BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF MECHANICAL ENGINEERING

FAKULTA STROJNÍHO INŽENÝRSTVÍ

INSTITUTE OF PHYSICAL ENGINEERING

ÚSTAV FYZIKÁLNÍHO INŽENÝRSTVÍ

SHAPING OF LIGHT FIELDS FOR CONTROL AND CORRECTION OF ELECTRON BEAMS

TVAROVÁNÍ OPTICKÝCH POLÍ PRO KONTROLU A KOREKCI ELEKTRONOVÝCH SVAZKŮ

BACHELOR'S THESIS

BAKALÁŘSKÁ PRÁCE

AUTHOR
AUTOR PRÁCE

Tomáš Brada

SUPERVISOR
VEDOUCÍ PRÁCE

Ing. Martin Hrtoň, Ph.D.

BRNO 2024

Assignment Bachelor's Thesis

Institut:	Institute of Physical Engineering
Student:	**Tomáš Brada**
Degree programm:	Physical Engineering and Nanotechnology
Branch:	no specialisation
Supervisor:	**Ing. Martin Hrtoň, Ph.D.**
Academic year:	2023/24

As provided for by the Act No. 111/98 Coll. on higher education institutions and the BUT Study and Examination Regulations, the director of the Institute hereby assigns the following topic of Bachelor's Thesis:

Shaping of light fields for control and correction of electron beams

Brief Description:

Ability to control properties of an electron beam is the fundamental cornerstone of the modern electron microscopy. To reach the current level of image resolution and quality, a variety of technological solutions was developed and implemented. Among them, the ponderomotive interaction with light represents a novel approach to shaping electron wavefunction in free space. A key prerequisite facilitating this method is the ability to tailor the optical field at the scale of the electron beam diameter. That can be accomplished through various avenues, this work shall focus, however, on one particular solution utilizing a spatial light modulator (SLM). This programmable device allows to modify the wavefront of a passing light beam and ensure the desired spatial profile of the electric field intensity within the target area. Implementation of this techniques requires to find the phase profile that needs to be imprinted on the SLM in order to produce the desired output. This task is complicated by various factors stemming from the presence of other optical elements within the system, properties of the source etc. The goal of this work should be to anticipate those factors, take them into account during the theoretical calculation of the phase profile fed to the SLM and possibly also test the whole concept experimentally. Ultimately, the outcome of this work should serve as a stepping stone towards development of an electron microscope that employs the electron beam shaping by intense light fields.

Bachelor's Thesis goals:

1. Carry out research on the current literature dealing with calculations of phase holograms for SLMs.
2. Using a suitable algorithm, calculate phase holograms that produce intensity profiles relevant to electron beam shaping by light fields.
3. Test the obtained holograms in experimental conditions.

Recommended bibliography:

GERCHBERG, R. W. a SAXTON, W. O. A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures. Optik. 1972, roč. 35, č. 2, s. 237-246.

CHIRITA MIHAILA, Marius Constantin; WEBER, Philipp; SCHNELLER, Matthias; GRANDITS, Lucas; NIMMRICHTER, Stefan et al. Transverse Electron-Beam Shaping with Light. Physical Review X. 2022, roč. 12, č. 3, 031043. ISSN 2160-3308. Dostupné z: https://doi.org/10.1103/PhysRevX.12.031043.

KOMRSKA, Jiří. Vlnová optika: část Difrakce světla: přednášky pro 4. ročník oborů "Fyzikální inženýrství" a "Přesná mechanika a optika" na FSI VUT v Brně. Brno: Akademické nakladatelství CERM, 2004. ISBN 80-214-2822-8.

Deadline for submission Bachelor's Thesis is given by the Schedule of the Academic year 2023/24

In Brno,

L. S.

_____	_____
prof. RNDr. Tomáš Šikola, CSc.	doc. Ing. Jiří Hlinka, Ph.D.
Director of the Institute	FME dean

Abstrakt
Elektrony v elektronovém mikroskopu jsou běžně řízeny statickými elektrickými a magnetickými poli. Existuje však také možnost modifikovat dráhu elektronu skrze ponderomotivní potenciál vytvořený laserovým pulsem. Tato práce se zabývá tvorbou speciálně tvarovaných laserových svazků pro potřeby tohoto zařízení. V první části je vysvětlen princip fungování elektronového mikroskopu řízeného světlem. Následuje přestavení prostorového modulátoru světla, který bude k tvarování svazku využit. Poté je rozebráno hledání fázové masky pro prostorový modulátor, což zahrnuje komentář simulačních metod propagace světla, rešerši iterativních algoritmů pro problém hledání fáze a postup vývoje vlastního algoritmu. Na závěr je vyvinutý algoritmus otestován v experimentu.

Summary
Electrons in an electron microscope are commonly controlled by static electric and magnetic fields. However, there is also the possibility to modify the electron trajectory through the ponderomotive potential created by a laser pulse. This work focuses on creating specially shaped laser beams for this purpose. The first part explains the principle of operation of a light-controlled electron microscope. This is followed by an introduction to the spatial light modulator, which will be used for beam shaping. Then, the search for a phase mask for the spatial modulator is discussed, including commentary on simulation methods for light propagation, a review of iterative algorithms for phase retrieval, and the development process of a custom algorithm. Finally, the developed algorithm is tested in an experiment.

Klíčová slova
ponderomotivní síla, tvarování svazku, prostorový modulátor světla, Gerchbergův-Saxtonův algoritmus

Keywords
ponderomotive force, beam shaping, spatial light modulator, Gerchberg-Saxton algorithm

Hereby I declare that I have created this work autonomously under a scientific supervision of Ing. Martin Hrtoň, Ph.D. All sources, references and literature used or excerpted during the elaboration of this work are properly cited and listed in the complete reference.

Tomáš Brada

Contents

1. Introduction

The development of advanced imaging techniques has revolutionized the field of microscopy, providing insights into the microscopic world. Among these advancements, electron microscopy has emerged as a powerful tool, enabling researchers to visualize structures at the atomic level with a remarkable precision. However, despite its capabilities, traditional electron microscopy faces limitations in manipulating and controlling the electron beam with a high enough precision. The integration of the laser into electron microscopy offers a new way to manipulate electron beams, leading to the concept of a light-controlled electron microscope.

This bachelor thesis focuses on one specific problem in the development of a light-controlled electron microscope, which is the delivery of an on-demand shaped laser beam for enhancing the functionality and precision of the electron beam. The shaping of the laser beam is achieved using a Spatial Light Modulator (SLM), a device capable of modulating the phase of light to generate complex light patterns. By incorporating the SLM into the electron microscope setup, it becomes possible to manipulate the electron beam with unprecedented control, opening new avenues for advanced imaging and analysis.

The integration of a shaped laser beam into electron microscopy offers several advantages. It allows for the precise control of the electron beam's trajectory, enabling targeted illumination of specific regions of a sample. Another application involves correction of abberations where we can take advantage of wide variety of reachable laser beam shapes which provide much more freedom than conventional correction techniques. As the use of an SLM introduces flexibility and adaptability, it opens a door to dynamical alterations of the beam profile to suit different experimental requirements such as imaging weak-contrast or beam-sensitive samples.

This thesis is structured as follows: Chapter 1 provides an overview of the principles of electron microscopy and the challenges associated with beam control. Chapter 2 discusses the theoretical background of laser beam shaping and the role of Spatial Light Modulators. Chapter 3 details the experimental setup and methodology employed in the integration of the SLM into the electron microscope. Chapter 4 presents the results of the beam shaping experiments, highlighting the improvements in beam control and imaging capabilities. Finally, Chapter 5 concludes the thesis with a summary of the findings and potential future directions for research.

The successful development of a light-controlled electron microscope represents a significant step forward in the field of microscopy. By harnessing the power of shaped laser beams, this innovative approach promises to enhance the precision, versatility, and capabilities of electron microscopy, paving the way for new discoveries and applications in science and technology.

2. Beam Shaping

Laser light has various applications across many disciplines of science, each placing different requirements on laser beam properties. The choice of laser source determines the initial intensity distribution and beam size, but these parameters can be further manipulated by placing appropriate optical elements in the path of the beam. Many of these optical elements serve a single preset function. For example, an axicon (specialized type of lens which has a conical surface) can create Bessel beams. The focus of this work is not on these single-purpose elements. Instead, I will concentrate on a general-purpose device capable of generating arbitrary laser beams. This device is called Spatial Light Modulator (SLM) and its properties and functionality are discussed later in section 3.2.

2.1. Examples of Shaped Beams Applications

To demonstrate that there are many applications of shaped laser beams I will name a few examples from research and industry:

- Ultrashort laser welding - typical Gaussian beam can be reshaped into a Bessel beam that is capable to meet the higher precision demands in advanced joining techniques. [1]

- Trapping Rydberg atoms - excitation and trapping Rydberg atoms is important in the field of quantum computation, sensing, imaging and others. [2]

- Parallelization of micromachining - laser can be divided into a grid of spots that can be used to manufacture repetitive structures in a shorter time. [3]

- Optical tweezers - shaped laser beam can be used to manipulate small particles. By utilizing a Spatial Light Modulator, it is possible to smoothly move the particles across small areas without changing anything apart from the phase mask uploaded on the SLM from a computer. [4]

- Generating optical vortices - the Spatial Light Modulator has the capability to generate laser beams with a spiral phase that winds around a phase singularity [5]. These beams, known as optical vortices, possess angular momentum. They have numerous applications, including optical tweezers, optical encryption, and other advanced technologies.

These applications and many more are the subject of intense research nowadays.

2.2. Control and Correction of Electron Beams

Another potential application could be preparing laser pulses with a custom spatial intensity distribution that can interact with electron beams in a predictable way. This is a potentially promising way to modulate electron beams in electron microscopes. This work will be focused on creating shaped laser beams for this particular application.

2.2.1. Interaction of Electrons with a Laser Pulse

An intuitive description of what happens when an electron passes through a focused laser field is offered by the concept of **ponderomotive force**.

Details about the ponderomotive force can be found in chapter V.§30. of [6] or in chapter 21.1. of online source [7]. This force occurs in situations when an object is exposed to a potential that consists of a smoothly varying component and rapidly varying oscillations. Position of the object x can be written as a sum of a slower motion X and rapid oscillations ξ:

$$x(t) = X(t) + \xi(t). \tag{2.1}$$

Interestingly, $X(t)$ is not independent of the oscillations. A good example that demonstrates how motion can be changed by a rapidly oscillating force is a pendulum in a gravitational field. Under normal conditions, the only stable position is when the pendulum points downward. However, when the pivot point of the pendulum starts to oscillate in the vertical direction, the pendulum becomes also stable when it points straight up.

Ponderomotive force can also act on a particle with a charge q in an oscillating electric field $E\cos(\omega t)$. The force on this particle is

$$F = qE\cos(\omega t). \tag{2.2}$$

Invoking the Newton's second law, the acceleration can be expressed as

$$a = \frac{F}{m} = \frac{qE}{m}\cos(\omega t), \tag{2.3}$$

where m is the mass of the particle. Position x can then be calculated by a two-fold integration over time:

$$x = \frac{-a}{\omega^2} = \frac{qE}{m\omega^2}\cos(\omega t). \tag{2.4}$$

The particle experiencing a harmonic motion has a time-averaged energy

$$U = \frac{1}{2}m\omega^2\langle x^2\rangle = \frac{q^2E^2}{4m\omega^2}. \tag{2.5}$$

Finally, force acting on the particle can be expressed as a negative gradient of the energy:

$$F_p = -\frac{e^2}{4m\omega^2}\nabla(E^2). \tag{2.6}$$

Equation (2.6) shows that the charged particle moves towards areas of a weaker field strength (this applies to both positively and negatively charged particles). That means that if we made an electron beam interact with an electric field whose intensity increases with a radial distance, the electron beam would be focused because off-axis electrons would be pushed towards the axis.

2.2.2. Light-Controlled Electron Microscope

Work presented in this thesis is a part of a larger project of the team Artemis from the Institute of Physical Engineering at the Faculty of Mechanical Engineering, Brno University of Technology. The project's goal is to design and assemble an experimental

setup similar to one described in [8] (see also Figure 2.1). It is a scanning electron microscope equipped with an ultrafast laser producing a beam that is diverted into two branches. While the first branch is used to eject an electron pulse from the electron gun via fotoemission, the laser in the second branch is shaped and then interacts with the electron beam as it passes through the electron column. The objectives are to demonstrate lensing, beam scanning and abberation correction which is schematically depicted in Figure 2.1. Later, the focus will shift towards exploration of applications involving unconventionally shaped electron beams.
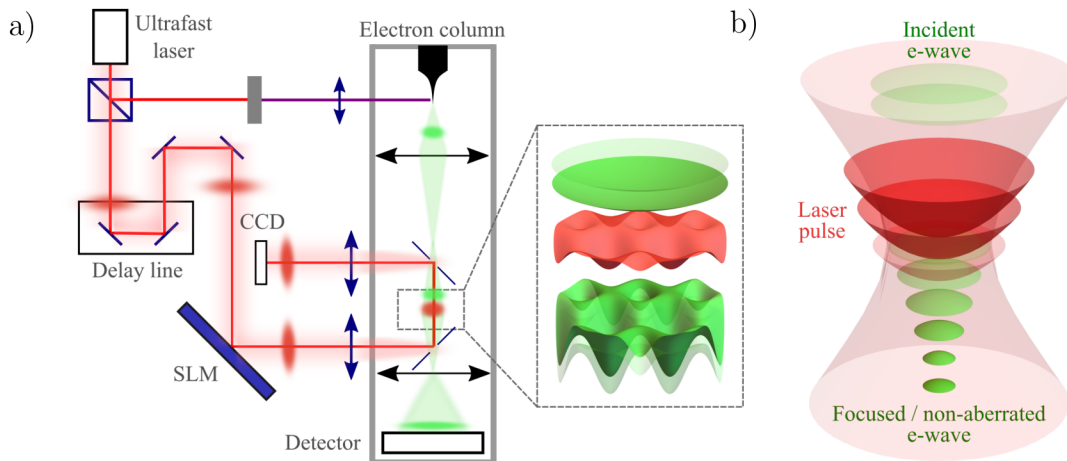

Figure 2.1: a) Scheme of electron microscope equipped with ultrafast laser (CCD - Charged-Couple Devide - type of camera sensor, SLM - Spatial Light Modulator - details in Chapter 3), b) Correction of electron beam abberations by interaction with shaped laser pulse. Taken from the archive of Dr. Andrea Konečná, with her kind permission.

