2 \exp[-N^2\pi(x^2+y^2)] \tag{10}$$

which provides the following result:

$$g(x) = g(x) * \delta(x) \tag{11}$$

The '*comb*' function, $\delta_{T_x}(x)$, is an infinite train of Dirac-delta functionals uniformly separated by T_x and is defined as follows:

$$\delta_{T_x}(x) = \sum_{n=-\infty}^{+\infty} \delta(x - nT_x) \tag{12}$$

Where T_x represents the sampling interval between adjacent Dirac-delta functionals. The process of sampling a function $u(x)$ that is continuous in x can be described as a multiplication of $u(x)$ with the *comb* function. The resulting discrete function $u_s(x)$ is defined as follows:

$$u_s(x) = u(x)\delta_{T_x}(x) \tag{13}$$

The Fourier transform of $u(x)$, which is denoted by $U(f_x)$, is defined as follows:

$$\begin{aligned} U(f_x) &= \mathrm{FT}\{u(x)\}(f_x) \\ &= \int_{-\infty}^{+\infty} u(x) \exp(-2\pi f_x x)\, dx \end{aligned} \tag{14}$$

Here, $\mathrm{FT}\{u(x)\}(f_x)$ represents the Fourier Transform operator that is transforming the function from the space domain x into the spatial frequency domain f_x. The inverse Fourier transform performs the opposite operation as follows:

$$\begin{aligned} u(x) &= \mathrm{FT}^{-1}\{U(f_x)\}(x) \\ &= \int_{-\infty}^{+\infty} U(f_x) \exp(2\pi f_x x)\, df_x \end{aligned} \tag{15}$$

One of the core properties of the Fourier transform that is of interest here relates to convolution; the Fourier Transform of the convolution of the functions $g(x)$ and $h(x)$ is given by the product of their respective Fourier transforms as defined in the equations below:

$$\mathrm{FT}\left\{g(x) * h(x)\right\}(f_x) = \mathrm{FT}\left\{g(x)\right\}(f_x) \times \mathrm{FT}\left\{h(x)\right\}(f_x) \tag{16}$$

$$\mathrm{FT}\left\{g(x) \times h(x)\right\}(f_x) = \mathrm{FT}\left\{g(x)\right\}(f_x) * \mathrm{FT}\left\{h(x)\right\}(f_x) \tag{17}$$

The Fourier Transform of the *comb* function with sampling interval T_x is itself a *comb* function with sampling interval of $1/T_x$:

$$\mathrm{FT}\left\{\delta_{T_x}(x)\right\}(f_x) = \delta_{\frac{1}{T_x}}(f_x) \tag{18}$$

The Fourier transform of the discrete sampled function $u_s(x)$ can, therefore, be defined

in terms of a convolution:

$$\begin{aligned} U_s(f_x) &= \mathrm{FT}\{u_s(x)\}(f_x) \\ &= U(f_x) * \frac{1}{T_x}\delta_{\frac{1}{T_x}}(f_x) \\ &= \frac{1}{T_x}\sum_{n=-\infty}^{+\infty} U\left(f_x - \frac{n}{T_x}\right) \end{aligned} \tag{19}$$

This function, which is sometimes referred to as the 'Discrete Time Fourier Transform' (DTFT) is continuous and infinitely periodic in f_x. It can be described as the superposition of an infinite number of copies of $U(f_x)$, where each copy is shifted in f_x by an integer multiple of $1/T_x$. The Nyquist sampling condition can be derived from Equation 19. It states that the continuous function $u(x)$ can be recovered from $u_s(x)$ on the condition that (i) $U(f_x)$ has a finite support bounded within $\pm\Delta f_x/2$, and (ii) the sampling interval T_x is sufficiently small such that

$$\frac{1}{T_x} \geq \Delta f_x \tag{20}$$

The condition in Equation 20 guarantees that the shifted copies of $U(f_x)$ in the DTFT do not have overlapping support. In this case it is possible to recover $U(f_x)$ by multiplying $U_s(f_x)$ with a rect function:

$$U(f_x) = U_s(f_x) \times \mathrm{rect}\left(f_x T_x\right) \tag{21}$$

where

$$\begin{aligned} \mathrm{rect}(f_x T_x) &= 1 \forall\, |x| < \frac{1}{2T_x} \\ \mathrm{rect}(f_x T_x) &= 0 \forall\, |x| \geq \frac{1}{2T_x} \end{aligned} \tag{22}$$

The effect of multiplying by the rect function is to isolate a single period of $U_s(f_x)$, i.e. to isolate the copy for which $n = 0$ in Equation 19. The Fourier transform of a rect function is given by:

$$\mathrm{FT}\left\{\mathrm{rect}(f_x T_x)\right\} = T_x \mathrm{sinc}\left(\frac{x}{T_x}\right) \tag{23}$$

Therefore, by invoking the property of the Fourier transform given in Equation 16 we can derive Shannon's interpolation formula, which states that $u(x, y)$ can be recovered from its sampled values $u_x(x)$ so long as the Nyquist condition is met:

$$u(x) = u_s(x) * \mathrm{sinc}\left(\frac{x}{T_x}\right) \tag{24}$$

The condition known as 'aliasing' occurs when the Nyquist condition is not satisfied, and the copies of $U(f_x)$ overlap to some extent. In this case, a single copy of $U(f_x)$ can not be isolated and, therefore, $u(x)$ can not be perfectly recovered using Shannon's interpolation formula.

Sampling the DTFT in the frequency domain will lead to a function, $\widetilde{U}_s(f_x)$, that

is discrete and periodic in f_x and with a FT that is also discrete and periodic in x . Choosing a sampling interval of T_{fx} in f_x, this discrete function can once again be defined using a *comb* function:

$$\begin{aligned}\widetilde{U}_s(f_x) &= U_s(f_x) \times \delta_{T_{fx}}(f_x) \\ &= \left[U(f_x) * \frac{1}{T_x} \delta_{\frac{1}{T_x}}(f_x) \right] \times \delta_{T_{fx}}(f)\end{aligned} \tag{25}$$

The function $\widetilde{U}_s(f)$ is periodic in f_x due to the convolution of $U_s(f_x)$ with one *comb* function and it is also discrete in f_x as a result of multiplication with a second *comb* function. The inverse Fourier transform of this function, denoted by $\widetilde{u}_s(x)$ can be described as a convolution of $u_s(x)$ with a *comb* function as follows:

$$\begin{aligned}\widetilde{u}_s(x) &= u_s(x) * \frac{1}{T_{fx}} \delta_{\frac{1}{T_{fx}}}(x) \\ &= [u(x) \times \delta_{T_x}(x)] * \frac{1}{T_{fx}} \delta_{\frac{1}{T_{fx}}}(x)\end{aligned} \tag{26}$$

The effect of the sampling process in the frequency domain is to convolve $u_s(x)$ with a *comb* function of period $1/T_f$. The resulting function is, therefore, both discrete and infinitely periodic in x. In order to avoid aliasing in the space domain, the support of $u(x)$ must be smaller than the period $1/T_{fx}$, which is similar to the Nyquist condition:

$$\frac{1}{T_{fx}} \geq \Delta x \tag{27}$$

It is mathematically impossible for a function to be bounded in both x and f_x; [61] however, a function may be described as being approximately zero valued outside of finite supports in both x and f_x given by Δx and Δf_x, respectively. In such a case, it may be assumed that aliasing can be negligible in both domains if sufficiently small sampling intervals are applied.

The discrete functions $\widetilde{u}_s(x)$ and $\widetilde{U}_s(f_x)$ are related by the discrete Fourier transform (DFT), which is derived by substituting the discrete variables nT_x and mT_{fx} into Equation 14 in place of x and f_x:

$$\widetilde{U}_s(mT_{fx}) = T_x \sum_{n=-\infty}^{+\infty} \widetilde{u_s}(nT_x) \exp(-2\pi n T_x m T_{fx}) \tag{28}$$

where m and n are both integers taking the range from $-\infty$ to $+\infty$. Since $\widetilde{u}_s(x)$ is infinitely periodic in x with period $1/T_{fx}$, we can consider only the N samples in the integer range of n that make up a single period. Similarly, we can consider only N samples over the integer range of m. This implies a relationship between the sampling intervals in x and f_x as follows:

$$T'_{fx} = \frac{1}{NT_x} \tag{29}$$

Considering only N samples over a single period in both x and f_x allows for a simplification of Equation 28 as follows:

$$\widetilde{U}_s(mT_{fx}) = \mathrm{DFT}\{\widetilde{u_s}(nT_x)\} = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \widetilde{u_s}(nT_x) \exp\left(-\frac{2\pi nm}{N}\right) \tag{30}$$

where the range of m is given by $m = -N/2 \rightarrow N/2-1$. Equation 30 is known as the discrete Fourier transform (DFT). Similarly the inverse DFT is given by:

$$\widetilde{u}_s(nT_x) = \mathrm{DFT}^{-1}\left\{\widetilde{U}_s(mT_{fx})\right\} = \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} \widetilde{U}_s(mT_{fx}) \exp\left(\frac{2\pi nm}{N}\right) \tag{31}$$

The total number of operations required to directly calculate the DFT is N^2. However, the fast Fourier transform algorithm exploits redundancies in this calculation, and requires in the order of $N\log_2(N)$ calculations using a Radix 2 implementation (this requires N to be a power of 2).

Assuming that $u(x)$ and $U(f_x)$ are approximately zero valued outside of the supports Δx and Δf_x, respectively, and that the Nyquist sampling condition is true in both domains, i.e. $\Delta x \leq 1/T_{fx}$ and $\Delta f_x \leq 1/T_x$, then DFT is a mapping between two discrete and infinite functions, $\widetilde{u}_s$ and $\widetilde{U}_s$, which can be defined in terms of the original continuous function, u and its Fourier transform, U, as follows:

$$\widetilde{u}_s(nT_x) = u(nT_x) \forall n = -N/2 \rightarrow N/2-1 \tag{32}$$

$$\widetilde{U}_s(mT_{fx}) = U(mT_{fx}) \forall m = -N/2 \rightarrow N/2-1 \tag{33}$$

All of the analysis above can be easily generalized to two dimensions, x and y, by considering sampling separately for these two dimensions and the corresponding spatial frequency dimensions f_x and f_y. For the two-dimensional case the Nyquist condition would require an approximately finite support in all four dimensions: Δx, Δy, Δf_x, and Δf_y. Going forward, a two-dimensional analysis will be applied, which is necessitated by the angular spectrum method (ASM). In many cases the notation that was previously introduced will be extended to two-dimensions. Note that the 2-D $comb$ functions are separable in both dimensions, and the convolution operators can be considered separately over the two dimensions, as can all of the integrals/summations.

The Angular Spectrum Method (ASM) can be used to calculate the samples of a diffracted wavefield, $u_z(x,y)$, following propagation of the original wavefield $u(x,y)$ a distance z. We stipulate that the these two wavefields have support of $(\Delta x, \Delta y)$ and $(\Delta x_z, \Delta y_z)$ in space and both have spatial frequency support of $(\Delta f_x, \Delta f_y)$, which is a consequence of the properties of the propagation of the angular spectrum, which is the basis of the ASM. Propagation of the angular spectrum [62] is described as follows:

$$u_z(x) = \mathrm{FT}^{-1}\left\{\mathrm{FT}\{u_0(x)\}\, H_z(f)\right\}(f) \tag{34}$$

where,

$$H_z(f_x,f_y)=\begin{cases}\exp\left(j2\pi z\sqrt{\frac{1}{\lambda^2}-{f_x}^2-{f_y}^2}\right),\text{for}\frac{1}{\lambda^2}\geq {f_x}^2-{f_y}^2\\ \exp\left(-2\pi z\sqrt{{f_x}^2+{f_y}^2-\frac{1}{\lambda^2}}\right),\text{for}\frac{1}{\lambda^2}\leq {f_x}^2-{f_y}^2\end{cases} \tag{35}$$

Propagation of the angular spectrum is illustrated on the left side of Fig. 1, in which the continuous functions $u_0(x)$, $U_0(f_x)$, $H_z(f_x)$, and $u_z(x)$ are illustrated. In Equation 35, the transfer function, H_z has an infinite support in (f_x,f_y), which tends to zero at $1/\lambda^2={f_x}^2-{f_y}^2$. However, it should be noted that the support of H_z that needs to be considered in the propagation of the angular spectrum is limited by the support of the signal, i.e. $(\Delta f_x,\Delta f_y)$.

The ASM is based on discretisation of Equations 34 and 35 and is illustrated on the right side of Fig. 1. The ASM is implemented using the DFT in place of the Fourier Transform operator in Equation 34. Firstly $u(x)$ is sampled and the DFT is calculated as described in Equation 30. The transformation between the two discrete and infinitely periodic functions $\widetilde{u}_s(nT_x)$ and $\widetilde{U}_s(mT_f)$ is illustrated in Fig. 1. The latter function is then multiplied by an infinitely periodic and discrete version of H_z given by:

$$\begin{aligned}&\widetilde{H}_{zs}(m_xT_{fx},m_yT_{fy})=\\&\left[\left\{H_z(f_x,f_y)\times\text{rect}\left(\frac{f_x}{\Delta f_x},\frac{f_y}{\Delta f_y}\right)\right\}*\frac{1}{T_xT_y}\delta_{\frac{1}{T_x},\frac{1}{T_y}}(f_x,f_y)\right]\\&\times\delta_{T_{fx},T_{fy}}(f_x,f_y)\end{aligned} \tag{36}$$

where $\text{rect}\left(\frac{f_x}{\Delta f_x},\frac{f_y}{\Delta f_y}\right)$ is unity for $(f_x,f_y)<(\pm\frac{\Delta f_x}{2},\pm\frac{\Delta f_y}{2})$ and $(\Delta f_x,\Delta f_y)=(\frac{\Delta x}{2},\frac{\Delta y}{2})$. The 2D *comb* function can be defined in terms of the 1D equivalent as follows:

$$\delta_{T_{fx},T_{fy}}(f_x,f_y)=\delta_{T_{fx}}(f_x)\delta_{T_{fy}}(f_y) \tag{37}$$

Here, the rect function represents the bounding of H_z within the support of U, and the two $comb$ functions denote periodicity and sampling that is matched to that of $\widetilde{U}_s(mT_f)$. In terms of discrete samples, the centre period of Equation 38 can be rewritten as follows:

$$\begin{aligned}&\widetilde{H}_{zs}(m_xT_{fx},m_yT_{fy})=H_z(m_xT_{fx},m_xT_{fx})\\&\forall\, m_x=-N_x/2\rightarrow N_x/2-1\\&n_y=-N_y/2\rightarrow N_y/2-1\end{aligned} \tag{38}$$

The overall algorithm is, therefore, given by:

$$\widetilde{u}_{zs}(n_xT_x,n_yT_x)=\text{DFT}^{-1}\left\{\text{DFT}\left\{\widetilde{u}_s\right\}\times\widetilde{H}_{zs}\right\} \tag{39}$$

Note, that the periodicity of $\widetilde{u}_{zs}$ is identical to that of $\widetilde{u}_s$ and is given by $(\Delta x,\Delta y)=(N_xT_x,N_yT_y)=(1/T_{fx},1/T_{fy})$. However, as illustrated in Fig. 1, in many cases, the propagated field u_z has a significantly smaller support that that of u, which was given

by $(\Delta x_z, \Delta y_z) = (M_x T_x, M_y T_y) = (1/T'_{fx}, 1/T'_{fy})$, such that:

$$\begin{aligned} \Delta x_z &\ll \Delta x \\ \Delta y_z &\ll \Delta y \end{aligned} \tag{40}$$

Examples of this include the focusing of a plane wave by a lens, which is examined in some detail in this thesis. In such a case, there will clearly be a large number of zero valued samples in the calculated function $\widetilde{u}_{zs}$. We can conclude that a smaller number of samples M_x, M_y is required to fully describe u_z such that

$$\begin{aligned} M_x &\ll N_x \\ M_y &\ll N_y \end{aligned} \tag{41}$$

Although less samples are required to represent the continuous function, the sampling interval (T_x, T_y) remains suitable, since the inverse FT of u_z, denoted as U_z has the same support as that of U. If the support of u_z is given by the reduced values of $M_x T_x$ and $M_y T_y$, this implies a larger sampling interval is sufficient in the frequency domain in order to fully represent the signal, which is defined by rewriting Equation 29:

$$\begin{aligned} T'_{fx} &= \frac{1}{M_x T_x} \\ T'_{fy} &= \frac{1}{M_y T_y} \end{aligned} \tag{42}$$

Clearly the sampling intervals required to represent U_z, i.e. T'_{fx}, T'_{fy} will be much larger than the intervals required for U, i.e. T_{fx}, T_{fy}. If this larger set of intervals is applied to sample the DTFT, $U_s(f_x, f_y)$, given in Equation 19, the result is given by

$$\begin{aligned} \widetilde{U}_s(f_x, f_y) &= U_s(f_x, f_y) \times \delta_{T'_{fx}, T'_{fx}}(f_x, f_y) \\ &= \left[U(f_x, f_y) * \frac{1}{T_x T_y} \delta_{\frac{1}{T_x}, \frac{1}{T_y}}(f_x, f_y) \right] \times \delta_{T'_{fx}, T'_{fy}(f_x, f_y)} \end{aligned} \tag{43}$$

The inverse Fourier transform of this function, denoted by $\widetilde{u}'_s(x)$ is described as a convolution of $u_s(x)$ with a *comb* function as follows:

$$\begin{aligned} \widetilde{u}'_s(x, y) &= u_s(x, y) * \frac{1}{T'_{fx} T'_{fy}} \delta_{\frac{1}{T'_{fx}}, \frac{1}{T'_{fy}}}(x, y) \\ &= [u(x, y) \times \delta_{T_x, T_y}(x, y)] * \frac{1}{T'_{fx} T'_{fy}} \delta_{\frac{1}{T'_{fx}}, \frac{1}{T'_{fy}}}(x, y) \end{aligned} \tag{44}$$

$\widetilde{u}_s(x)$ and $\widetilde{u}'_s(x)$ defined in Equations 26 and 44 may look similar; however, there exists an important distinction. Whereas for the former, the Nyquist condition defined in

Equation 27 is satisfied, this is not so for the latter:

$$\begin{aligned} \frac{1}{T'_{fx}} &< \Delta x \\ \frac{1}{T'_{fy}} &< \Delta y \end{aligned} \tag{45}$$

In this case the period of the convolving *comb* function in Equation 44 is smaller than the support of u and the result is, therefore, aliasing. The DFT equations are rewritten as follows:

$$\begin{aligned} \widetilde{U}_s(m_x T_{fx}, m_y T_{fy}) &= \mathrm{DFT}\left\{\widetilde{u'_s}(n_x T_x, n_y T_y)\right\} \\ &= \sum_{n_x=-\frac{M_x}{2}}^{\frac{M_x}{2}-1} \sum_{n_y=-\frac{M_y}{2}}^{\frac{M_y}{2}-1} \widetilde{u'_s}(n_x T_x, n_y T_y) \\ &\quad \times \exp\left(-\frac{2\pi n_x m_x}{M_x}\right) \exp\left(-\frac{2\pi n_y m_y}{M_y}\right) \end{aligned} \tag{46}$$

$$\begin{aligned} \widetilde{u}'_s(n_x T_x, n_y T_y) &= \mathrm{DFT}^{-1}\left\{\widetilde{U}_s(m_x T_{fy}, m_x T_{fy})\right\} \\ &= \sum_{m_x=-\frac{M_x}{2}}^{\frac{M_x}{2}-1} \sum_{m_y=-\frac{M_y}{2}}^{\frac{M_y}{2}-1} \widetilde{U}_s(m_x T_{fx}, m_y T_{fy}) \\ &\quad \times \exp\left(\frac{2\pi n_x m_x}{M_x}\right) \exp\left(\frac{2\pi n_y m_y}{M_y}\right) \end{aligned} \tag{47}$$

where the range of n_x and m_x is given by $n_x, m_x = -M_x/2 \rightarrow M_x/2-1$ and the range of n_y and m_y is given by $n_y, m_y = -M_y/2 \rightarrow M_y/2-1$. In this case the DFT is a mapping between two two-dimensional discrete and infinite functions, $\widetilde{u}'_s$ and $\widetilde{U}_s$, which can be defined in terms of the original continuous function, u and its Fourier transform, U, as follows:

$$\begin{aligned} \widetilde{u}'_s(n_x T_x, n_y T_y) = \sum_{k_x=-\infty}^{+\infty} \sum_{k_y=-\infty}^{+\infty} & u_x T_x + k_x M_x T_x, n_y T_y + k_y M_y T_y) \\ & \forall n_x = -M_x/2 \rightarrow M_x/2-1 \\ & n_y = -M_y/2 \rightarrow M_y/2-1 \end{aligned} \tag{48}$$

$$\begin{aligned} \widetilde{U}_s(m_x T_{fx}, m_y T_{fy}) = & U(m_x T_{fx}, m_y T_{fy}) \\ & \forall m_x = -M_x/2 \rightarrow M_x/2-1 \\ & m_y = -M_y/2 \rightarrow M_y/2-1 \end{aligned} \tag{49}$$