This project brings many theoretical and practical problems and it requires cooperation among many members of our team.

This thesis specifically concentrates on the laser beam shaping process before its interaction with the electron beam. Both theoretical and practical aspects of this problem will be addressed. One of the outputs should be a simple experimental optical setup for beam shaping that would be able to tailor the Gaussian beam from the laser source to a chosen testing intensity distribution, e.g. one possessing a parabolic radial profile (Figure 2.2) that should have, considering the discussion accompanying Equation (2.6), a converging effect on the passing electron beam.
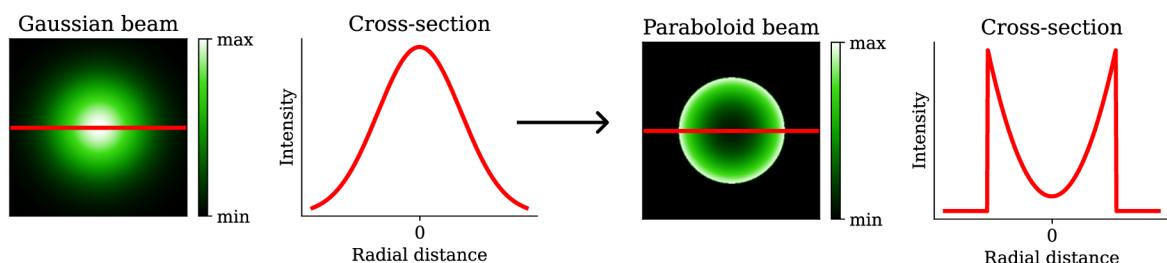

Figure 2.2: Reshaping of a Gaussian laser beam into a beam with a parabolic radial profile.

3. Spatial Light Modulator (SLM)

In this chapter, I will introduce a device that is commonly used for beam shaping. This device is called SLM, which stands for Spatial Light Modulator. In its essence, it is a liquid crystal display similar to a regular LCD monitor, but in this case, liquid crystals are used to spatially modify amplitude or phase of an incident beam. My focus will be on phase-only SLMs. Their main advantage is that they do not, in principle, dissipate any power from the initial beam, which makes them more energy efficient.

I will first describe the physical principle behind the device. Then I will break down the design of the SLM. At the end of the chapter, a specific setup for beam shaping will be discussed.

3.1. Physical Principle

3.1.1. Phase Shift

The fundamental principle behind the phase tunable alteration of light involves the use of materials with refractive index n that can be modified. When light propagates trough a medium with a higher n, its spatial frequency k increases, which results in a different output phase ϕ compared to light travelling the same distance in air. This concept is illustrated in Figure 3.1, where two waves start with the same phase, but they end up with a mutual phase shift after travelling a distance d.
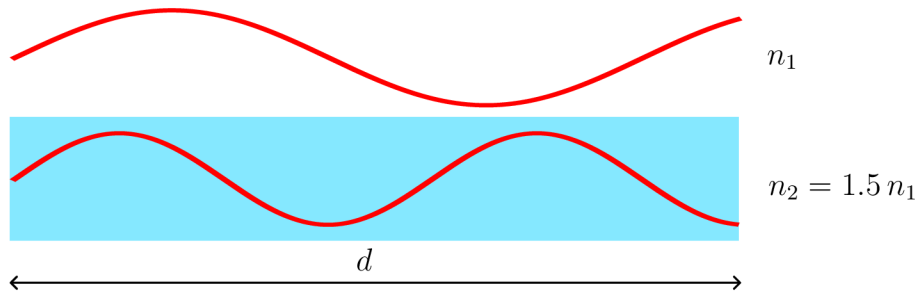
Figure 3.1: Light waves travelling in media with different refraction indices.

This phase shift can be expressed as

$$\Delta\phi = \frac{2\pi}{\lambda}(n_2 - n_1)d. \tag{3.1}$$

3.1.2. Anisotropic Materials

Now we need a material that can actively change its refractive index n. This can be achieved with some certain kinds of **anisotropic materials** (chapter 6.3. in [9]) which I will dive into now. First, fundamentals of the underlying theory will be outlined and next one specific approach for modifying refractive index will be described.

A dielectric optical medium is called anisotropic if its optical macroscopic properties depend on the propagation direction and polarization of a passing light wave. The macroscopic properties are determined by microscopical properties like shape and orientation of molecules and their organisation in space.

In **crystals** (anisotropic dielectric media), the electric displacement field $\vec{D}$ is a linear combination of the three components of the electric field $\vec{E}$, with the vacuum permittivity ε_0 as the constant of proportionality. This relationship is given by $\vec{D} = \varepsilon_0 \hat{\varepsilon} \vec{E}$, which can be mathematically expressed as a matrix multiplication:

$$\begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} = \varepsilon_0 \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}, \tag{3.2}$$

where the 3×3 electric permittivity tensor describes the material properties of the crystal. This tensor is symmetrical which means that it contains 6 independent elements. Due to the symmetry, we can find a coordinate system, in which off-diagonal elements vanish and the matrix is purely diagonal. That implies that $\vec{E}$ and $\vec{D}$ are parallel in these directions and the associated permittivity components are called principal permittivity indices. Therefore, the permittivity tensor reduces to

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon_x & 0 & 0 \\ 0 & \varepsilon_y & 0 \\ 0 & 0 & \varepsilon_z \end{pmatrix}. \tag{3.3}$$

Every principal permittivity has its corresponding index of refraction:

$$n_x = \sqrt{\varepsilon_x}, \quad n_y = \sqrt{\varepsilon_y}, \quad n_z = \sqrt{\varepsilon_z}. \tag{3.4}$$

Here, ε_0 denotes the permittivity of free space. Phase velocity of light propagation can be calculated as $v = c/n$. Hence, it is obvious that phase velocity of light travelling through an anisotropic medium depends on the direction of propagation.

3.1.3. Index Ellipsoid

There is a convenient way to geometrically visualize the relationship between the direction of light propagation and the phase velocity. It comes from the fact that electric field energy density can be calculated as

$$w_E = \frac{1}{2}\vec{E} \cdot \vec{D} = \frac{1}{2}\left[\frac{D_x^2}{\varepsilon_0 n_x^2} + \frac{D_y^2}{\varepsilon_0 n_y^2} + \frac{D_z^2}{\varepsilon_0 n_z^2}\right]. \tag{3.5}$$

This can be rewritten as

$$\frac{D_x^2}{\frac{2\varepsilon_0 w_{EF}}{n_x^2}} + \frac{D_y^2}{\frac{2\varepsilon_0 w_{EF}}{n_y^2}} + \frac{D_z^2}{\frac{2\varepsilon_0 w_{EF}}{n_z^2}} = \frac{\alpha^2}{n_x^2} + \frac{\beta^2}{n_y^2} + \frac{\gamma^2}{n_z^2} = 1 \tag{3.6}$$

which can be recognized as the equation of an ellipsoid.

Figure 4.1 a) shows an outline of the ellipsoid under consideration. Keeping in mind that every polarization state can be decomposed into two mutually orthogonal linear polarizations, one can determine the refractive indices perceived by an electromagnetic wave travelling in the direction of $\vec{k}$ using the following procedure.

First, we need to draw vector $\vec{k}$ and a plane that is perpendicular to it and passes through the origin. The intersection of the plane and the ellipsoid forms an ellipse and its principal axes give us the index of refraction for the two polarizations.
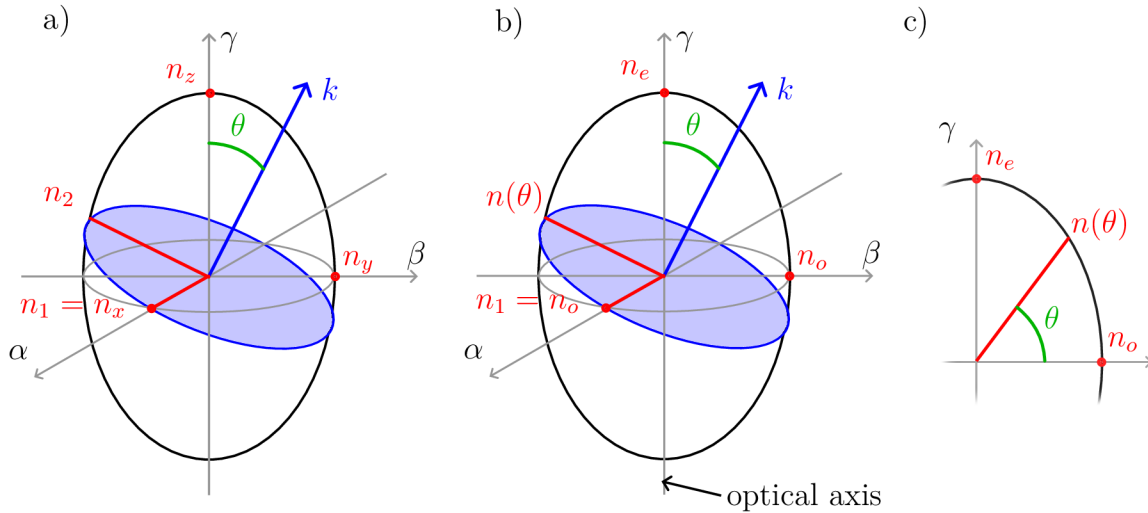
Figure 3.2: Index ellipsoid for a) a general case where $n_x \neq n_y \neq n_z$ b) an uniaxial crystal where $n_o = n_x = n_y \neq n_e = n_z$, c) a sectional view of the index ellipsoid of an uniaxial crystal showing geometry that is important for the derivation of a formula for $n(\theta)$.

3.1.4. Uniaxial Crystals

We will be particularly interested in crystals with a certain kind of symmetry, where two out of the three principal refractive indices are equal. Crystals with this property are called **uniaxial**. Their refractive indices are usually denoted as ordinary n_o ($= n_x = n_z$) and extraordinary n_e ($= n_z$) which is indicated in figure 4.1 b). It turns out that for one polarization, refractive index is always n_o, while for the other, the refractive index depends on the angle between the direction in which the light wave propagates and the optical axis of the uniaxial material (marked in figure 4.1 c)). Following formula can be derived

$$\frac{1}{n^2(\theta)} = \frac{\cos^2\theta}{n_o^2} + \frac{\sin^2\theta}{n_e^2}, \tag{3.7}$$

which gives us a very convenient way of designing a material with a tunable refractive index! More specifically, if we succeed in creating a material composed of aligned uniaxial crystals with a proper orientation with respect to the passing light wave, we can modify the refractive index (and consequently also the light wave) by rotating the crystals.

3.1.5. Liquid Crystals

Now we need to find a material that is composed of uniaxial crystals that can be controllably rotated. Liquid crystals meet this requirement.

Matter in a liquid-crystal state consists of elongated (typically ellipsoidal) molecules that possess an orientation order (like crystals) but lack positional order (like liquid). For more details see Section 6.5 in [9].

There are three types of liquid crystal arrangements. SLMs typically use nematic liquid crystals where molecules tend to be parallel but their positions are random (illustrated in Figure 3.3). Their orientation can be changed by applying external electric field. After the field is removed, they return to their original positions. Therefore, liquid crystals
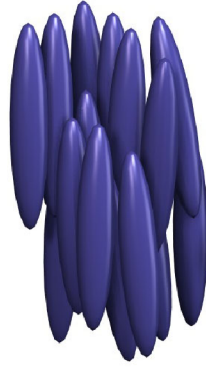
Figure 3.3: Nematic crystals [10].

represent a controllable anisotropic environment. This property is widely used in LCD displays. Their rapid development ensured that many practical issues of liquid crystals were addressed and resolved which makes their usage for other applications much easier.

3.2. Design of SLM

3.2.1. Display Structure

To take advantage of the properties of liquid crystals, we need to come up with a design of a pixel that contains these crystals and has electrodes to which voltage can be applied. One pixel then serves as a building block for a whole display that can contain millions of them.

Figure 3.4 shows the different layers of a pixel. Depending on the type of SLM, pixel can be transparent or reflective. Our laboratory is equipped with the reflective one, where light travels back and forth which means that it passes twice through the layer of liquid crystals. The whole display has a glass cover. Underneath it, there is a transparent electrode that facilitates the generation of a homogeneous electric field across the liquid crystal layer and alignment layers that surround it. These layers determine orientation of crystals when no electric field is present. At the very bottom, there is a sillicon layer with electrodes. That is why this type of devices is called LCoS (Liquid Crystals on Sillicon).
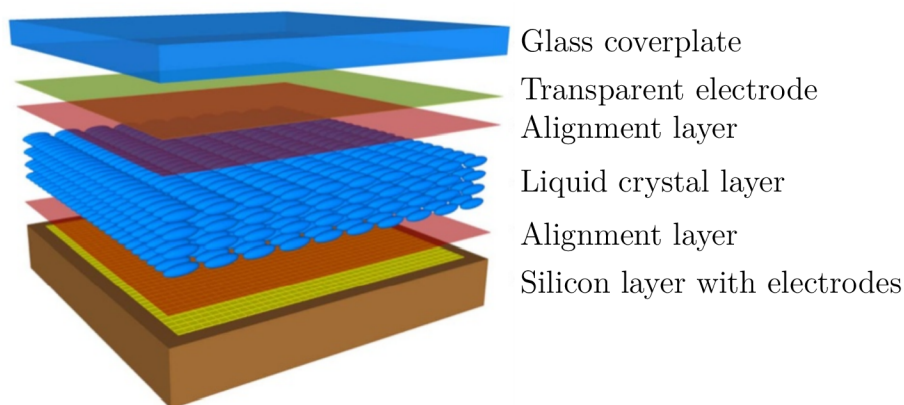

Figure 3.4: The schematic structure of an Liquid Crystal on Silicon device (taken over and modified from [11]).