The complexity of the two dimensional FFT algorithm is in the order of $M_x M_y \log_2(M_x) \log_2(M_y)$ calculations using a Radix 2 implementation. The effect of aliasing is described by the infinite summation given in Equation 48. This infinite summation assumes that the support of u is infinite in both x and y, and, therefore, the values of each of the infinite copies must be included in the summation. This is not so since we have already defined u to have finite support x and Δy. In such case, only a finite number of copies

may be included in the summation, i.e. the integers satisfied by:

$$\begin{aligned} |k_x| &< \frac{\Delta x}{2M_x T_x} + \frac{1}{2} \\ |k_y| &< \frac{\Delta y}{2M_y T_y} + \frac{1}{2} \end{aligned} \tag{50}$$

Even though the function $\widetilde{u}'_s$ is celarly aliased, it can serve as an input to the ASM and produce an accurate calculation of u_z such that the values of the sampled function $\widetilde{u}_{zs}(n_x T_x, n_y T_y)$ defined in Equation 39 can be related to the continuous function u_z:

$$\begin{aligned} \widetilde{u}_{zs}(n_x T_x, n_y T_y) &= u_z(n_x T_x, n_y T_y) \\ \forall\, n_x &= -M_x/2 \rightarrow M_x/2 - 1 \\ n_y &= -M_y/2 \rightarrow M_y/2 - 1 \end{aligned} \tag{51}$$

C Appendix C: Results for the case of an ideal lens using Algorithm 2

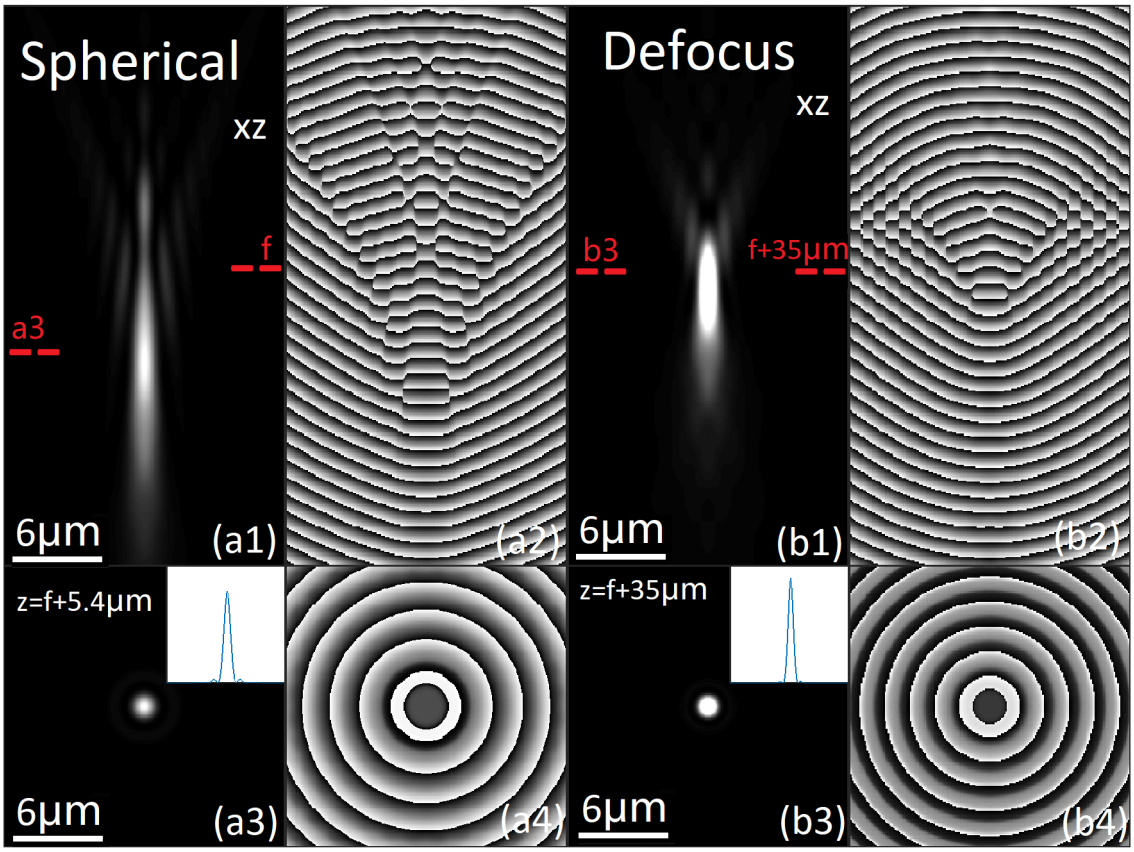

Figure 3: Intensity and phase distribution of focused spot for an ideal lens defined in Equation 4.20 and Equation 4.32 for the cases of spherical ($w_{11} = 2 \times 10^{-7}$) and defocus aberrations ($w_4 = -2 \times 10^{-6}$) with a Gaussian laser, calculated using Algorithm 2. (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$. Note that the focal plane is indicated by the f symbol in the image and the brightest spot occurs at a different distance, $z = f + 5.4\mu$m. The intensity and phase of the XY plane at this distance are shown in (a3) and (a4). A similar set of results are shown for defocus aberration in (b1)-(b4). Note that for the case of the XZ plane shown in (b1) and (b2) the bright spot occurs at a large distance from the focus, albeit at a different distance $z = f + 35\mu$m.

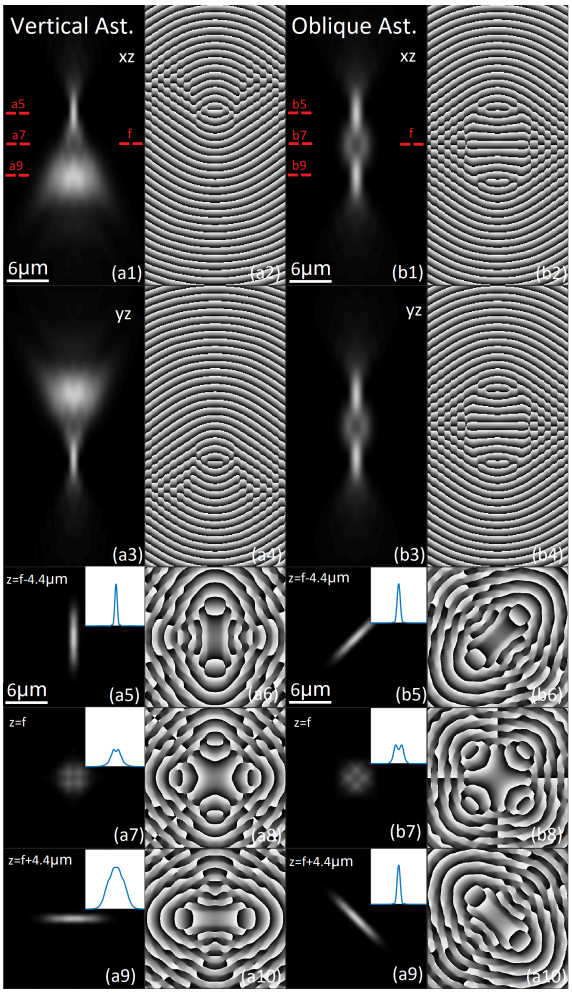

Figure 4: Intensity and phase distribution of focused spot for an ideal lens defined in Equation 4.20 and Equation 4.32 with primary astigmatism and with a Gaussian laser, calculated using Algorithm 2. The case of vertical astigmatism is shown in figures (a1-a10) for $w_6 = 4 \times 10^{-7}$; (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$; the brightest spot appears at distance $f - 4.4\mu$m and the intensity and phase distribution at this distance in the XY plane is shown in (a5) and (a6); here, the light has converged along the X-axis but is distributed over a wider range along the Y-axis. (a3) and (a4) show the intensity and phase distributions of the YZ slice for $y = 0$; now the brightest spot appears at distance $f + 4.4\mu$m and the intensity and phase distribution at this distance in the XY plane is shown in (a9) and (a10); here, the light has converged along the Y-axis but is distributed over a wider range along the X-axis and the phase distribution is similar the pattern in (a6) rotated by $\pi/2$. The intensity and phase distributions at the focal plane are shown in (a7) and (a8). The same results are shown for the case of oblique astigmatism ($w_5 = 4 \times 10^{-7}$) in (b1-b10). Vertical and oblique astigmatism are related by a $\pi/4$ rotation of the 3D light volume; the intensity and phase distributions in the XZ (b1 and b2) and YZ (b3 and b4) plane are both identical and are symmetrical and the two focal points of light can easily be seen. The XY distributions shown for the three planes in (b5-b10) are equivalent to those shown for the vertical astigmatism case with $\pi/4$ rotation.