One pixel can cause phase delay that is controlled by a locally applied voltage. Display consisting of many pixels can imprint a spatially varying phase mask to the incident beam and in this way, it is possible to mimic the function of a number of optical elements. For example, we can choose a phase mask that causes a larger phase shift on pixels that are further from the centre. This situation is depicted in figure 3.5. After being reflected from the SLM, an incident plane wave is transformed into a converging spherical wave, which means that the display acts as a converging lens.
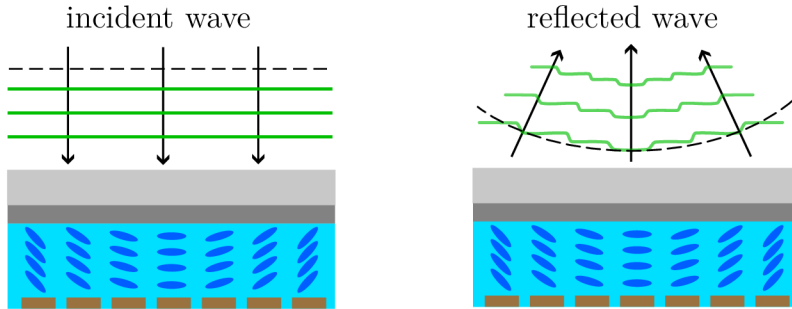
Figure 3.5: The schematic structure of an Liquid Crystal on Silicon device (taken over and modified from [11]).

The phase mask can not be perfectly continuous due to the finite number of pixels. However, a sufficiently large number of pixels can yield good enough results. Parameters of the SLM display will be discussed in the next subsection.

3.2.2. PLUTO-2.1 LCOS Spatial Light Modulator

There are many manufacturers of phase-only SLMs on the market. each with slight variations in their device design. I will focus on the design of the PLUTO-2.1 LCOS Spatial Light Modulator (Figure 3.6) because I used it in my experiment. However, other models are similar in principle.

Figure 3.6: PLUTO-2.1 LCOS Spatial Light Modulator.

The device consists of a driver unit that has data and power inputs. The driver unit is connected to the SLM display by a flex cable. Many different displays can be connected to the driver unit. Because of the intended application, a display that can handle high energy laser pulses is needed. This requirement is fulfilled by the PLUTO-2.1-NIR-149 model.

I will list some specifications of the device in the table for a better picture of the SLM's capabilities (tab: 3.1).

display area	15,36 × 8,64 mm^2 (0,7" diagonal)
resolution	1920 × 1080
pixel pitch	8,0 μm
fill factor	93 %
wavelength range	1000 - 1100 nm
max. phase shift	$2,9\pi$ Rad at 1064 nm
reflectivity	93 %
signal formats	HDMI - HDTV resolution
frame rate	60 Hz
adressing	8 Bit (256 grayscale levels)

Table 3.1: Selected specifications of the PLUTO-2.1 LCOS Spatial Light Modulator with PLUTO-2.1-NIR-149 display [12].

When the light field interacts with the display, each pixel has the capability to induce unique modifications. It is crucial that the maximum phase shift is no less than 2π. This requirement guarantees the ability to imprint arbitrary phase masks onto the incident optical field, as complex exponentials are periodic functions with a period of 2π.

3.2.3. Experimental Setup with SLM

In order to create spatial laser intensity distribution determined by the phase mask loaded on the SLM, a setup shown in Fig. 3.7 was employed. It consists of a laser source, a beam expander, the previously indtroduced SLM, a lens, a CCD camera, and a computer.
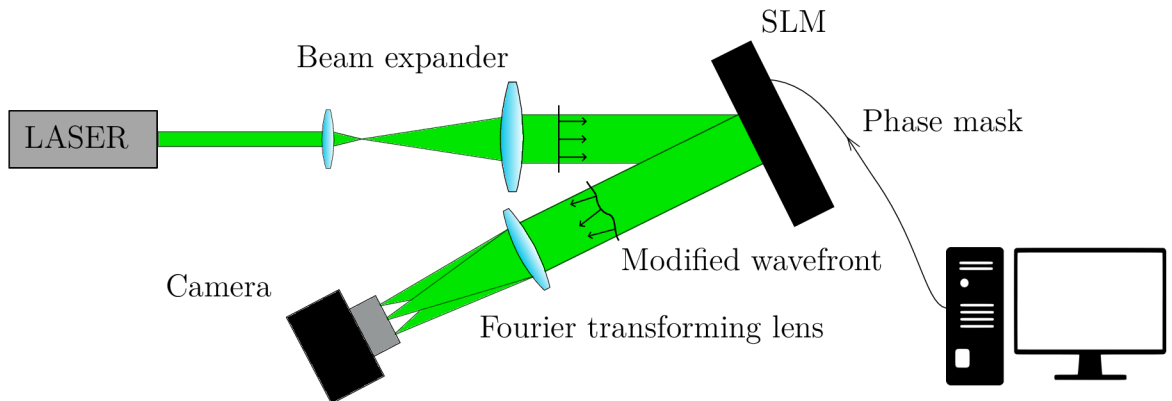

Figure 3.7: Experimental setup for SLM utilization.

The **laser source** produces quasi-monochromatic light with a linear polarization that is parallel to the long axis of the liquid crystals in the SLM, which ensures that the incident laser beam will experience a change in the refractive index as we rotate the crystal according to (3.7). The other polarization would not be modified by the SLM because its refractive index is always n_o regardless of the tilt of the crystal. In practice, this is achieved with the help of an adjustable polarization filter. The **Beam expander** ensures that a larger part of SLM display is covered by the laser. Without it, we would suffer a loss in resolution. The **SLM** fulfills the task of imprinting a phase mask to the laser. After the reflection from the SLM, the laser reaches the **lens**. The final shaped beam is then observed by **camera** in the focal plane of the lens. It turns out that the field at the SLM display and the field in the focal plane of the lens are linked together via two-dimensional Fourier transform. This relation is derived in detail in Chapter 4.

4. Light Propagation from the SLM to the Image Plane

In order to calculate the propagation of light, a solid mathematical foundation is necessary. Given that this thesis relies on computer simulations, it is necessary that these mathematical expressions can be computed efficiently on a personal computer. Fortunately, diffraction theory offers tools that meet all these requirements.

4.1. Diffraction Integrals

The propagation of light can be effectively described using diffraction integrals. These integrals are derived in accordance with the Huygens-Fresnel principle. First, it is essential to establish the geometrical context of the problem.

The coordinate system is set up in such a way that light propagates along the z-axis. The object plane is situated at $z = 0$, and each point A in the initial plane is defined by coordinates x' and y'. Point B represents the observation point in the image plane where the wavefunction will be computed. The distance between the points A and B is denoted as s.
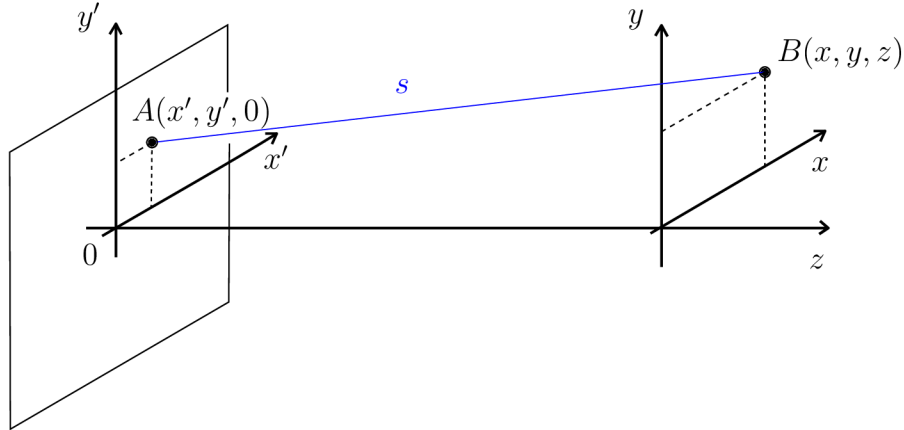
Figure 4.1: Geometry involved in solving the diffraction integral.

The incident wave $\psi(x, y, 0) = \psi_0(x', y')$ interacts with the initial plane, which is bounded by an aperture. Later, this plane will serve as the surface of a Spatial Light Modulator (SLM). A wave is fully characterized by its electric field E and phase ϕ at each point on a plane perpendicular to the propagation axis. The propagated wavefunction $\psi(x, y, z > 0)$ can be represented through double integrals, referred to as diffraction integrals, as detailed in the foundational source for this chapter [13].

The general diffraction integral, in accordance with the Huygens-Fresnel theory, can be expressed as follows:

$$\psi(P) = -\frac{ik}{2\pi} \iint_{S_0} \psi_0(x', y') \frac{\exp(iks)}{s} \cos \vartheta \, dS_0, \tag{4.1}$$

Here, k represents the wavenumber of light, S_0 is the surface of the initial plane where the light intensity is non-zero, and ϑ is the angle between a line perpendicular to the wavefront at the point B and the line connecting the points A and B.

Diffraction integral from equation (4.1) can be rewritten using the coordinates introduced introduced above:

$$\psi(x, y, z) = \frac{1}{i\lambda} \iint \psi(x', y', 0) \frac{z \exp\left(ik\sqrt{(x - x')^2 + (y - y')^2 + z^2}\right)}{(x - x')^2 + (y - y')^2 + z^2} \mathrm{d}x' \mathrm{d}y'. \tag{4.2}$$

While (4.2) is the most accurate formula for calculating scalar diffraction, its practical application is limited. In the most general case, it must be evaluated for every pair of points from the object and image planes, causing the computation time to increase with the number of pixels N as N^4 leading to unreasonable computation times. It can be reduced by using the same discretization for both planes and by adopting the paraxial Fresnel approximation leading to a computational complexity that increases only as N^2.

4.1.1. Fresnel Diffraction Integral

The problematic part of Formula (4.2) is the spherical term, since it ties together the two different sets of coordinates. These sets can be decoupled by assuming that the radial distance $\rho^2 = (x - x')^2 + (y - y')^2$ is much smaller than the distance z between the object plane and the image plane. This condition is known as the paraxial approximation and it allows us to replace the distance s by only few leading terms of its Taylor expansion, namely:

$$s = \sqrt{\rho^2 + z^2} = z\sqrt{1 + \frac{\rho^2}{z^2}} = z\left[1 + \frac{\rho^2}{2z^2} - \frac{\rho^4}{8z^3} + ...\right] \approx z + \frac{\rho^2}{2z}. \tag{4.3}$$

The distance s appears in several places of Equation (4.2) that are, from the mathematical standpoint, not equivalent. While the term $1/s$ can be replaced by $1/z$ without introducing any substantial error, the rapidly oscillating exponential e^{iks} requires a more careful approach. Provided the radial distance ρ satisfies the condition

$$k\frac{\rho^4}{8z^3} \ll 2\pi, \tag{4.4}$$

it is sufficient to retain the first two terms of the Taylor series in the argument of the exponential and the spherical wave then acquires the following simplified form:

$$\frac{\mathrm{e}^{iks}}{s} \approx \frac{1}{z}\mathrm{e}^{ikz}\,\mathrm{e}^{\frac{ik}{2z}\left[(x-x')^2 + (y-y')^2\right]}. \tag{4.5}$$

If we substitute spherical wave in the Rayleigh-Sommerfeld integral by this approximate expression, we obtain a formula known in the literature as the Fresnel diffraction integral

$$\psi(x, y, z) = \frac{1}{i\lambda}\frac{\mathrm{e}^{ikz}}{z}\iint \psi(x', y', 0)\,\mathrm{e}^{\frac{ik}{2z}\left[(x-x')^2 + (y-y')^2\right]}\,\mathrm{d}x'\,\mathrm{d}y'. \tag{4.6}$$

This expression can be recast into a more convenient form by expanding the argument of the exponential into individual terms and taking out those terms that do not depend on the integration variables in front of the integral.

$$\psi(x,y,z) = \frac{1}{i\lambda}\frac{\mathrm{e}^{ikz}}{z}\,\mathrm{e}^{\frac{ik}{2z}(x^2+y^2)}\iint \psi(x',y',0)\,\mathrm{e}^{\frac{ik}{2z}(x'^2+y'^2)}\,\mathrm{e}^{-\frac{ik}{z}(xx'+yy')}\,\mathrm{d}x'\,\mathrm{d}y' \qquad (4.7)$$

One can recognize the core of the 2D Fourier transform and one additional term that, as we will learn in the next section, cancels out with the phase transformation introduced by a thin lens. But first, we need to derive the phase transformation effect of the lens.

4.2. Thin Lens as a Phase Transformation

A lens inserted into an optical system modifies the phase of a passing electromagnetic wave. It is caused by the fact that optical paths and the phase delay of the optical ray depend on the position, namely thickness of the lens at the point of entry. We will consider a thin lens which means the points at which the light ray enters and exits the lens have the same distance from the optical axis.

In the following paragraphs, phase transformation will be calculated. For simplicity, we will consider a thin lens with only one curved surface. The result can be generalized to a lens with both sides curved which is shown in section 3.1 of [9].

Optical path Λ over a curve C can be computed as this integral
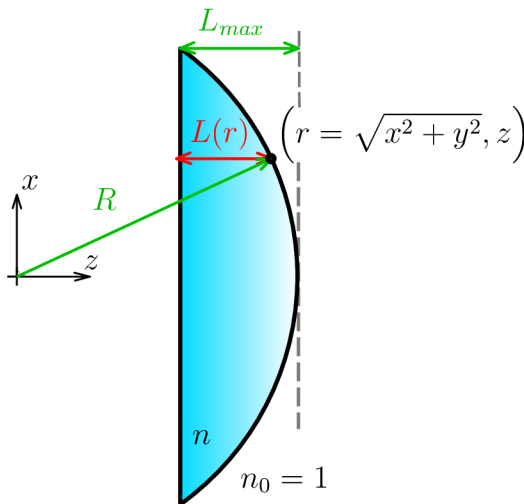
$$\Lambda = \int_C n\mathrm{d}s. \qquad (4.8)$$

Figure 4.2: Thin lens.