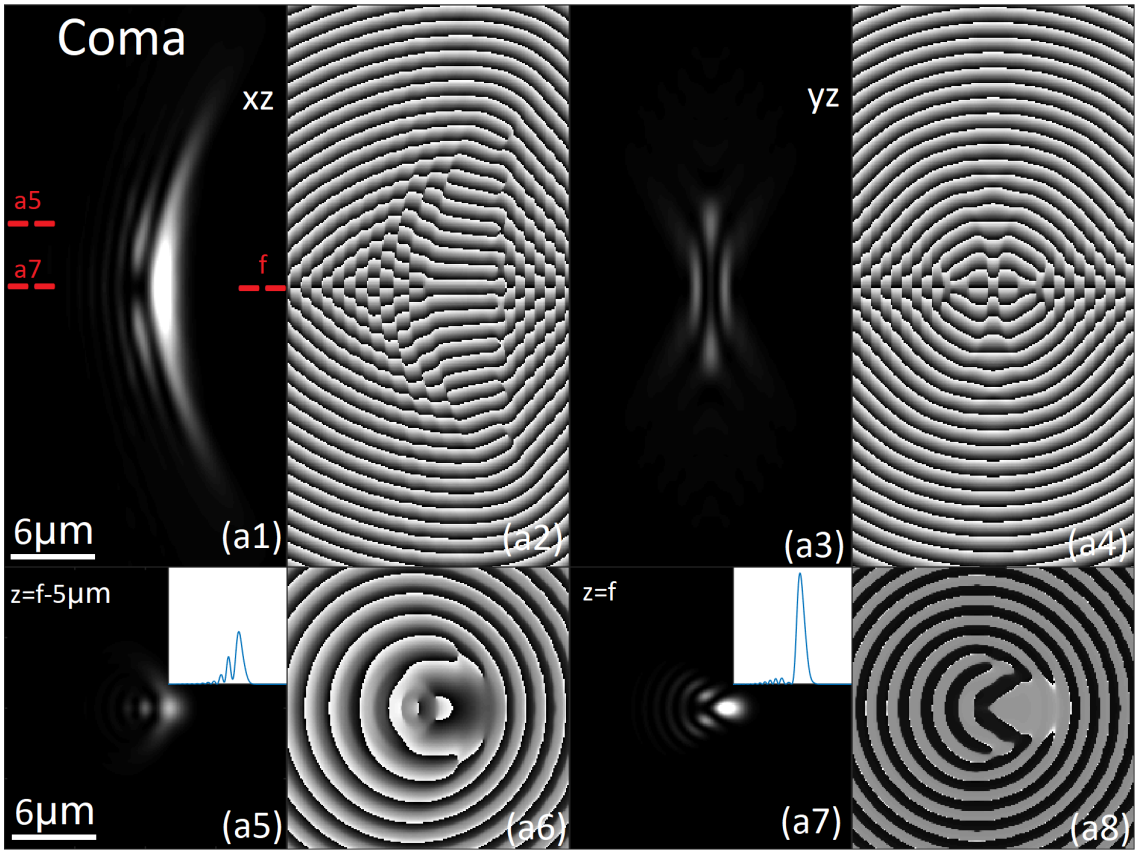

Figure 5: Intensity and phase distribution of focused spot for an ideal lens (defined in Equation 4.20 and Equation 4.32) with vertical coma ($w_7 = 3 \times 10^{-7}$), and with a Gaussian laser, calculated using Algorithm 2; (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$, and the focused light clearly has a curved shape in this plane. The XY intensity and phase distributions are shown for two different distances: at $z = f - 5\mu$m in (a5) and (a6) and at $z = f$ in (a7) and (a8) from which it can be seen the focused spot appears elongated along the X-axis and appears to move slightly as a function of distance z. Also shown is the intensity and phase distributions of the YZ slice for $x = 0$ in (a3) and (a4), which cuts through the curved spot.

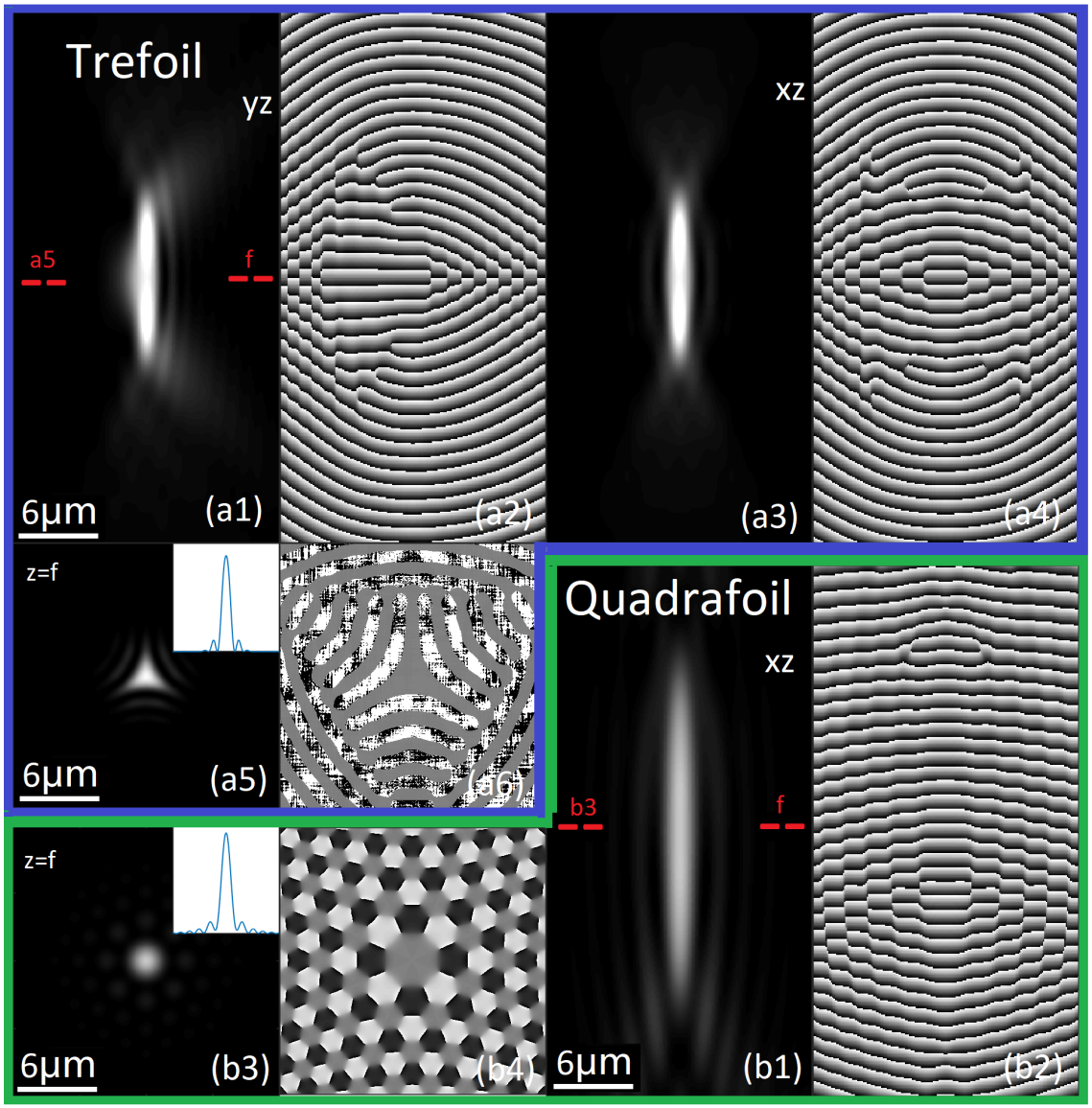

Figure 6: Intensity and phase distribution of focused spot for an ideal lens (defined in Equation 4.20 and Equation 4.32) with Gaussian laser illumination, and with vertical trefoil aberration ($w_9 = 3 \times 10^{-7}$) shown in (a1-a6) and vertical quadrafoil aberration ($w_{14} = 1 \times 10^{-6}$) shown in (b1-b4), calculated using Algorithm 2; for the trefoil case (a1) and (a2) show the intensity and phase distributions of the XZ slice for $y = 0$, and the focused light is symmetrical about the Z-axis in this plane; the intensity and phase distributions of the YZ slice for $x = 0$ is shown in (a3) and (a4), which is asymmetrical about the Z-axis. The intensity and phase distribution in the XY plane at the focal plane $z = f$ are shown in (a5) and (a6) in which the characteristic triangular shape of the focused light is seen. For the quadrafoil case (b1) and (22) show the intensity and phase distributions of the XZ slice for $y = 0$, and once again the focused light is symmetrical about the Z-axis but in this case there are two side lobes present around the focused spot, albeit weaker than for the case of the RS lens; the intensity and phase distribution in the XY plane at the focal plane $z = f$ are shown in (b3) and (b4) in which the focused spot is surrounded by eight smaller spots (weaker than for the case of the RS lens). It is the two spots immediately to the left and right that correspond to the two side lobes in the XZ-slice.