In our case (see Figure 4.2), the above integral becomes a sum with two contributions. Light first travels trough glass with a refractive index n and then trough air with a refractive index $n_0 = 1$. Considering the simple geometry of the problem, the distance

travelled by a ray in the glass depends only on the thickness function $L(r)$, which can be analytically expressed as

$$
\begin{aligned}
L(r) &= R - (L_{max} - z) \\
&= R - \left(L_{max} - \sqrt{R^2 - x^2 - y^2} \right) \\
&= R - \left(L_{max} - R\sqrt{1 - \frac{x^2 + y^2}{R^2}} \right) \\
&\approx R - \left[L_{max} - R\left(1 - \frac{x^2 + y^2}{2R^2} \right) \right].
\end{aligned} \tag{4.9}
$$

Phase delay $\Delta\phi = k\Lambda$, where k stands for the wavenumber, can be, in our case, computed as follows:

$$
\Delta\phi = knL(r) + k\left[L_{max} - L(r) \right]. \tag{4.10}
$$

Finally, we can introduce a phase function $f_{lens} = \mathrm{e}^{i\Delta\phi}$ that fully captures the deformation of the wavefront due to the lens.

$$
\begin{aligned}
f_{lens} &= \mathrm{e}^{i\Delta\phi} \\
&= \mathrm{e}^{iknL(x) + ik[L_{max} - L(r)]} \\
&= \mathrm{e}^{ikn\left\{ R - \left[L_{max} - R\left(1 - \frac{x^2 + y^2}{2R^2} \right) \right] \right\}} \mathrm{e}^{ikL_{max}} \mathrm{e}^{-ik\left\{ R - \left[L_{max} - R\left(1 - \frac{x^2 + y^2}{2R^2} \right) \right] \right\}} \\
&= \mathrm{e}^{const} \mathrm{e}^{-ikn\frac{x^2 + y^2}{2R}} \mathrm{e}^{ik\frac{x^2 + y^2}{2R}} \\
&= \mathrm{e}^{const} \mathrm{e}^{-ik\frac{x^2 + y^2}{2R}(n-1)} \\
&= \mathrm{e}^{const} \mathrm{e}^{-ik\frac{x^2 + y^2}{2f}}
\end{aligned} \tag{4.11}
$$

where any contributions that do not depend on r can be dropped. Also note that in the last step, we used the well known relation between focal length and parameters of the lens that is known from basic optics:

$$
\frac{1}{f} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2} \right), \tag{4.12}
$$

where R_1 and R_2 are radii of the lens surfaces and n is its index of refraction.

4.3. Fourier Transforming Property of a Lens

General formula for computing 2D Fourier transform of function $f(x, y)$ is

$$
F(f_x, f_y) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty}\!\!\!\int f(x, y)\, \mathrm{e}^{-i2\pi(f_x x + f_y y)} \mathrm{d}x\mathrm{d}y, \tag{4.13}
$$

where f_x and f_y are spatial frequencies in x and y directions, respectively. This operation and its counterpart, the inverse Fourier transform

$$f(x,y) = \frac{1}{\sqrt{2\pi}} \iint\limits_{-\infty}^{\infty} F(f_x, f_y) \, e^{i2\pi(f_x x + f_y y)} \mathrm{d}f_x \mathrm{d}f_y, \tag{4.14}$$

allow us to switch between spatial and frequency domains.

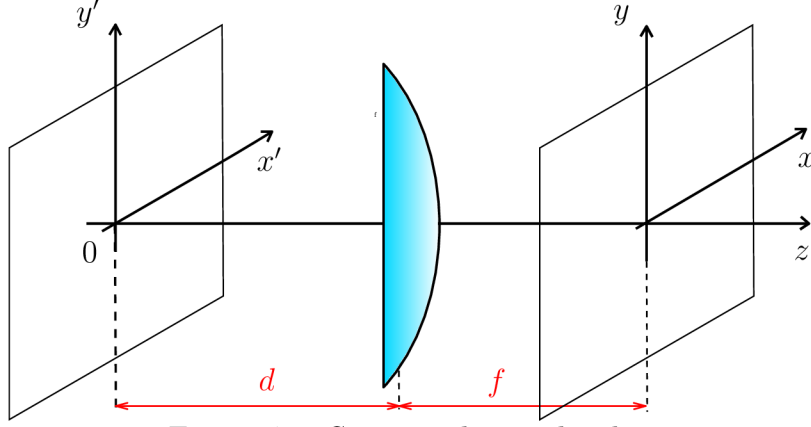Let us now consider an optical system that is schematically drawn in Figure 4.3.

Figure 4.3: Setup with one thin lens.

If we assume that the distance d from the object plane to the lens is very small $(d \longrightarrow 0)$, we can imprint the effect of the lens on the passing electromagnetic wave directly into the initial field, namely

$$\psi(x', y', 0) = \psi_0(x', y', 0) f_{lens} = \psi_0(x', y', 0) \, e^{-ik \frac{x'^2 + y'^2}{2f}}. \tag{4.15}$$

Substitution of $\psi(x', y', 0)$ into the Fresnel diffraction integral (4.7) yields

$$\psi(x,y,z) = \frac{1}{i\lambda} \frac{e^{ikz}}{z} e^{\frac{ik}{2z}(x^2+y^2)} \iint\limits_{-\infty}^{\infty} \psi_0(x',y',0) \, e^{-ik\frac{x'^2+y'^2}{2f}} \, e^{\frac{ik}{2z}(x'^2+y'^2)} e^{-\frac{ik}{z}(xx'+yy')} \mathrm{d}x' \mathrm{d}y'. \tag{4.16}$$

If we identify the focal length f as the propagation distance $(z = f)$, then the two exponentials in the expression cancel out, leaving only an integral corresponding to a Fourier transform of the initial field distribution (4.13). It is also convenient to express the angular wavenumber as $k = 2\pi/\lambda$ so that the form and meaning of spatial frequncies becomes clearer.

$$\psi(x,y,z) = \frac{1}{i\lambda} \frac{e^{ikf}}{f} e^{\frac{ik}{2f}(x^2+y^2)} \iint\limits_{-\infty}^{\infty} \psi_0(x',y',0) \, e^{-i2\pi(\frac{xx'}{\lambda f} + \frac{yy'}{\lambda f})} \mathrm{d}x' \mathrm{d}y' \tag{4.17}$$

By comparing (4.17) with (4.13), we obtain

$$\begin{aligned} f_x &= x/\lambda f, \\ f_y &= y/\lambda f. \end{aligned} \tag{4.18}$$

The Fourier transform relation between the field distribution in the object plane and the image plane (which in our case coincides with the focal plane of the lens) is not a

complete one, due to the presence of the quadratic phase factor before the integral. In the context of the electron beam shaping, however, the quantity of interest is only the intensity ($I \propto |\psi|^2$) of the field in the image plane and this quadratic phase factor is rendered irrelevant.

There is an useful intuition behind the Fourier transforming property of a thin lens. Initial field can be expressed in the form of an angular spectrum. That means that it is a superposition of plane waves travelling at different angles from the image plane. If a plane wave goes through an ideal lens, it is focused to a single point in the focal plane (see Figure 4.4). Spatial coordinates within the image plane are therefore linked to the angular coordinates of the object plane.
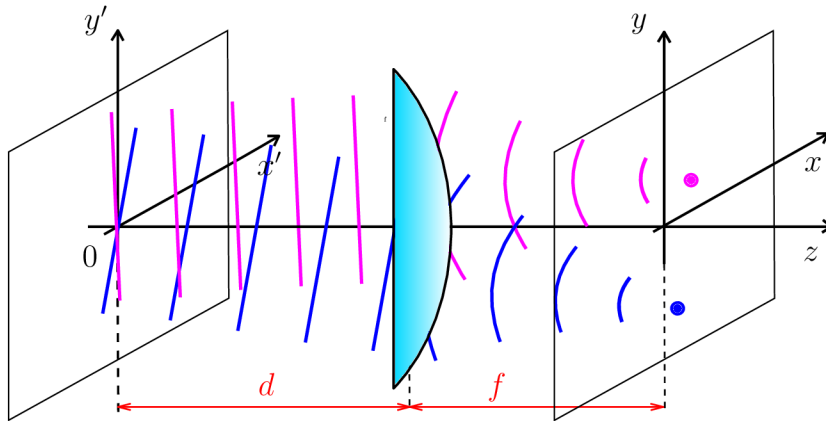

Figure 4.4: Lens focusing plane waves into points within the focal plane.

5. Iterative Algorithms for Phase Retrieval

So far, we have an active optical element for phase modulation called SLM (3.2.2), a setup in which the SLM can be effectively used to shape laser beams (3.2.3), and a mathematical tool for computation of the light propagation from the SLM to the camera in this setup (4.3). We know what is the initial intensity distribution of the laser beam because it depends only on our source and we can easily measure it. The target intensity distribution on the camera is then an user input that can be arbitrarily chosen.

The last missing ingredient is a phase mask that is uploaded on the SLM to shape the beam in accordance with the user input. For a limited number of cases, the phase mask can be derived analytically, but there is no general approach that would achieve the same for an arbitrary intensity distribution. Therefore, iterative computer algorithms are a common approach to this problem.

The algorithms are divided into two categories, each allowing for further modifications. The Gerchberg-Saxton algorithm and its extensions fall into the first category, while the other one comprises the gradient descent method and its derivatives. I decided to examine and employ the first option. In the next sections, basic principles of both approaches will be described.

5.1. Gerchberg-Saxton Algorithm

The Gerchberg-Saxton (GS) algorithm is a computational method used in the field of optics and signal processing for solving the phase retrieval problem. Phase retrieval is the process of determination of the phase of a complex-valued signal from its magnitude (intensity) measurements alone. This is a crucial problem in various applications such as X-ray imaging, astronomical imaging, adaptive optics, and binary optical design [14].

The GS algorithm was introduced independently by R. W. Gerchberg and W. O. Saxton in the late 1970s [15]. It is particularly useful in situations where the phase information of a signal is not accessible, and only the intensity (magnitude) of the signal is available. The algorithm iteratively refines an estimate of the phase by alternating between the real and Fourier domains.

Individual steps of the GS algorithm are described in the following list and illustrated by the diagram in Figure 5.1:

1. Initialization: begin with an initial estimate of the phase. This could be a random guess or be based on some prior knowledge, if available.

2. Fourier Transform: take the Fourier transform of the initial laser field ① with the imprinted estimated phase to obtain an estimate of the complex-valued signal ② in the Fourier domain.

3. Target Magnitude Constraint: replace the magnitude of this complex-valued signal with the magnitude that is required to obtain ③.

4. Inverse Fourier Transform: perform an inverse Fourier transform to obtain an updated estimate of the object in the real space. Phase ④ can be separated from this field.

5. Initial Magnitude Constraint: Replace the magnitude with the known intensity of initial laser beam ①.

6. Iteration: Repeat steps 2-5 until convergence criteria are met. This could be a pre-defined number of iterations, a tolerance level for change between iterations, or other convergence criteria.
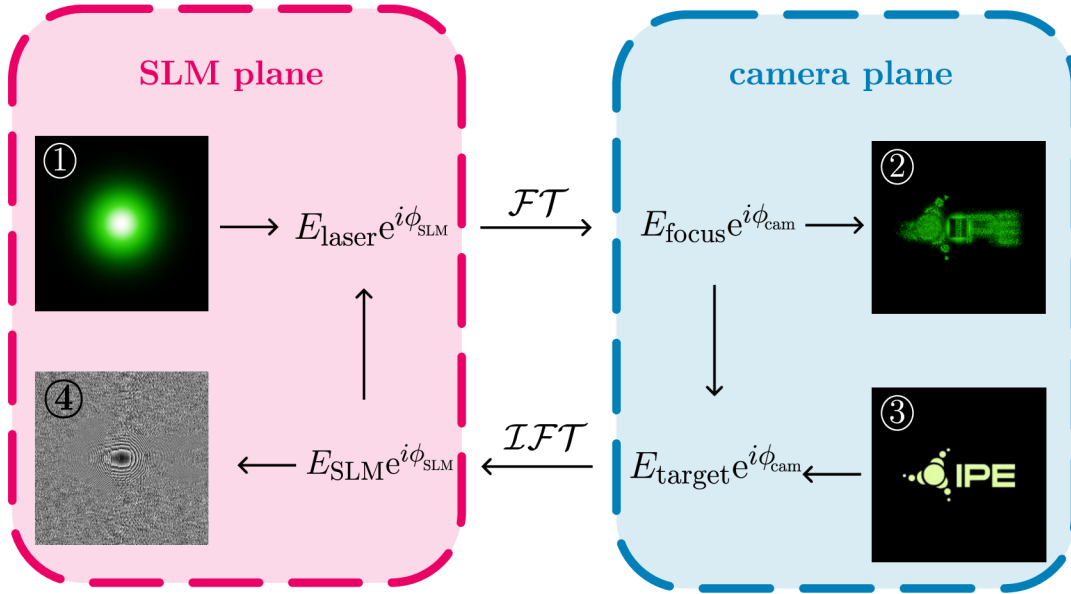

Figure 5.1: Diagram of the Gerchberg-Saxton algorithm - computing the phase mask ④ for shaping laser ① into the instensity distribution ③ on the output ②.

The key idea behind the GS algorithm is the alternating update between the real and Fourier domains while maintaining consistency with fields amplitudes in both domains. Through iterative refinement, the algorithm aims to converge to a solution where the estimated phase accurately reconstructs the target field intensity.

5.2. Gradient Descent Method

Gradient descent is a fundamental optimization algorithm widely used in various fields such as machine learning, deep learning, and numerical optimization. Its primary goal is to minimize a given objective function, often referred to as the cost or loss function, by iteratively adjusting the parameters in the direction of the steepest descent, as defined by the negative of the gradient.

In the context of optics, gradient descent is a general optimization technique that can be applied to a wide range of problems, including the phase retrieval. It can incorporate various modifications and constraints by defining an appropriate cost function [16].