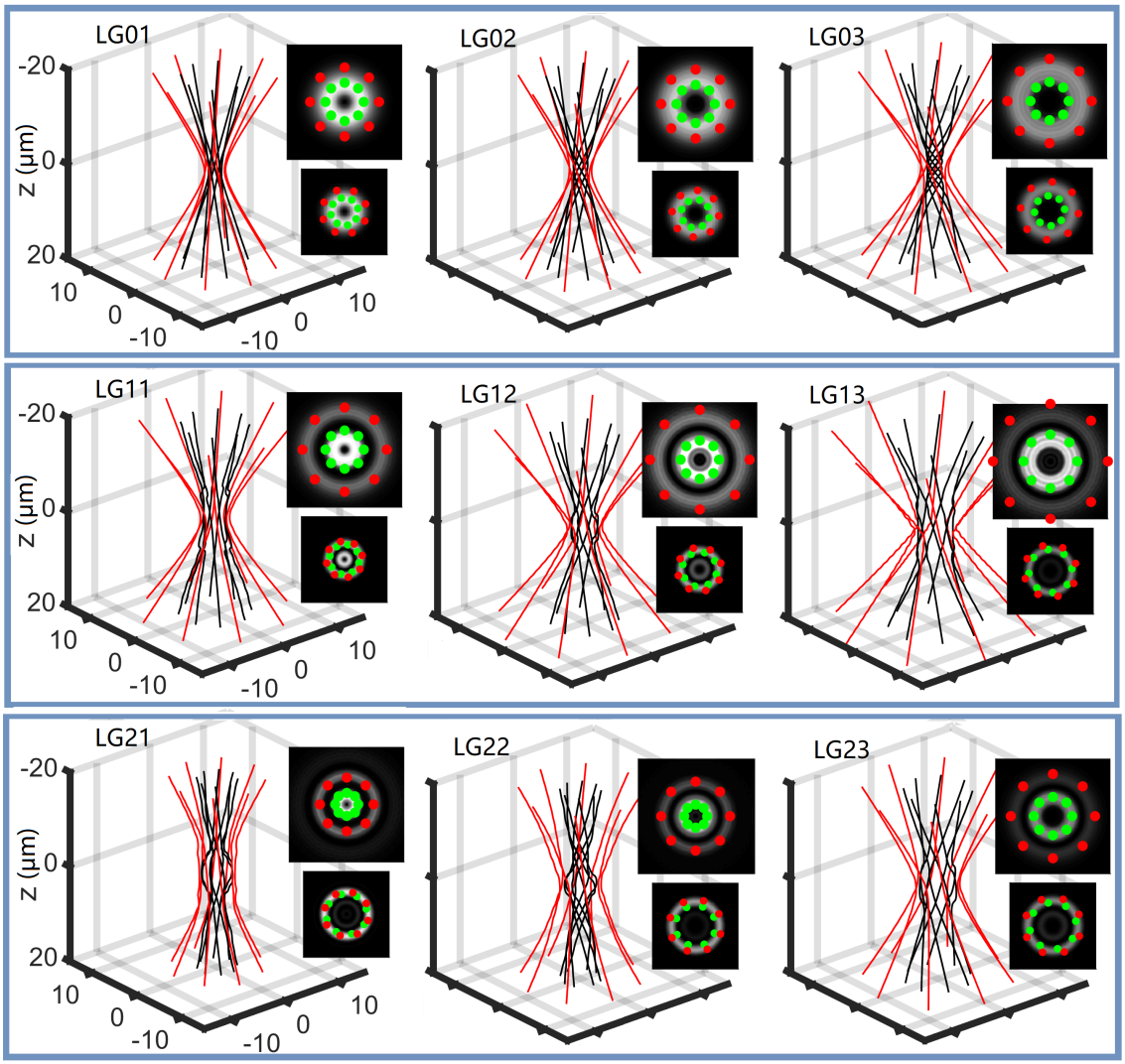

Figure 7: Nonlinear ray tracing for the different Laguerre-Gaussian modes focused using the same Ideal lens with numerical aperture of 0.6 investigated in Section 4 and Fig. 5 in Chapter 4: The values of l and p vary across rows and columns, respectively. Also shown in the figure are the intensity patterns at the focal plane and 20 µm above, which highlight the ray positions in these planes. It is interesting to note that the spin rate of the rays increases linearly as a function of l and also as a function of the position of the ray relative to the centre of the donut.

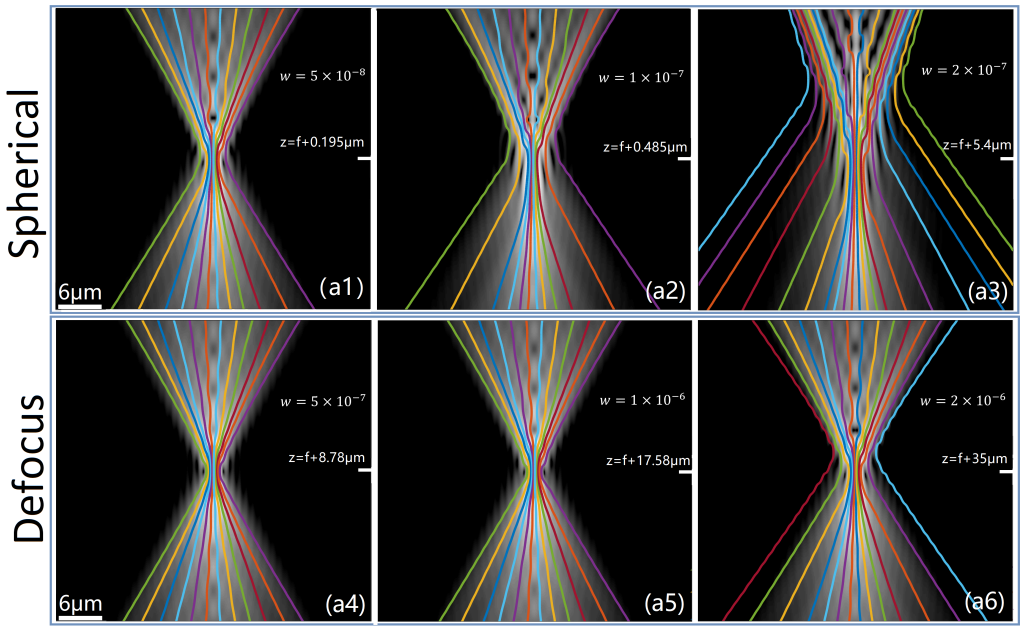

Figure 8: Non-linear ray tracing applied to different cases of spherical (a1)-(a3) and defocus (b1-b3) aberration on the Ideal lens with numerical aperture of 0.6 and TEM_{00} illumination computed using Algorithm 2. This simulation is based on the identical method and parameters described in Section 5 and Fig. 6 in Chapter 4. Three different weights, w, of spherical and defocus aberration are applied, shown in the figures. Note that the centre figures on both rows (a3) and (b3) corresponds to the specific results given Fig. 6. in Chapter 4. Note that both algorithms model aberrations in fundamentally different ways and yet both produce remarkably similar results.

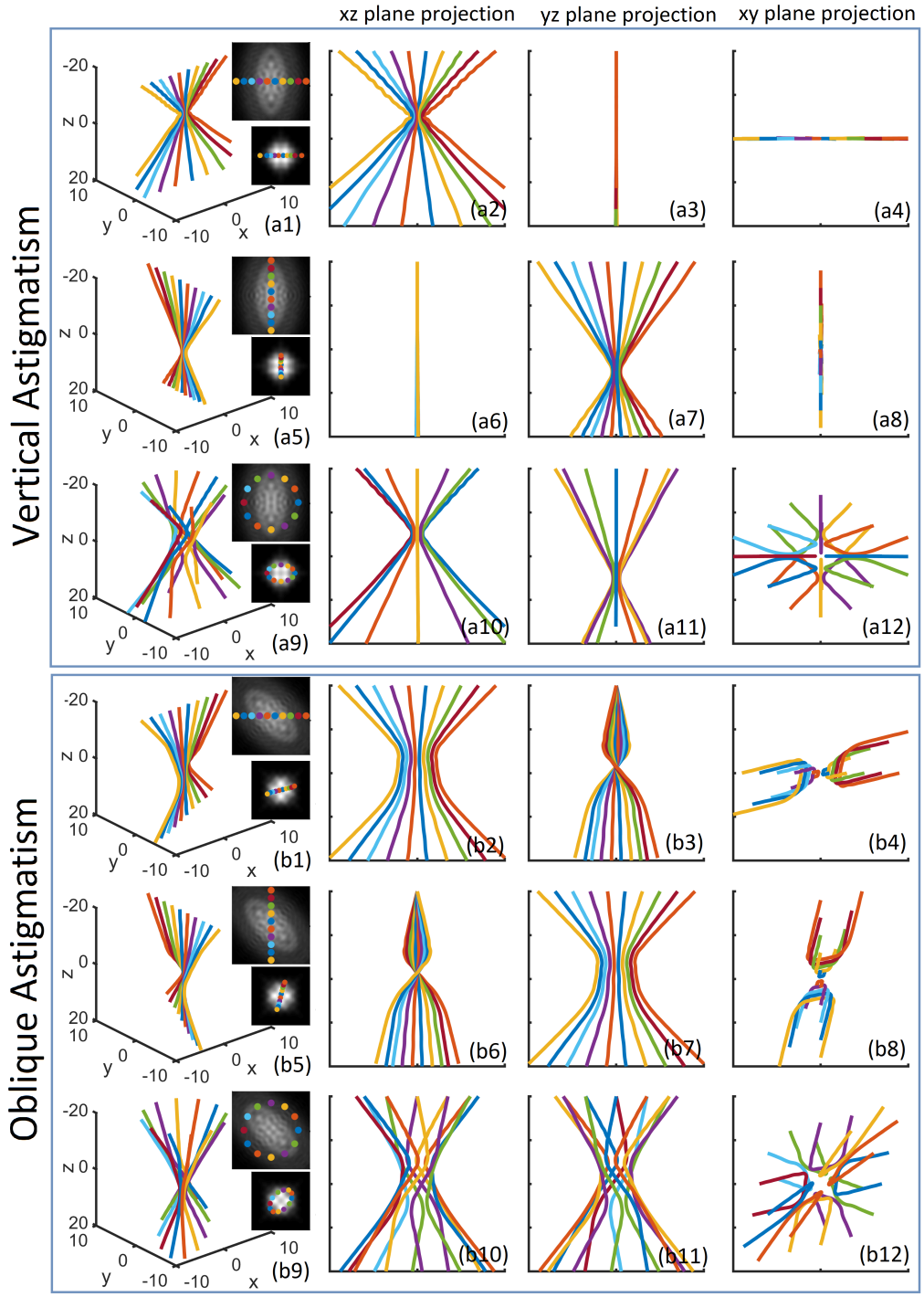

Figure 9: The results of non-linear ray tracing applied to vertical astigmatism aberration are presented in Fig. 9 (a1)-(a12), again for the case of the ideal lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 2. All of the parameters outlined in Chapter 4 are identical here (including the value of w used in the investigation of vertical astigmatism) except for the value of T_z. In the left most column a 3D plot is shown that traces various sets of 10 or 12 rays from $z = -20\mu m$ before the focal plane to $z = 20\mu m$ after, and also shows two inset (log of) intensity images that correspond to the intensity at the plane $z = -20\mu m$ before the focal plane, and at the focal plane. The larger inset image has dimensions of $20\mu m \times 20\mu m$ and the smaller inset image has dimensions of $10\mu m \times 10\mu m$. Similar results for horizontal astigmatism are presented in Fig. 9 (b1)-(b12)

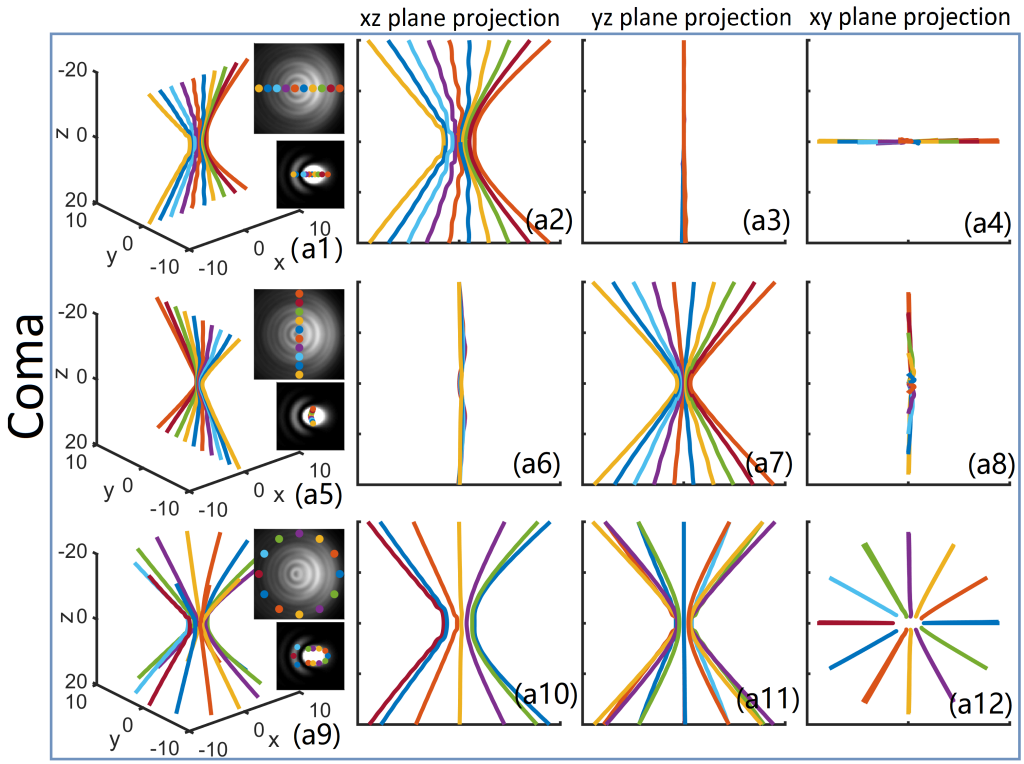

Figure 10: Non-linear ray tracing applied to coma aberration on the ideal lens with numerical aperture of 0.6 and TEM_{00} illumination computed using Algorithm 2 as described in Chapter 4. Various projections of different sets of 10 rays originating in the plane at $z = -20\mu m$: (a1)-(a4) ray origins uniformly distributed along the x-axis; (a5)-(a8) ray origins uniformly distributed along the y-axis; (a9)-(a12) ray origins uniformly distributed over a circle in xy plane. For all three cases two inset (log) intensity images are shown corresponding to the ray origin plane at $z = -20\mu m$ and the focal plane. The larger inset image has dimensions of $20\mu m \times 20\mu m$ and the smaller inset image has dimensions of $6\mu m \times 6\mu m$.

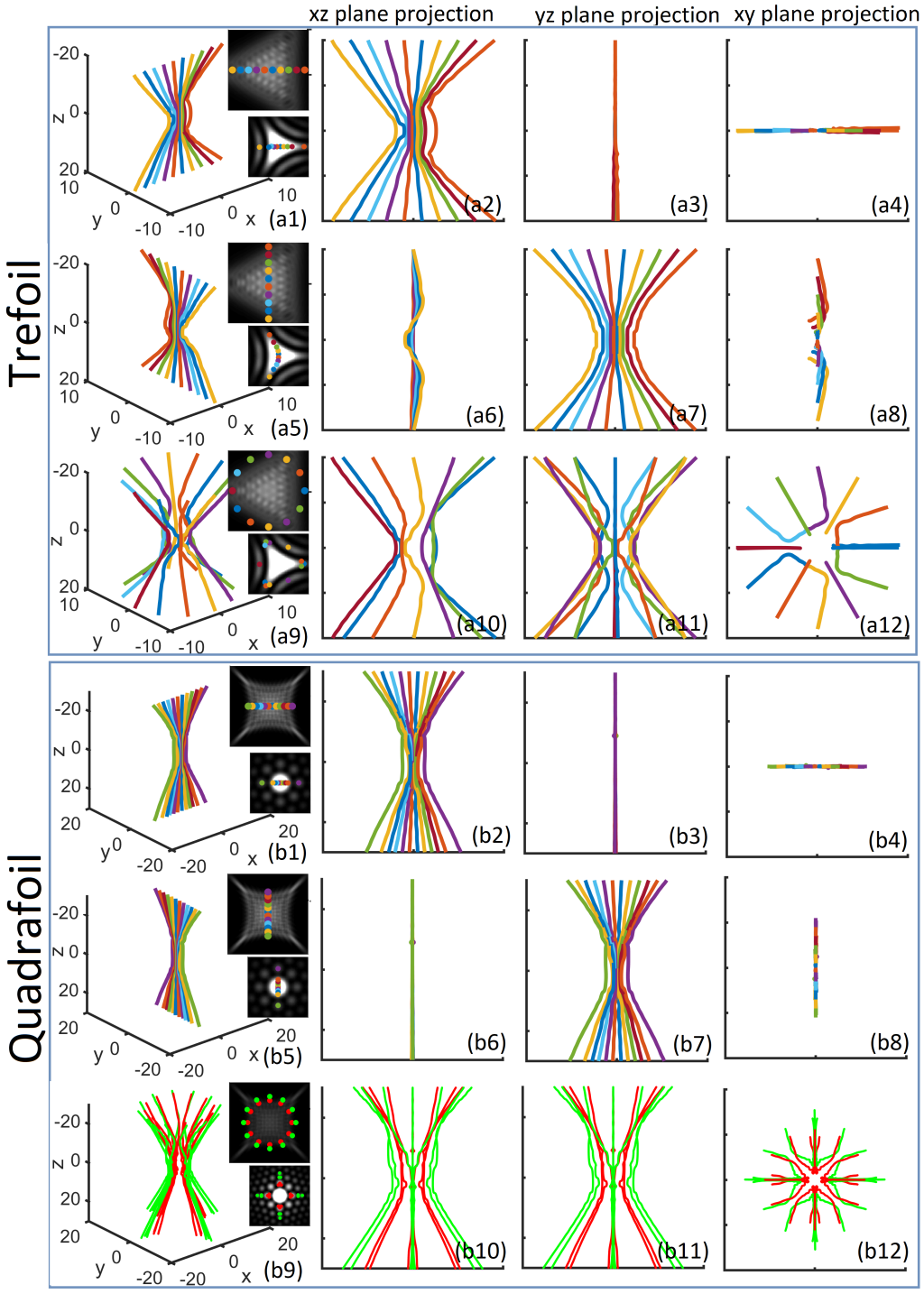

Figure 11: The results of non-linear ray tracing applied to Trefoil aberration are presented in Fig. 11 (a1)-(a12), for the case of the ideal lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 2. All of the parameters outlined in Chapter 4 are identical here (including the value of w used in the investigation for Trefoil) except for the value of T_z. The larger inset image has dimensions of 20μm×20μm and the smaller inset image has dimensions of 10μm×10μm. The results of non-linear ray tracing applied to Quadrafoil aberration are presented in Fig. 11 (b1)-(b12), again for the case of the ideal lens with numerical aperture of 0.6 and with TEM_{00} illumination computed using Algorithm 2. All of the parameters outlined in Chapter 4 are identical here except here we use $w = 5 \times 10^{-6}$, which was was found to produce excellent results for non-linear ray tracing

References

[1] J. C. Maxwell, "Viii. a dynamical theory of the electromagnetic field," *Philosophical transactions of the Royal Society of London*, no. 155, pp. 459–512, 1865.

[2] J. C. Maxwell, *A treatise on electricity and magnetism*, vol. 1. Oxford: Clarendon Press, 1873.

[3] E. F. Nichols and G. F. Hull, "The pressure due to radiation," in *Proceedings of the American Academy of Arts and Sciences*, vol. 38, pp. 559–599, JSTOR, 1903.

[4] T. H. Maiman *et al.*, "Stimulated optical radiation in ruby," 1960.

[5] A. Ashkin, "Acceleration and trapping of particles by radiation pressure," *Physical review letters*, vol. 24, no. 4, p. 156, 1970.

[6] A. Ashkin and J. Dziedzic, "Optical levitation of liquid drops by radiation pressure," *Science*, vol. 187, no. 4181, pp. 1073–1075, 1975.

[7] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, "Observation of a single-beam gradient force optical trap for dielectric particles," *Optics letters*, vol. 11, no. 5, pp. 288–290, 1986.

[8] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[9] J. Hecht, *Understanding fiber optics*. Jeff Hecht, 2015.

[10] J. C. Maxwell, "Xxv. on physical lines of force: Part i.–the theory of molecular vortices applied to magnetic phenomena," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 21, no. 139, pp. 161–175, 1861.

[11] P. Y. Ufimtsev, *Fundamentals of the physical theory of diffraction*. John Wiley & Sons, 2014.

[12] Y. Z. Umul, "Modified theory of physical optics," *Optics Express*, vol. 12, no. 20, pp. 4959–4972, 2004.

[13] M. Kanka, R. Riesenberg, and H. Kreuzer, "Reconstruction of high-resolution holographic microscopic images," *Optics letters*, vol. 34, no. 8, pp. 1162–1164, 2009.

[14] M. Hillenbrand, A. Hoffmann, D. P. Kelly, and S. Sinzinger, "Fast nonparaxial scalar focal field calculations," *JOSA A*, vol. 31, no. 6, pp. 1206–1214, 2014.

[15] S. Chu, L. Hollberg, J. E. Bjorkholm, A. Cable, and A. Ashkin, "Three-dimensional viscous confinement and cooling of atoms by resonance radiation pressure," *Physical review letters*, vol. 55, no. 1, p. 48, 1985.

[16] A. Ashkin and J. M. Dziedzic, "Optical trapping and manipulation of viruses and bacteria," *Science*, vol. 235, no. 4795, pp. 1517–1520, 1987.

[17] D. G. Grier, "Optical tweezers in colloid and interface science," *Current opinion in colloid & interface science*, vol. 2, no. 3, pp. 264–270, 1997.

[18] J. E. Curtis, B. A. Koss, and D. G. Grier, "Dynamic holographic optical tweezers," *Optics communications*, vol. 207, no. 1-6, pp. 169–175, 2002.

[19] D. G. Grier, "A revolution in optical manipulation," *nature*, vol. 424, no. 6950, pp. 810–816, 2003.