5.3. Development of the GS Algorithm

I chose to code and further develop the GS algorithm because of its simplicity, low computational time, and high number of available information sources. Gradient descent

```python
1  import numpy as np
2  import matplotlib.pyplot as plt
3
4  size = (512,512) # number of pixels of the complex optical field
5
6  # function for creating intial Gaussian beam
7  def Gaussian_beam(size):
8      R = size[0]/7 # 7 controls beams width
9      x, y = np.indices(size) - (size[0] - 1) / 2
10     gaussian_array = np.exp(-0.5 * ((x / R)**2 + (y / R)**2))
11     return gaussian_array
12
13 # example of the target field - paraboloid beam
14 def paraboloid_beam(size):
15     R = size[0]/6 # 6 controls beams width
16     x, y = np.indices(size) - (size[0] - 1) / 2
17     r = np.sqrt(x**2 + y**2)
18     paraboloid_beam = np.where(r<R, (r/R)**2+0.2, 0)
19     return paraboloid_beam
20
21 E_SLM = Gaussian_beam(size) # electric field of the laser
22 I_goal = paraboloid_beam(size) # target intensity of the electric field
23 phase = np.random.randint(0,2*np.pi,size) # initial phase guess
24
25 # algorithm core
26 iterations = 35
27 for i in range(iterations):
28     # field at SLM plane
29     SLM_field = np.abs(E_SLM) * np.exp(1j*phase)
30     # FFT propagation from the SLM to the camera
31     camera_field = np.fft.fftshift(np.fft.fft2(np.fft.fftshift(SLM_field)))
32     phase = np.angle(camera_field)
33     # replacing field amplitude with targer amplitude
34     G = np.sqrt(I_goal) * np.exp(1j*phase)
35     # IFFT propagation from the camera to the SLM
36     SLM_field = np.fft.ifftshift(np.fft.ifft2(np.fft.ifftshift(G)))
37     # storing current phase in the SLM plane
38     phase = np.angle(SLM_field)
39
40 # propagating initial field with imprinted computed phase from SLM to camera
41 camera_field = np.fft.fftshift(np.fft.fft2(np.fft.fftshift(E_SLM * np.exp(1j*phase))))
42
43 # plotting results
44 plt.rcParams['font.size'] = 15
45 fig, axs = plt.subplots(1, 3, figsize=(12,4))
46 axs[0].imshow(E_SLM**2)
47 axs[0].set_title("Initial laser beam intensity")
48 axs[1].imshow(phase, cmap="gray")
49 axs[1].set_title("Computed phase mask")
50 axs[2].imshow(np.abs(camera_field)**2)
51 axs[2].set_title("Intensity at camera")
52 plt.tight_layout()
53 plt.show()
```

Listing 5.1: Example of a simple Python code for running the GS algorithm - used versions of software and packages: Python 3.11.5, numpy 1.24.3, matplotlib 3.7.2.
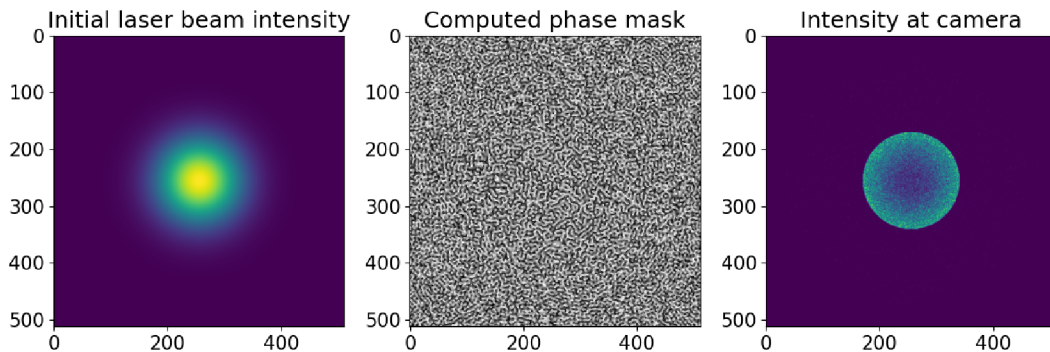
Figure 5.2: Plotted results of the simple GS algorithm from Listing 5.1 (it would be appropriate to add axes labels but the intention was to keep the code as short as possible).

methods would offer more degrees of freedom thanks to the possibility of tailoring the cost function but the GS algorithm can also be modified and upgraded which was demonstrated in [17] and other studies.

The first step was to code the conventional version of the GS algorithm. Listing 5.1 provides an example of a simple Python code, the purpose of which is to demonstrate an implementation of the algorithm and also offer a clear and accessible code for anyone interested in exploring the GS algorithm themselves. The algorithm has 2 main inputs. The variable `E_SLM` on line 21 contains the optical field represented by `2D numpy ndarray` of complex numbers and the variable `I_goal` on line 22 contains the desired intensity distribution on the output represented by a `2D numpy ndarray` of positive real numbers.

This particular example takes a Gaussian beam created by the function `Gaussian_beam` as input and its goal is to find a phase mask that can shape this beam into the intensity distribition given by the function `paraboloid_beam`. After given number of iterations, algorithm will stop and the phase distribution from the SLMs plane will be stored in the variable `phase`. Phase ϕ can be added to the original field E_0 by the following operation: $E_1 = E_0 e^{i\phi}$, where i is the imaginary unit. It is clear that $\left| e^{i\phi} \right| = 1$, which means that this operation does not change the intensity I of the complex field since $I \propto |E|^2$. At the end, this code shows a figure similar to the Figure 5.2 with the initial field intensity (left), the computed phase mask (middle) and the shaped beam after the simulated propagation by the FFT (right).

By defining an alternative form of variables `E_SLM` and `I_goal`, it is possible to compute phase masks for various cases of beam shaping.

5.3.1. Improvement of the GS Algorithm

The conventional Gerchberg-Saxton algorithm (described in Section 5.1 and implemented utilizing the code from Listing 5.1) often converges to some local optimum which results in an intensity distribution that is similar in shape to the desired one, but it also contains some additional noise, as shown in Figure 5.3. Hence, there is still a lot of room for improvements.

The next step is to make adjustments to the algorithm that would reduce noise and improve the overall performance. This raises the need for a metric that would evaluate
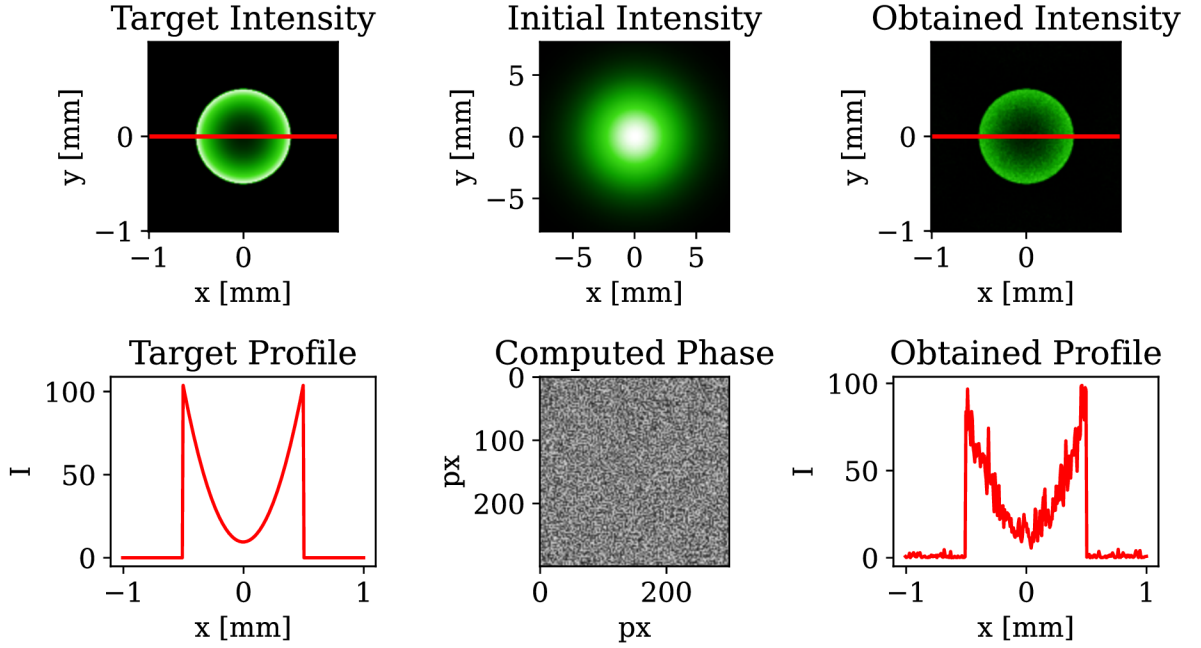
results (35 iterations)

Figure 5.3: Results of the conventional GS algorithm with a random initial phase distribution.

how well does the result match the expectation. Root mean square deviation is a suitable option. It can be mathematically written down as

$$RMSD = \frac{1}{N} \sum_{n=0}^{N-1} \left[I_{\text{sim},n} - I_{\text{tar},n} \right]^2, \qquad (5.1)$$

where the sum is taken over the area of interest (i.e. only the area covered by the beam without a background), N is the number of pixels in the area of interest and n is the index of a pixel in this area, I_{sim} denotes the intensity of the beam that is computed by the numerical propagation of the initial beam with the imprinted phase mask, and I_{tar} is the target intensity.

Initial Phase Guess

The initial phase guess can play a role in algorithm's convergence and it would be advantageous to set it based on some prior knowledge instead of setting it randomly. It turns out that a good choice of the initial phase guess is

$$\varphi_i = m_x x + m_y y + 4R \left[\gamma x^2 + (1-\gamma)y^2 \right], \qquad (5.2)$$

where $m_x x$ and $m_y y$ are linear terms that diffract the laser away from the optical axis, R determines the curvature of the quadratic part and γ is related to the aspect ratio [18]. Figure 5.4 shows results of the algorithm from Listing 5.1 for 2 different initial phase guesses. The first row is for a random initial phase distribution which results in a very chaotic computed phase mask that contains a lot of phase singularities (points where

23

the exact value of the phase is not defined beacause all values in a range from 0 to 2π can be found in the immediate neighbourhood of the singularity). Introduction of the phase singularities to the field creates unwanted optical vortices that are characteristic by having zero intensity in the center and they manifest themselves as a little dark spots in the resulting intensity. Book [19] provides many interesting information about vortices, including a proof that the overall topological charge is conserved troughout the propagation. That means that the number of vortices in the object plane will differ only slightly from the image plane (slight variation may be caused by annihilation of oppositely oriented vortices). It is not always possible to make a good enough phase guess to eliminate majority of vortices and further improvements are needed, which is described in [18].
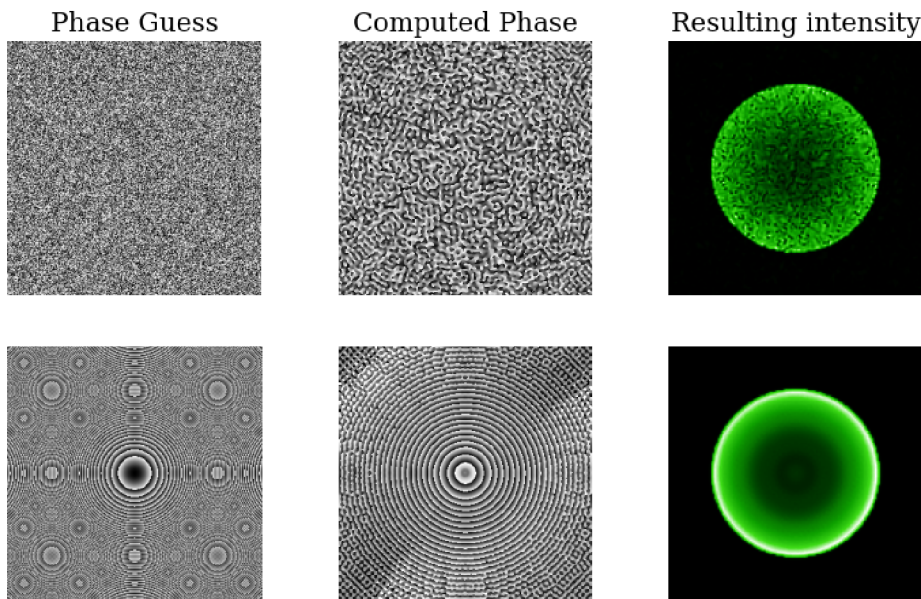

Figure 5.4: Impact of a phase guess on the final results, the first row - the algorithm starts with a randomly generated phase, the second row - the algorithm starts with a phase generated according to (5.2) - the phase guess is shown after the modulo 2π operation which causes moiré patterns.

Adaptive Weighting

The next possible improvement is to allow the target intensity to be adaptively adjusted to enforce larger changes in areas that are far from the desired result [20]. In the conventional GS algorithm, the field amplitude in the camera plane is replaced with the target amplitude (③ in Figure 5.1). In the weighted GS algorithm (WGS), the same procedure is done in the first iteration but in the following iterations, weights $\mathcal{W}$ are used instead. These weights are computed in the following way:

$$\mathcal{W}_{i+1} = \mathcal{W}_i \left(\frac{\mathcal{T}_i}{\mathcal{F}_i} \right)^p,\qquad(5.3)$$

where i is the index of the current iteration, $\mathcal{T}$ is the target field amplitude, $\mathcal{F}$ is the current field amplitude (② in Figure 5.1), and p is a parameter that controls the speed

of change of $\mathcal{W}$. For $p = 0$, the WGS becomes GS again, because $\mathcal{W}$ will never change and it will remain identical to the target amplitude. It is crucial to find the optimal value of the parameter p to ensure that the algorithm converges to a state with a low $RMSD$ (5.1). A numerical experiment was set up, where algorithms with different values of p were compared. In each iteration, $RMSD$ was computed. Results of this experiment are plotted in Figure 5.5.
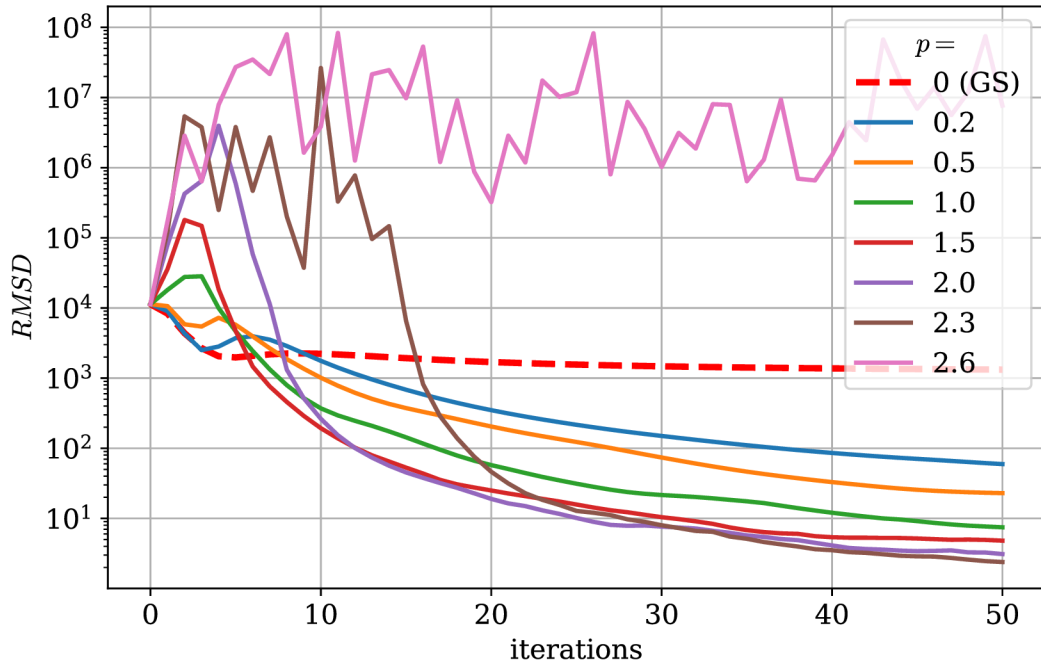

Figure 5.5: Numerical experiment aimed at finding the optimum valie of the parameter p in the WGS algorithm.