[20] D. G. Grier and Y. Roichman, "Holographic optical trapping," *Applied optics*, vol. 45, no. 5, pp. 880–887, 2006.

[21] A. Ashkin, "Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime," *Biophysical journal*, vol. 61, no. 2, pp. 569–582, 1992.

[22] K. C. Neuman and S. M. Block, "Optical trapping," *Review of scientific instruments*, vol. 75, no. 9, pp. 2787–2809, 2004.

[23] P. Zemánek, A. Jonáš, L. Šrámek, and M. Liška, "Optical trapping of rayleigh particles using a gaussian standing wave," *Optics communications*, vol. 151, no. 4-6, pp. 273–285, 1998.

[24] Y. Harada and T. Asakura, "Radiation forces on a dielectric sphere in the rayleigh scattering regime," *Optics communications*, vol. 124, no. 5-6, pp. 529–541, 1996.

[25] J. Wang and P. S. Lee, "Progress and prospects in stretchable electroluminescent devices," *Nanophotonics*, vol. 6, no. 2, pp. 435–451, 2017.

[26] P. J. Pauzauskie, A. Radenovic, E. Trepagnier, H. Shroff, P. Yang, and J. Liphardt, "Optical trapping and integration of semiconductor nanowire assemblies in water," *Nature materials*, vol. 5, no. 2, pp. 97–101, 2006.

[27] K. Ladavac and D. G. Grier, "Microoptomechanical pumps assembled and driven by holographic optical vortex arrays," *Optics express*, vol. 12, no. 6, pp. 1144–1149, 2004.

[28] J. W. Goodman, *Introduction to Fourier Optics*. Roberts & Company Publishers, 2004.

[29] G. Lazarev, A. Hermerschmidt, S. Krüger, and S. Osten, "Lcos spatial light modulators: trends and applications," *Optical Imaging and Metrology: Advanced Technologies*, pp. 1–29, 2012.

[30] T. A. Bartlett, W. C. McDonald, J. N. Hall, P. I. Oden, D. Doane, R. S. Ketchum, and T. Byrum, "Recent advances in the development of the texas instruments phase-only microelectromechanical systems (mems) spatial light modulator," *Emerging Digital Micromirror Device Based Systems and Applications XIII*, vol. 11698, pp. 103–116, 2021.

[31] Y.-X. Ren, R.-D. Lu, and L. Gong, "Tailoring light with a digital micromirror device," *Annalen der physik*, vol. 527, no. 7-8, pp. 447–470, 2015.

[32] M. E. Friese, T. A. Nieminen, N. R. Heckenberg, and H. Rubinsztein-Dunlop, "Optical alignment and spinning of laser-trapped microscopic particles," *Nature*, vol. 394, no. 6691, pp. 348–350, 1998.

[33] N. Simpson, D. McGloin, K. Dholakia, L. Allen, and M. Padgett, "Optical tweezers with increased axial trapping efficiency," *Journal of Modern Optics*, vol. 45, no. 9, pp. 1943–1949, 1998.

[34] M. Padgett and R. Bowman, "Tweezers with a twist," *Nature photonics*, vol. 5, no. 6, pp. 343–348, 2011.

[35] R. Bowman, A. Jesacher, G. Thalhammer, G. Gibson, M. Ritsch-Marte, and M. Padgett, "Position clamping in a holographic counterpropagating optical trap," *Optics express*, vol. 19, no. 10, pp. 9908–9914, 2011.

[36] G. Gibson, J. Courtial, M. J. Padgett, M. Vasnetsov, V. Pas'ko, S. M. Barnett, and S. Franke-Arnold, "Free-space information transfer using light beams carrying orbital angular momentum," *Optics express*, vol. 12, no. 22, pp. 5448–5456, 2004.

[37] A. Ashkin and J. M. Dziedzic, "Optical trapping and manipulation of viruses and bacteria," *Science*, vol. 235, no. 4795, pp. 1517–1520, 1987.

[38] S. M. Block, D. F. Blair, and H. C. Berg, "Compliance of bacterial flagella measured with optical tweezers," *Nature*, vol. 338, no. 6215, pp. 514–518, 1989.

[39] A. Ashkin, "History of optical trapping and manipulation of small-neutral particle, atoms, and molecules," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 6, pp. 841–856, 2000.

[40] M. Pitzek, R. Steiger, G. Thalhammer, S. Bernet, and M. Ritsch-Marte, "Optical mirror trap with a large field of view," *Optics express*, vol. 17, no. 22, pp. 19414–19423, 2009.

[41] G. Thalhammer, R. Steiger, M. Meinschad, M. Hill, S. Bernet, and M. Ritsch-Marte, "Combined acoustic and optical trapping," *Biomedical optics express*, vol. 2, no. 10, pp. 2859–2870, 2011.

[42] G. Thalhammer, R. Steiger, S. Bernet, and M. Ritsch-Marte, "Optical macro-tweezers: trapping of highly motile micro-organisms," *Journal of Optics*, vol. 13, no. 4, p. 044024, 2011.

[43] A. D. Mehta, M. Rief, J. A. Spudich, D. A. Smith, and R. M. Simmons, "Single-molecule biomechanics with optical methods," *Science*, vol. 283, no. 5408, pp. 1689–1695, 1999.

[44] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, "Stretching dna with optical tweezers," *Biophysical journal*, vol. 72, no. 3, pp. 1335–1346, 1997.

[45] T. G. Mason, K. Ganesan, J. H. van Zanten, D. Wirtz, and S. C. Kuo, "Particle tracking microrheology of complex fluids," *Physical review letters*, vol. 79, no. 17, p. 3282, 1997.

[46] J. Solon, I. Levental, K. Sengupta, P. C. Georges, and P. A. Janmey, "Fibroblast adaptation and stiffness matching to soft elastic substrates," *Biophysical journal*, vol. 93, no. 12, pp. 4453–4461, 2007.

[47] J. Guck, S. Schinkinger, B. Lincoln, F. Wottawah, S. Ebert, M. Romeyke, D. Lenz, H. M. Erickson, R. Ananthakrishnan, D. Mitchell, *et al.*, "Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence," *Biophysical journal*, vol. 88, no. 5, pp. 3689–3698, 2005.

[48] K. Berg-Sørensen and H. Flyvbjerg, "Power spectrum analysis for optical tweezers," *Review of Scientific Instruments*, vol. 75, no. 3, pp. 594–612, 2004.

[49] C. Bustamante, Z. Bryant, and S. B. Smith, "Ten years of tension: single-molecule dna mechanics," *Nature*, vol. 421, no. 6921, pp. 423–427, 2003.

[50] J. R. Moffitt, Y. R. Chemla, S. B. Smith, and C. Bustamante, "Recent advances in optical tweezers," *Annu. Rev. Biochem.*, vol. 77, pp. 205–228, 2008.

[51] W. J. Greenleaf, M. T. Woodside, and S. M. Block, "High-resolution, single-molecule measurements of biomolecular motion," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 36, pp. 171–190, 2007.

[52] S. B. Smith, Y. Cui, and C. Bustamante, "Overstretching b-dna: the elastic response of individual double-stranded and single-stranded dna molecules," *Science*, vol. 271, no. 5250, pp. 795–799, 1996.

[53] J. Levitz, C. Pantoja, B. Gaub, H. Janovjak, A. Reiner, A. Hoagland, D. Schoppik, B. Kane, P. Stawski, A. F. Schier, *et al.*, "Optical control of metabotropic glutamate receptors," *Nature neuroscience*, vol. 16, no. 4, pp. 507–516, 2013.

[54] K. C. Neuman and A. Nagy, "Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy," *Nature methods*, vol. 5, no. 6, pp. 491–505, 2008.

[55] D. R. Gossett, H. T. Tse, S. A. Lee, Y. Ying, A. G. Lindgren, O. O. Yang, J. Rao, A. T. Clark, and D. Di Carlo, "Hydrodynamic stretching of single cells for large population mechanical phenotyping," *Proceedings of the National Academy of Sciences*, vol. 109, no. 20, pp. 7630–7635, 2012.

[56] K. Svoboda and S. M. Block, "Biological applications of optical forces," *Annual review of biophysics and biomolecular structure*, vol. 23, no. 1, pp. 247–285, 1994.

[57] J. Guck, R. Ananthakrishnan, H. Mahmood, T. J. Moon, C. C. Cunningham, and J. Käs, "The optical stretcher: a novel laser tool to micromanipulate cells," *Biophysical journal*, vol. 81, no. 2, pp. 767–784, 2001.

[58] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophysical journal*, vol. 82, no. 5, pp. 2775–2783, 2002.

[59] K. Franze and J. Guck, "The biophysics of neuronal growth," *Reports on Progress in Physics*, vol. 73, no. 9, p. 094601, 2010.

[60] Z. Bryant, M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli, and C. Bustamante, "Structural transitions and elasticity from torque measurements on dna," *Nature*, vol. 424, no. 6946, pp. 338–341, 2003.

[61] R. N. Bracewell and R. N. Bracewell, *The Fourier transform and its applications*, vol. 31999. McGraw-Hill New York, 1986.

[62] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[63] D. C. Ghiglia and M. D. Pritt, *Two-dimensional phase unwrapping: theory, algorithms, and software*, vol. 4. Wiley New York, 1998.

[64] Y. Gao, S. Zhang, T. Li, Q. Chen, S. Li, and P. Meng, "Adaptive unscented kalman filter phase unwrapping method and its application on gaofen-3 interferometric sar data," *Sensors*, vol. 18, no. 6, p. 1793, 2018.

[65] W. T. Welford, *Aberrations of optical systems*. Routledge, 2017.

[66] A. Ashkin and J. Dziedzic, "Optical levitation by radiation pressure," *Applied Physics Letters*, vol. 19, no. 8, pp. 283–285, 1971.

[67] A. Ashkin, J. M. Dziedzic, and T. Yamane, "Optical trapping and manipulation of single cells using infrared laser beams," *Nature*, vol. 330, no. 6150, pp. 769–771, 1987.

[68] Y. Pang and R. Gordon, "Optical trapping of a single protein," *Nano letters*, vol. 12, no. 1, pp. 402–406, 2012.

[69] J. Ando, G. Bautista, N. Smith, K. Fujita, and V. R. Daria, "Optical trapping and surgery of living yeast cells using a single laser," *Review of Scientific Instruments*, vol. 79, no. 10, p. 103705, 2008.