The optimal value of p was found to be approximately 2. Algorithms with a higher p were unstable and sometimes did not even converge to a better result than the one calculated by the pure GS. Lower values of p cause the algorithm to converge to a local optimum which is far from the global optimum. The conventional GS algorithm has a property, that the error decreases with each iteration, which can be proven [15]. The WGS does not possess this property but it is still better in the long run.

Mixed Region Amplitude Freedom

Considering that the electron beams in the light-controlled electron microscope (described in section 2.2.2) will interact only within a small area covered by the reshaped laser beam, it might be beneficial not to enforce the intensity distribution in the entire image plane. In other words, the intensity in regions near borders could be left to develop randomly which would provide more freedom for the algorithm to adjust the phase mask. An algorithm with this modification is often referred to as Mixed Region Amplitude Freedom (MRAF). A way to transfer this principle into the algorithm is to replace the step 3 step in the GS algorithm described in 5.1, where the field amplitude is normally replaced

by the target amplitude while the phase is kept the same. Modification from GS to MRAF is done by this transition:

$$E_{\text{target}}\, e^{i\phi} \longrightarrow \left(|ME_{\text{target}}|_{\text{SR}} + |(1-M)E_{\text{actual}}|_{\text{NR}} \right) e^{i\phi}, \qquad (5.4)$$

where M is a mixing parameter ($0 < M < 1$), SR means signal region (the area that we are interested in), NR represents the noise region (area of an uncontrolled development), and $|E_{\text{actual}}|$ is an amplitude of the field that was created by propagation from the SLM plane[17].

The MRAF algorithm did not show significant improvements in conducted numerical tests. In most cases, the change was almost imperceptible. Some effects were achieved only with a very attenuated weighting ($p = 0.1$) in an experiment whose results are shown in Figure 5.6. The goal was to shape a Gaussian beam into a logo of the Institute of Physical Engineering. A very low noise level was achieved in the area around the logo at the cost of losing a large part of the laser power in the noise region around the borders of the evaluated area. This could be useful for applications like parallel micromachining or atom trapping where it is desirable to have some areas with zero laser intensity.
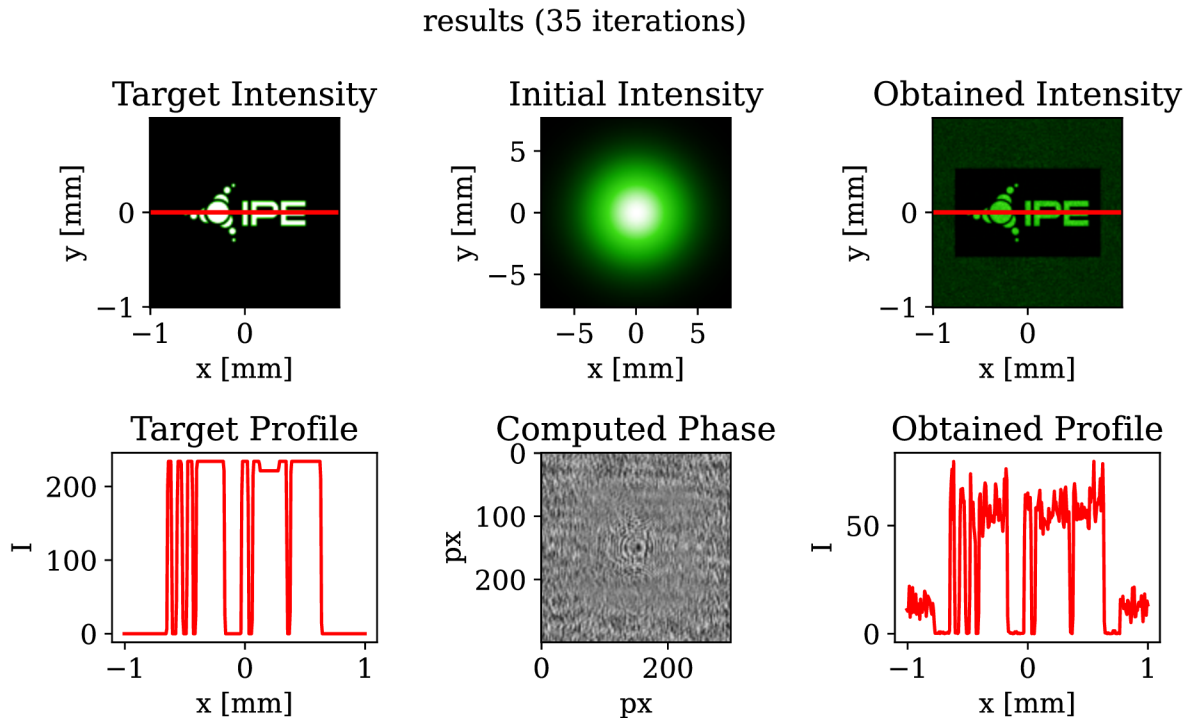

Figure 5.6: Results of the MRAF algorithm - creating logo of the Institute of Physical Engineering.

The MRAF algorithm is often used to create atom traps (for example in [17]). However, for the light-controlled electron microscope (section 2.2.2), which is the principal motivation for this work, MRAF does not bring any significant improvement.

Conclusion from the Algorithm's Development

Three major potential upgrades of the GS algorithm were tested. It was shown that an appropriate choice of the initial phase guess can the reduce the number of optical

vortices resulting in fewer unwanted dark spots in the obtained intensity distribution. The most significant improvement was attained by introducing adaptive weighting which allows the algorithm to converge to a much better local optimum. Example of a WGS's performance is in Figure 5.7, where the obtained intensity profile is almost identical to the target intensity distribution.
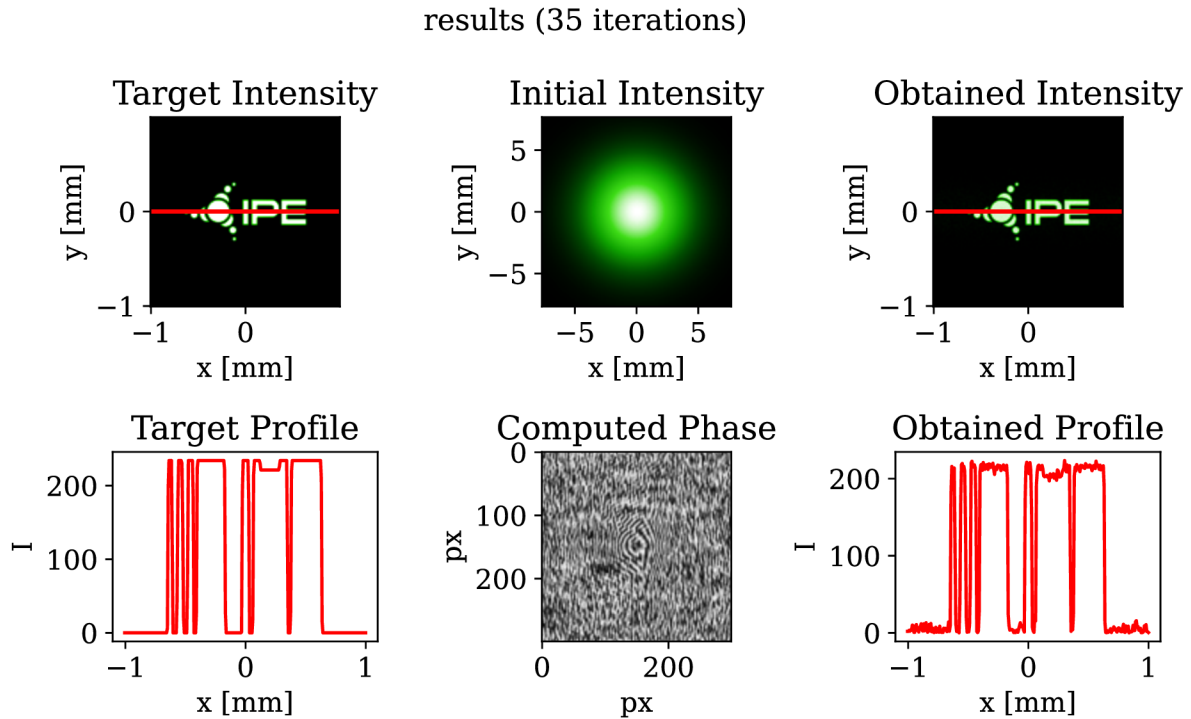

Figure 5.7: Results of the WGS algorithm with a parabolic initial phase guess.

The MRAF version of the GS algorithm was also tested, but no improvements relevant to our project of shaping electron beams, were observed.

The development of the algorithm was inspired by sources cited in this section and also by [21], where an open Python code is provided.

The next step was to verify the results of the developed algorithm. The first approach involved simulating the laser propagation using the angular spectrum method, while the second approach was conducting an actual experiment in a laboratory.

5.3.2. Propagation with the ASM

Propagation of the laser was simulated using a single Fourier transform in the algorithm so far. The validity of this propagation method was derived in chapter 4. This method is based on the Fresnel diffraction integral 4.7, which is only valid under the paraxial approximation. Another numerical method for simulating light propagation is the Angular Spectrum Method (ASM). This general-purpose method can be used for free propagation between any two planes.

The core idea behind ASM is the fact that the complex field in a plane can be Fourier-analyzed and the spatial Fourier components can be recognized as plane waves travelling in different directions away from the plane (chapter 3.10. in [22]). The field amplitude at any other point can be calculated by summing up the contributions from all these plane

waves, taking taking into account their phase accumulated during their propagation. From a complex field represented by $U(x, y, 0)$, a field in another plane can be calculated as

$$U(x, y, z) \propto \mathcal{F}^{-1}\Big\{\mathcal{F}\{U(x, y, 0)\}\, \mathrm{e}^{ik_z z}\Big\}, \tag{5.5}$$

where $\mathcal{F}$ denotes the Fourier transform, $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, k_z is the component of the wave vector along the optical axis, and z marks the position of the plane where the field is computed. Notably, z can be negative, indicating backpropagation. The ASM allows for the inclusion of properties of a real optical system, such as a finite aperture size.

Results of a simulation corresponding to the setup are depicted in Figure 3.7. In the first step, a computed phase mask was imprinted to a Gaussian beam which was then propagated to the lens. The effect of the lens on the wavefront was included by another phase change according to (4.11). In the last step, the beam was propagated to the focal plane of the lens. Results of this numerical experiment are shown in Figure 5.8.
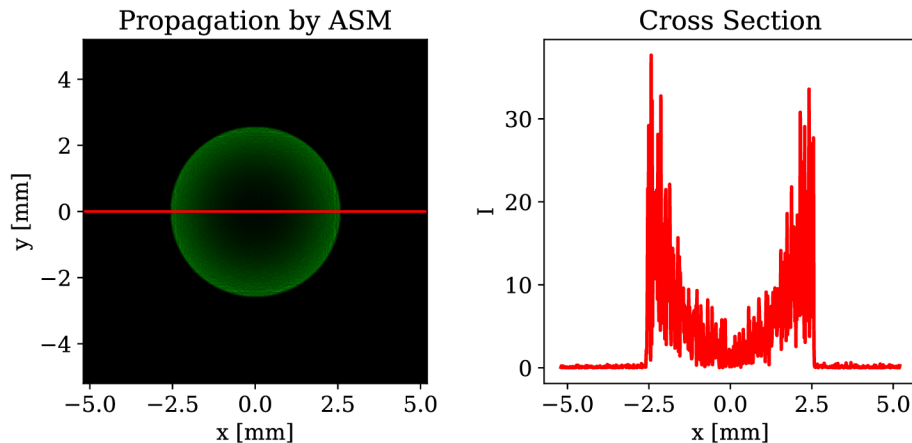

Figure 5.8: Results of the ASM propagation - the electric field intensity in the focal plane of the lens.

Propagation with the ASM showed that the the phase mask calculated by the WGS algorithm indeed produces in the focal plane of the lens a beam with a parabolic radial profile, but it contains a very high level of noise. That means that the functionality of the WGS algorithm should be tested in a laboratory with a real equipment before making any conclusions about its performance.

Note about ASM and FFT Propagation

So far, two methods of a light propagation have been employed. One challenge I have encountered is the size of the evaluated field represented by a data matrix. While the ASM uses the same set of spatial coordinates for both the input and output plane, the propagation based on the Fresnel diffraction utilizing Fast Fourier Transform does not: the field distribution created in the focal plane of a lens corresponds to the angular spectrum of plane waves that travel at different angles relative to the optical axis. Therefore, to transform the set of angles into a set of real space coordinates in the focal plane of the lens, parameters like the focal distance, light wavelength, and discretization step within the SLM plane must be taken into account.

The highest value of a spatial frequency present in a Fourier representation of an optical field is

$$k_{x,\max} = \frac{2\pi}{d} = \frac{2\pi N}{D}, \tag{5.6}$$

where D is the width of the field, N denotes a number of pixels along one side of the field and d is the size of one pixel of the field.

From Figure 5.9, it is clear that we can calculate the maximal spatial position as

$$x_{\max} = \theta_{x,\max} f = \frac{k_{x,\max}}{k} f = k_{x,\max} \frac{\lambda f}{2\pi} = \frac{2\pi}{d} \frac{\lambda f}{2\pi} = \frac{\lambda f}{d} = \frac{\lambda f N}{D}. \tag{5.7}$$

Thus, we can see that the FFT propagation produces fields of different sizes depending on wavelength λ and focal length f. Therefore, with the same phase mask on the SLM, the size of the image can be changed by choosing a lens with a different focal length.
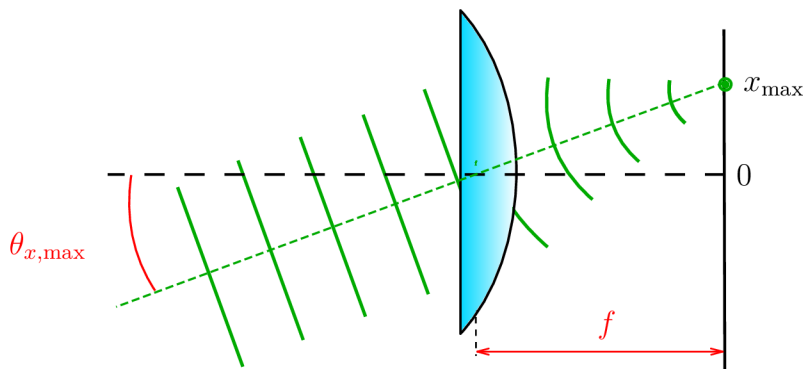

Figure 5.9: A sketch of a plane wave attributable to the highest spatial frequency present in the field formed within the focal plane of a lens.

5.4. Experimental Verification of Algorithm's Performance

The main results presented so far involve the algorithm developed for the purpose of beam shaping employing the experimental setup detailed in section 3.2.3. Potential shortcomings of the algorithm include inaccurate modeling of light propagation, neglecting the actual properties of the setup that affect the output, or other unforeseen factors. To evaluate its performance in real conditions, we decided to test it also experimentally.

5.4.1. Setup Realization

The ultimate goal of this work is to deliver shaped beams for a light-controlled electron microscope described in 2.1 where an ultrafast high-power laser is used. However, for the sake of practicality and safety this experiment was carried out with a conventional green laboratory laser. A photo of the actual realisation of this setup is in Figure 5.10.

The 515 nm laser was operated at approximately 50 mW power. The choice of laser affected also the requirements placed on the SLM display, so the PLUTO-2.1-NIR-149 was replaced with PLUTO-2.1-NIR-015 model while keeping the original driver unit. Results
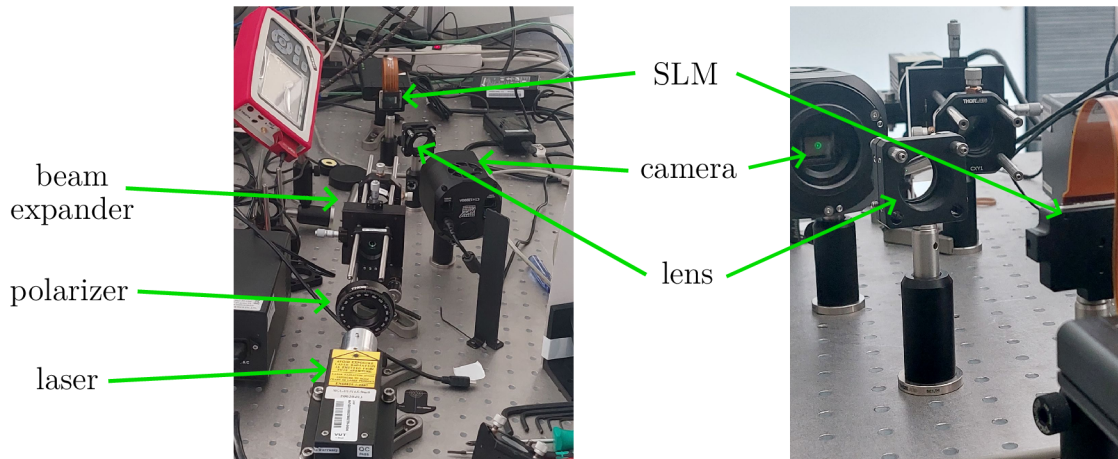
Figure 5.10: Experimental setup with the SLM from two different angles of view.

were captured on the C1+12000A CMOS camera by Moravian Instruments equipped with the Sony IMX304 chip.

It is important to keep the incident angle of the laser hitting the SLM small because higher incident angles may cause unexpected behaviour of the SLM. Study [23] shows that incident angle deviations less than 10° do not modify substantially its modulation properties.

5.4.2. Measurement

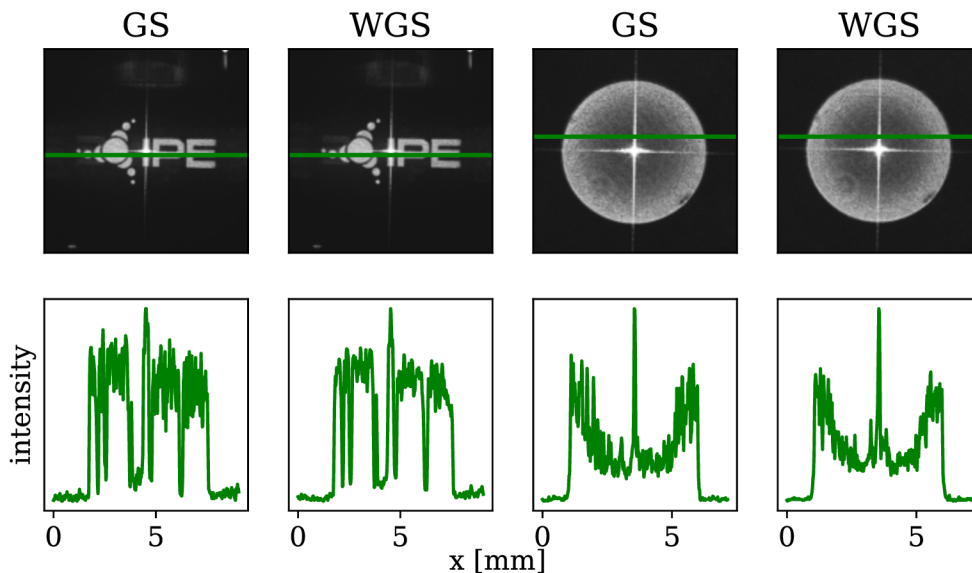First results of the experiment are shown in Figure 5.11.

Figure 5.11: Measured data - the first row shows grayscale images directly from the camera, the second row shows intensity cross sections along lines marked in the camera images.

The logo of the Institute of Physical Engineering and a beam with a parabolic radial profile were created. Phase masks were computed using both the conventional GS algorithm and the WGS algorithm. Intended shapes were indeed formed with a noise level higher than predicted by the algorithm that uses the FFT propagation method (example in Figure 5.7). Neither the results match the prediction made by the ASM propagation method (Figure. 5.8).

There is also an unwanted bright intensity peak right at the center of the image, which can be attributed to the reflection from the the outer glass coverplate of the SLM display. This reflected wave is not modified by the liquid crystals and therefore remains as a plane wave that is subsequently focused by the lens into a point in the observed focal plane. A possible solution for this problem is to deflect the image away from the optical axis by adding a phase wedge to the phase mask (Figure 5.12).
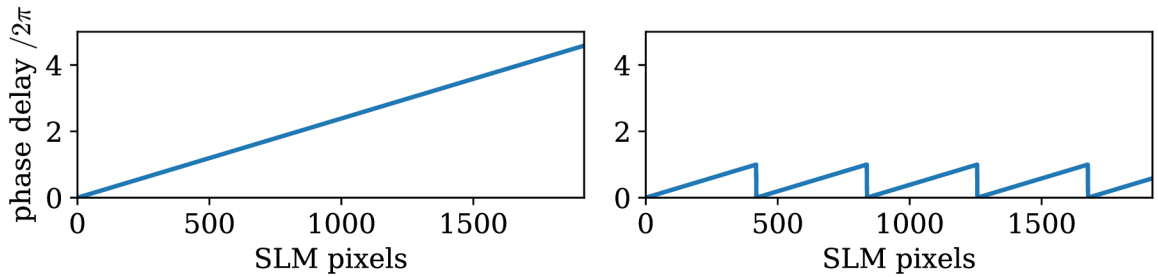

Figure 5.12: Left plot - a wedge shaped phase delay function for an image deflection, right plot - the same function after a modulo 2π operation that can be displayed with the SLM while preserving the effect of the original function (often referred to as a phase grating).

Results of the experiment with the applied phase grating are in Figure 5.13. Images were spatially separated from the bright central spots which are now easily removable by a conveniently placed screen.
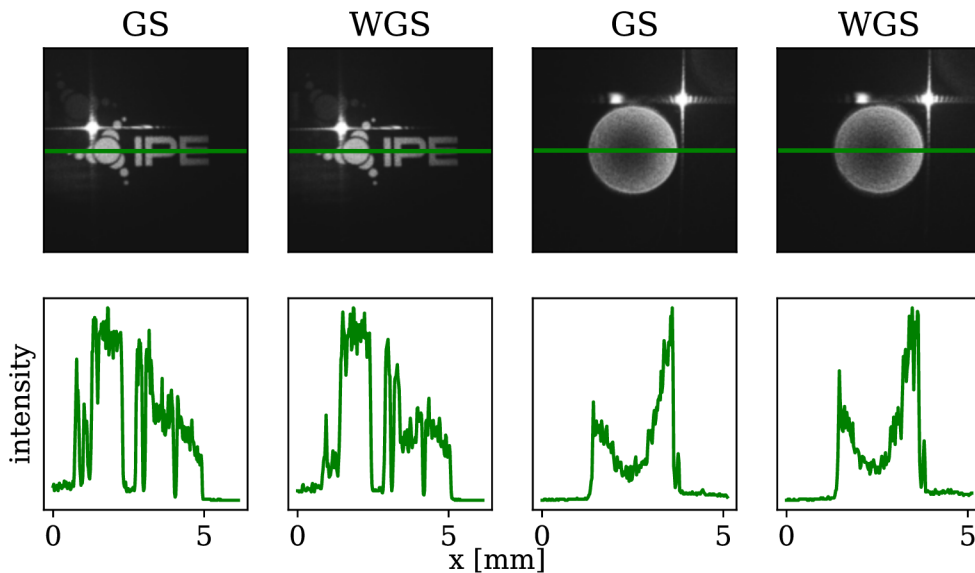

Figure 5.13: Images captured by the camera after a phase grating was added to the phase mask to spatially separate the engineered image from the parasitic bright spot.

However, the successful filtration of the bright spot was accompanied by another undesirable effect, namely gradual fading of the intensity as we get further and further away from the optical axis. This is clearly visible in Figure 5.13 where the intensity peaks in the vicinity of the bright spot. This could be compensated by intentionally creating a reverse defect in the target intensity profile that is given as input to the algorithm. Results of the measurement with with the incorporated compensation are shown in in Figure 5.14. In this case, the initial phase guess impacts the result more than the choice of algorithm.
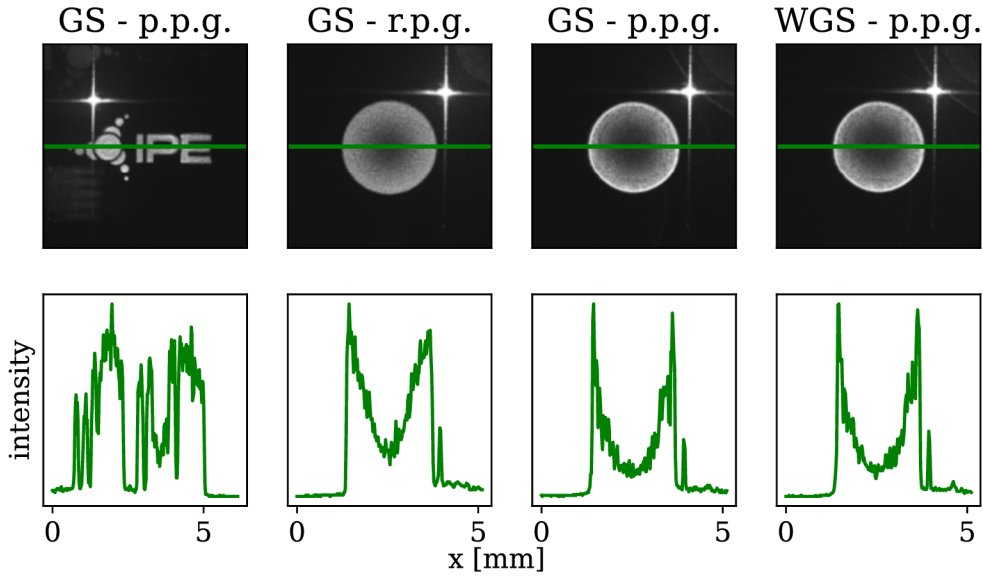

Figure 5.14: Measured data captured on the camera, a phase grating was added to the phase mask to spatially separate the image from the bright spot and the target intensity profile was adjusted to improve overall uniformity., p.p.g. - parabolic phase guess, r.p.g. - random phase guess.

5.4.3. Data analysis

The quality of the experimental results has been qualitatively discussed in the previous sections but has not been quantified yet. To fill this gap, we decided to use two metrics: the first metric will evaluate noise (local deviations from a smooth shape), and the second metric will assess the difference between the underlying smooth shape and the intended profile (e.g., a parabola).