[70] I. A. Favre-Bulle, A. B. Stilgoe, E. K. Scott, and H. Rubinsztein-Dunlop, "Optical trapping in vivo: theory, practice, and applications," *Nanophotonics*, vol. 8, no. 6, pp. 1023–1040, 2019.

[71] R. W. Applegate, J. Squier, T. Vestad, J. Oakey, and D. W. Marr, "Optical trapping, manipulation, and sorting of cells and colloids in microfluidic systems with diode laser bars," *Optics express*, vol. 12, no. 19, pp. 4390–4398, 2004.

[72] E. Bertseva, D. Grebenkov, P. Schmidhauser, S. Gribkova, S. Jeney, and L. Forró, "Optical trapping microrheology in cultured human cells," *The European Physical Journal E*, vol. 35, no. 7, pp. 1–8, 2012.

[73] J. Mas, A. C. Richardson, S. N. S. Reihani, L. B. Oddershede, and K. Berg-Sørensen, "Quantitative determination of optical trapping strength and viscoelastic moduli inside living cells," *Physical biology*, vol. 10, no. 4, p. 046006, 2013.

[74] M. Righini, P. Ghenuche, S. Cherukulappurath, V. Myroshnychenko, F. J. Garcia de Abajo, and R. Quidant, "Nano-optical trapping of rayleigh particles and escherichia coli bacteria with resonant optical antennas," *Nano letters*, vol. 9, no. 10, pp. 3387–3391, 2009.

[75] I. A. Vorobjev, H. Liang, W. H. Wright, and M. W. Berns, "Optical trapping for chromosome manipulation: a wavelength dependence of induced chromosome bridges," *Biophysical journal*, vol. 64, no. 2, pp. 533–538, 1993.

[76] W. H. Wright, G. Sonek, Y. Tadir, and M. W. Berns, "Laser trapping in cell biology," *IEEE Journal of Quantum electronics*, vol. 26, no. 12, pp. 2148–2157, 1990.

[77] R. Gussgard, T. Lindmo, and I. Brevik, "Calculation of the trapping force in a strongly focused laser beam," *JOSA B*, vol. 9, no. 10, pp. 1922–1930, 1992.

[78] T. Huber, A. Lambrecht, J. Schmidt, L. Karpa, and T. Schaetz, "A far-off-resonance optical trap for a ba+ ion," *Nature communications*, vol. 5, no. 1, pp. 1–7, 2014.

[79] M. J. Comstock, T. Ha, and Y. R. Chemla, "Ultrahigh-resolution optical trap with single-fluorophore sensitivity," *Nature methods*, vol. 8, no. 4, pp. 335–340, 2011.

[80] L. Mitchem and J. P. Reid, "Optical manipulation and characterisation of aerosol particles using a single-beam gradient force optical trap," *Chemical Society Reviews*, vol. 37, no. 4, pp. 756–769, 2008.

[81] N. Simpson, L. Allen, and M. Padgett, "Optical tweezers and optical spanners with laguerre–gaussian modes," *Journal of modern optics*, vol. 43, no. 12, pp. 2485–2491, 1996.

[82] N. Simpson, K. Dholakia, L. Allen, and M. Padgett, "Mechanical equivalence of spin and orbital angular momentum of light: an optical spanner," *Optics letters*, vol. 22, no. 1, pp. 52–54, 1997.

[83] M. Padgett and L. Allen, "Light with a twist in its tail," *Contemporary physics*, vol. 41, no. 5, pp. 275–285, 2000.

[84] L. Allen, M. W. Beijersbergen, R. Spreeuw, and J. Woerdman, "Orbital angular momentum of light and the transformation of laguerre-gaussian laser modes," *Physical review A*, vol. 45, no. 11, p. 8185, 1992.

[85] L. Allen, M. Padgett, and M. Babiker, "Iv the orbital angular momentum of light," in *Progress in optics*, vol. 39, pp. 291–372, Elsevier, 1999.

[86] A. Ashkin, "Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime," *Methods in cell biology*, vol. 55, pp. 1–27, 1997.

[87] M. Mir, B. Bhaduri, R. Wang, R. Zhu, and G. Popescu, "Quantitative phase imaging," *Progress in optics*, vol. 57, no. 133-37, p. 217, 2012.

[88] Y. Zhao, L. Cao, H. Zhang, D. Kong, and G. Jin, "Accurate calculation of computer-generated holograms using angular-spectrum layer-oriented method," *Optics express*, vol. 23, no. 20, pp. 25440–25449, 2015.

[89] T. Wilson, C. Shepparu, and K. Löschke, "Theory and practice of scanning optical microscopy," 1985.

[90] J. J. Stamnes, *Waves in focal regions: propagation, diffraction and focusing of light, sound and water waves*. Routledge, 2017.

[91] J. J. Braat, S. van Haver, A. J. Janssen, and P. Dirksen, "Assessment of optical systems by means of point-spread functions," *Progress in optics*, vol. 51, pp. 349–468, 2008.

[92] U. Schnars and W. Jüptner, "Direct recording of holograms by a ccd target and numerical reconstruction," *Applied optics*, vol. 33, no. 2, pp. 179–181, 1994.

[93] B. M. Hennelly and J. T. Sheridan, "Generalizing, optimizing, and inventing numerical algorithms for the fractional fourier, fresnel, and linear canonical transforms," *JOSA A*, vol. 22, no. 5, pp. 917–927, 2005.

[94] L. Mandel and E. Wolf, *Optical coherence and quantum optics*. Cambridge university press, 1995.

[95] M. Hillenbrand, D. P. Kelly, and S. Sinzinger, "Numerical solution of nonparaxial scalar diffraction integrals for focused fields," *JOSA A*, vol. 31, no. 8, pp. 1832–1841, 2014.

[96] N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast fourier transform approach," *JOSA A*, vol. 15, no. 4, pp. 857–867, 1998.

[97] F. Shen and A. Wang, "Fast-fourier-transform based numerical integration method for the rayleigh-sommerfeld diffraction formula," *Applied optics*, vol. 45, no. 6, pp. 1102–1110, 2006.

[98] L. Onural, "Exact analysis of the effects of sampling of the scalar diffraction field," *JOSA A*, vol. 24, no. 2, pp. 359–367, 2007.

[99] P. Picart and P. Tankam, "Analysis and adaptation of convolution algorithms to reconstruct extended objects in digital holography," *Applied optics*, vol. 52, no. 1, pp. A240–A253, 2013.

[100] T. Kozacki, K. Falaggis, and M. Kujawinska, "Computation of diffracted fields for the case of high numerical aperture using the angular spectrum method," *Applied optics*, vol. 51, no. 29, pp. 7080–7088, 2012.

[101] W. Zhang, H. Zhang, C. J. Sheppard, and G. Jin, "Analysis of numerical diffraction calculation methods: from the perspective of phase space optics and the sampling theorem," *JOSA A*, vol. 37, no. 11, pp. 1748–1766, 2020.

[102] M. Testorf, B. Hennelly, and J. Ojeda-Castañeda, *Phase-space optics: fundamentals and applications*. McGraw-Hill Education, 2010.

[103] J. J. Healy, B. M. Hennelly, and J. T. Sheridan, "Additional sampling criterion for the linear canonical transform," *Optics letters*, vol. 33, no. 22, pp. 2599–2601, 2008.

[104] A. E. Siegman, *Lasers*. University science books, 1986.

[105] R. J. Noll, "Zernike polynomials and atmospheric turbulence∗," *J. Opt. Soc. Am.*, vol. 66, pp. 207–211, Mar 1976.

[106] J. C. Wyant and K. Creath, "Basic wavefront aberration theory for optical metrology," *Applied optics and optical engineering*, vol. 11, no. part 2, pp. 28–39, 1992.

[107] T. Jannson and J. Sochacki, "Primary aberrations of thin planar surface lenses," *JOSA*, vol. 70, no. 9, pp. 1079–1084, 1980.

[108] A. M. Yao and M. J. Padgett, "Orbital angular momentum: origins, behavior and applications," *Advances in optics and photonics*, vol. 3, no. 2, pp. 161–204, 2011.

[109] S. Zwick, T. Haist, Y. Miyamoto, L. He, M. Warber, A. Hermerschmidt, and W. Osten, "Holographic twin traps," *Journal of Optics A: Pure and Applied Optics*, vol. 11, no. 3, p. 034011, 2009.

[110] M. Habaza, M. Kirschbaum, C. Guernth-Marschner, G. Dardikman, I. Barnea, R. Korenstein, C. Duschl, and N. T. Shaked, "Rapid 3d refractive-index imaging of live cells in suspension without labeling using dielectrophoretic cell rotation," *Advanced Science*, vol. 4, no. 2, p. 1600205, 2017.

[111] T. Moura, U. Andrade, J. Mendes, and M. Rocha, "Silicon microparticles as handles for optical tweezers experiments," *Optics Letters*, vol. 45, no. 5, pp. 1055–1058, 2020.

[112] W. H. Campos, J. M. Fonseca, J. B. Mendes, M. S. Rocha, and W. A. Moura-Melo, "How light absorption modifies the radiative force on a microparticle in optical tweezers," *Applied Optics*, vol. 57, no. 25, pp. 7216–7224, 2018.

[113] K. M. Oliveira, T. A. Moura, J. L. Lucas, A. V. Teixeira, M. S. Rocha, and J. B. Mendes, "Use of organic semiconductors as handles for optical tweezers experiments: Trapping and manipulating polyaniline (pani) microparticles," *ACS Applied Polymer Materials*, 2023.

[114] W. Campos, T. Moura, O. Marques, J. Fonseca, W. Moura-Melo, M. Rocha, and J. Mendes, "Germanium microparticles as optically induced oscillators in optical tweezers," *Physical Review Research*, vol. 1, no. 3, p. 033119, 2019.

[115] L. Oliveira, W. H. Campos, and M. S. Rocha, "Optical trapping and manipulation of superparamagnetic beads using annular-shaped beams," *Methods and Protocols*, vol. 1, no. 4, p. 44, 2018.

[116] M.-C. Zhong, A.-Y. Liu, and F. Ji, "Opto-thermal oscillation and trapping of light absorbing particles," *Optics Express*, vol. 27, no. 21, pp. 29730–29737, 2019.