Noise evaluation can be achieved by smoothing the measured data and computing the difference between the measured data and the smoothed data. To obtain a single comparable value, the variance (denoted as σ_1^2) of the noise is calculated. This calculation is performed only within a specific region of interest (RoI) to avoid bias near sharp edges caused by the smoothing process. An example of this metric's application is shown in Figure 5.15, where data from Figure 5.14 are evaluated, comparing the GS and WGS algorithms. The smoothing bias near edges is also notable in sectional plots. The WGS algorithm performed slightly better in terms of the noise level, having $\sigma_1^2 = 0.0019$ while the GS achieved $\sigma_1^2 = 0.0028$.
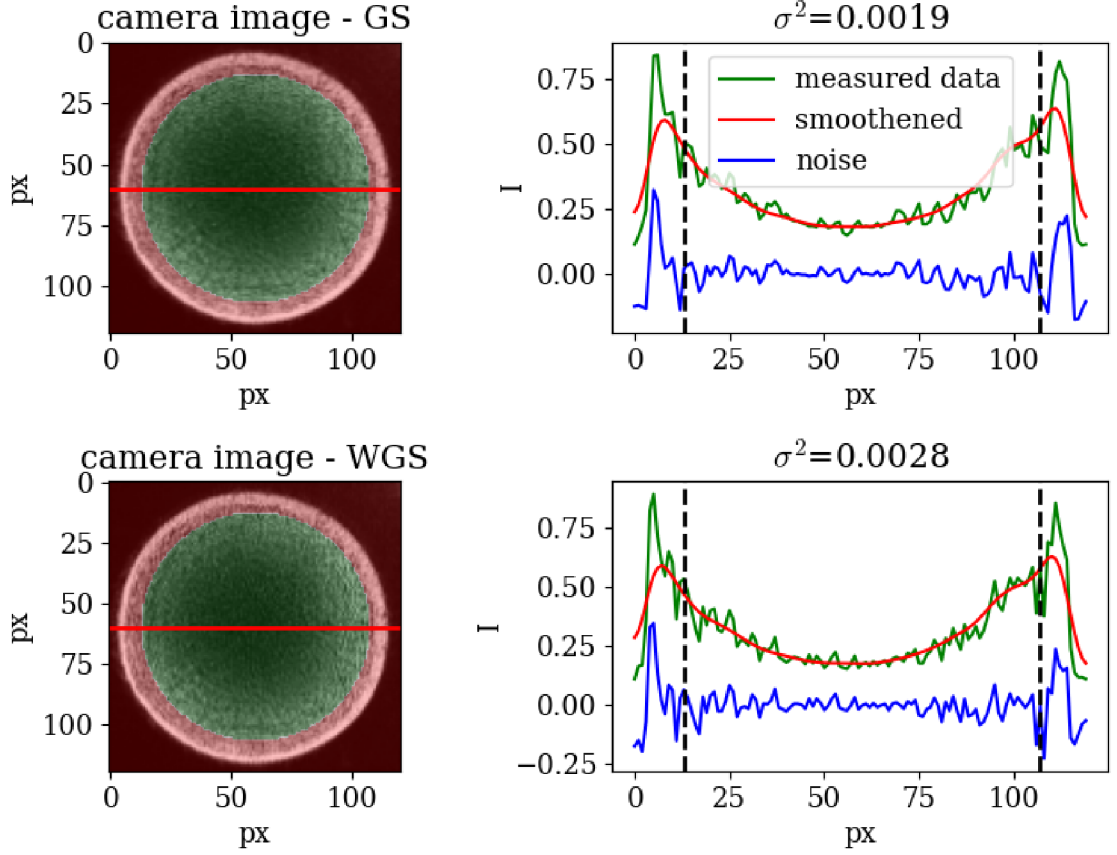
Figure 5.15: Noise quantification - comparison of the GS and the WGS algorithms, green areas in the camera images indicate the Region of Interest (RoI) where the noise is evaluated, dashed lines in cross sections on the right mark the borders of RoI.

The second metric requires data fitting which can be easily done for a simple shape like paraboloid with its analytical prescription

$$z = a\left[(x - x_0)^2 + (y - y_0)^2\right] + b. \tag{5.8}$$

An example of this fit is shown in Figure 5.16. The beam shaping accuracy was evaluated by examining the value of the fitting parameter a from Equation (5.8) and its standard deviation σ_2, which can be retrieved from the covariance matrix of the fit. We express a as $a_{\text{fit}} \pm \sigma_2$. The data visualized in Figure 5.14 were analyzed.

The GS algorithm with a random phase guess achieved $a = (1.21 \pm 0.0060) \times 10^{-4}$ m^{-1}, the GS with a parabolic initial phase guess had $a = (1.86 \pm 0.0089) \times 10^{-4}$ m^{-1}, and the WGS with a parabolic initial phase guess resulted in $a = (1.86 \pm 0.0075) \times 10^{-4}$ m^{-1}. It is evident that the initial phase guess significantly affects the value of a, more than the choice of algorithm.

Surprisingly, the GS with a random phase guess exhibited the lowest standard deviation, but for a smaller value of a, i.e. the parameter that controls the steepness of the created parabola. The steepness of the parabola determines the strength of the force acting on electrons, a crucial factor for our intended application (light-controlled electron microscope). So, in this situation, a higher value of a is more important than a low standard variation.
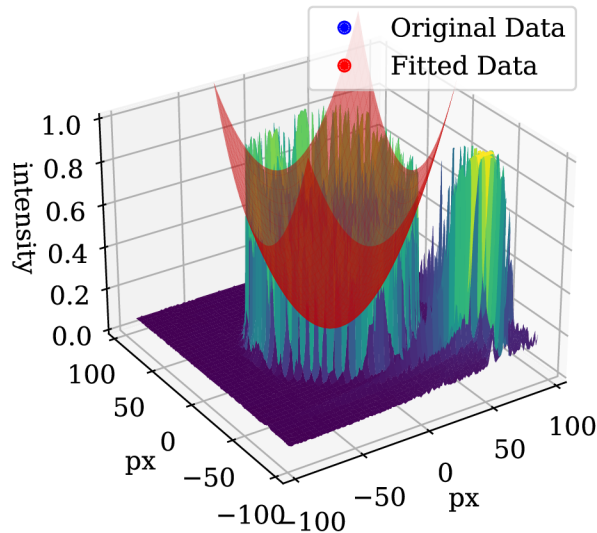
Figure 5.16: Fit of the measured data.

It is important to note that the statistical evaluation of the algorithm's performance was conducted on a very limited dataset. Due to time constraints, it was not possible to conduct a more extensive set of measurements which would be needed to obtain more reliable results.

5.4.4. Imperfections of the Setup

The developed algorithm can find a phase mask that can be imprinted on the incoming beam to produce an almost perfectly shaped beam with the desired intensity distribution at the computed output. However, the experimental results do not achieve this high quality, likely due to the physical model of the laser propagation derived under the assumption of a perfect setup, which may differ from the reality in the following characteristics:

- Lenses in the real setup have one-inch round aperutures, which is not reflected in the FFT propagation.

- Fill factor of the SLM display is 93% in reality (see parameters of the used SLM in Table 3.1) because there are walls separating individual pixels, while the simulation considers an idealized situation where pixels cover the whole display continuously and the phase mask is therefore imprinted perfectly.

- The simulation models each pixel as an independent unit capable of uniformly imprinting the required phase across its area. However, in practical applications, each pixel is affected by the electric fields of adjacent pixels, leading to field leakage into the surrounding area. This interaction can cause variations in the phase shift across a pixel, depending on the differences between it and its neighboring pixels.

- As study [24] shows, phase shift of a pixel suffesr from flickering, which means that the phase mask changes a little as time passes. The used physical model, however, considers it being stationary.

- The laser beam incident on the SLM is neither purely Gaussian nor a perfect plane wave. These deviations from ideality are omitted in the simulation.

- The coherence of the laser is also a variable that may be hard to include in the simulation because it is also dependent on the actually used power of the laser.

Given the numerous phenomena that are challenging to simulate, incorporating feedback from a camera might be beneficial. This feedback could help in compensating for the aforementioned sources of inaccuracy. Implementing this feedback mechanism is a planned step for future work.

6. Conclusion

Beam shaping plays a crucial role in various fields of physics, offering the ability to manipulate light in ways that enhance experimental precision and versatility. In the context of this thesis, the main focus was on delivering on-demand laser beams for a light-controlled electron microscope. This objective guided my choice of laser shapes, focusing on those with smoothly varying spatial intensity distributions. The use of a Spatial Light Modulator (SLM) for beam shaping introduces significant improvements in the control and functionality of the electron beam.

Throughout this thesis, the fundamental theories of beam shaping have been explored, providing a solid theoretical background for understanding the principles behind this technology. In addition to the theoretical aspects, a review of existing algorithms for phase retrieval has been provided. Two groups of frequently used algorithm were identified. The GS algorithm was selected over the family of gradient descent algorithms because of its simplicity, low computational time and availability of research papers about GS algorithm and its modifications.

This was followed by the development of my own weighted algorithm, which included modifications found in available research papers. The performance of the algorithm was tested and compared against the conventional GS algorithm. The first test was based on the Fresnel diffraction and propagation via Fourier transform and it predicted that the new algorithm should give almost perfect results. Propagation employing the angular spectrum method (ASM) resulted, on the other hand, in a very noisy image, even though the same phase mask was used. A laboratory experiment was set up in order to decide what is the real performance of the algorithm. The results obtained from this experiment indicate that the weighted GS algorithm is effective, although there is a room for an improvement. A noise was present but it was lower than predicted by the ASM. It is difficult to capture all the aspects of the real setup in the simulation of laser propagation, which is the reason why results of the experiment differ from the measured data.

The GS algorithm is frequently used in the field of atom trapping, where binary intensity profiles are required instead of smoothly varying ones. This may explain why some of the algorithms proposed in available research papers do not meet the requirements for our intended application. The algorithm developed in the framework of this work was also tested for shaping a beam into the logo of the Institute of Physical Engineering, which is a binary mask, and the results were reasonably good.

Consequently, further development is planned, particularly integration of a camera feedback into the iterative algorithm. This upgrade is expected to refine the beam shaping process by providing real-time adjustments based on the feedback, leading to more accurate and reliable results. The feedback may be able to address all imperfections in the setup at once instead of resolving every individual cause of inaccuracy one by one. A switch to a gradient descent algorithms will also be considered. Phase vortices and their elimination is also an interesting topic that should be further explored.

Bibliography

[1] Guodong Zhang, Razvan Stoian, Wei Zhao, and Guanghua Cheng. Femtosecond laser bessel beam welding of transparent to non-transparent materials with large focal-position tolerant zone. *Optics Express*, 26(2), 2018.

[2] R. M. W. van Bijnen, C. Ravensbergen, D. J. Bakker, G. J. Dijk, S. J. J. M. F. Kokkelmans, and E. J. D. Vredenbregt. Patterned rydberg excitation and ionization with a spatial light modulator. *New Journal of Physics*, 17(2), 2015-02-01.

[3] M. Silvennoinen, J. Kaakkunen, K. Paivasaari, and P. Vahimaa. Parallel microstructuring using femtosecond laser and spatial light modulator. *Physics Procedia*, 41:693–697, 2013.

[4] Ethan Schonbrun, Rafael Piestun, Pamela Jordan, Jon Cooper, Kurt D. Wulff, Johannes Courtial, and Miles Padgett. 3d interferometric optical tweezers using a single spatial light modulator. *Optics Express*, 13(10), 2005.

[5] Kaifei Peng, Xueju Shen, and Fuyu Huang. Vortex beam generation method based on spatial light modulator. *IOP Conference Series: Earth and Environmental Science*, 440(4), 2020-02-01.

[6] L. D. Landau and E. M. Lifshitz. *Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics)*. Butterworth-Heinemann, 3 edition, January 1976.

[7] Michael Fowler. Graduate classical mechanics. available at https://phys.libretexts.org/Bookshelves/Classical_Mechanics/Graduate_Classical_Mechanics_(Fowler).

[8] Marius Constantin Chirita Mihaila, Philipp Weber, Matthias Schneller, Lucas Grandits, Stefan Nimmrichter, and Thomas Juffmann. Transverse electron-beam shaping with light. *Physical Review X*, 12(3), 2022.

[9] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics*. Wiley Series in Pure and Applied Optics. Wiley, 2019.

[10] Wikipedia Contributors. Liquid crystal.

[11] Haining Yang and D P Chu. Phase flicker in liquid crystal on silicon devices. *Journal of Physics: Photonics*, 2(3):3, 2020-07-01.

[12] Holoeye. Pluto-2.1 lcos spatial light modulator. available at https://holoeye.com/products/spatial-light-modulators/pluto-2-1-lcos-phase-only-refl/.

[13] Jiří Komrska. *Vlnová optika*. Akademické nakladatelství CERM, Brno, 1. edition, 2004.

[14] Tieyu Zhao and Yingying Chi. Modified gerchberg–saxton (g-s) algorithm and its application. *Entropy*, 22(12), 2020.

[15] R. W. Gerchberg. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–246, 1972.

[16] Tiffany Harte, Graham D. Bruce, Jonathan Keeling, and Donatella Cassettari. Conjugate gradient minimisation approach to generating holographic traps for ultracold atoms. *Optics Express*, 22(22), 2014.

[17] Matthew Pasienski and Brian DeMarco. A high-accuracy algorithm for designing arbitrary holographic atom traps. *Optics Express*, 16(3), 2008.

[18] Paul Schroff, Arthur La Rooij, Elmar Haller, and Stefan Kuhr. Accurate holographic light potentials using pixel crosstalk modelling. *Scientific Reports*, 13(1), 2023.

[19] Paramasivam Senthilkumaran. *Singularities in Physics and Engineering*. IOP Publishing, 2018-11-01.

[20] Yang Wu, Jun Wang, Chun Chen, Chan-Juan Liu, Feng-Ming Jin, and Ni Chen. Adaptive weighted gerchberg-saxton algorithm for generation of phase-only hologram with artifacts suppression. *Optics Express*, 29(2), 2021.

[21] QPG-MIT. Slmsuite. available from: https://github.com/QPG-MIT/slmsuite.

[22] Joseph W. Goodman and Mary E. Cox. Introduction to fourier optics. *Physics Today*, 22(4):97–101, 1969-04-01.

[23] A. Lizana, N. Martín, M. Estapé, E. Fernández, I. Moreno, A. Márquez, C. Iemmi, J. Campos, and M. J. Yzuel. Influence of the incident angle in the performance of liquid crystal on silicon displays. *Optics Express*, 17(10), 2009.

[24] Haining Yang and D P Chu. Phase flicker in liquid crystal on silicon devices. *Journal of Physics: Photonics*, 2(3), 2020-07-01.