

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA
Department of Physics, Informatics and Mathematics

Dottorato di ricerca in Physics and Nanosciences

XXXVIII Cycle

Electron Beam Shaping with a Tunable Phase Plate

Candidate: Payam Habibzadeh Kavkani

Supervisor: Prof. Marco Beleggia

Co-Supervisor: Dr. Vincenzo Grillo

PhD Program Coordinator: Prof. Marco Affronte

Academic year 2024/2025

Dedicated to the people of Iran, in memory of those
who sacrificed everything for the homeland.
They will never be forgotten; their names will live on
in the heart of history.

Long live Iran.
Long live the people of Iran.
May the Lion-and-Sun flag fly high.

February 2026

تقدیم به فرزندان ایران و جان‌فدایان میهن؛
آنان که نام‌هایشان در هیچ کتابی نمی‌گنجد و تا ابد در دلِ تاریخ خواهند ماند.

زنده و پاینده باد ایران
زنده و پاینده باد مردم ایران
برافراشته باد پرچم سه‌رنگ شیر و خورشید

اسفند ۲۵۸۴

Abstract

Electron beam shaping means controlling the phase and amplitude of the electron wavefront to produce a prescribed wave function at specific planes of a transmission electron microscope (TEM). In practice, this is achieved by inserting electron-optical elements, called phase plates, into the electron beam's path to manipulate it. By placing a phase plate in the illumination system (condensers aperture planes), they offer the unique opportunity (not possible with conventional electron optics) to tune the incident probe to form structured beams, such as Bessel and vortex, for task-specific interactions. When placed in a focal plane of the imaging system, they impose tailored phase shifts on selected spatial frequencies, modifying the transfer function. For example, by adding a $\pi/2$ shift only on the direct beam, weak-phase objects appear as measurable intensity variations. There are different types of phase plates, each with their own pros and cons: thin-film phase plates (Zernike, Hilbert, Hole-free/Volta), electrostatic phase plates (Boersch, Zach, vortex generators), magnetic phase plates (Ampère), holographic diffractive masks, and Photonic phase plates.

Electrostatic phase plates are especially promising because they offer voltage-tunable and low-loss phase control with minimal scattering material in the beam path. Their operation relies on the Aharonov–Bohm effect: an electron acquires a phase proportional to the electrostatic potential it experiences, allowing a designed potential distribution to sculpt the electron wavefront. Our group is a world leader in electrostatic phase plates based on MEMS devices fabricated on silicon chips and mounted on biasing TEM holders. However, their design faces two main challenges:

(i) Limited numbers of biasing channels in commercial holders restrict the achievable potential profiles;
(ii) Complex electrode geometries are challenging to fabricate near the active region.

To overcome these limitations, this work introduces Current-Controlled Electrostatic Phase Plates, a novel concept that achieves a continuous potential gradients using simple few-channel architectures. In this design, boundary electrodes with specifically engineered shapes are fabricated from resistive doped silicon and connected through high-conductivity gold

layers. By driving a controlled current along these electrodes, a linear potential drop is established, reproducing complex boundary potential gradients with minimal independent biases. This approach enables compact, tunable, and reconfigurable devices. Using gradient-based optimization, we designed current-controlled vortex beam generators, orbital angular momentum (OAM) sorters, and electrostatic multipoles (quadrupoles, hexapoles, octupoles, and dodecapoles) for aberration compensation. The hexapole phase plate demonstrates bias-driven phase rotation, enabling alignment to any three-fold astigmatism orientation with a single device. This is a clear advantage over conventional multipole correctors, which require complex mechanical realignment. Finite-element simulations and wave-optical propagation confirm the desired phase profiles, and prototype MEMS chips have been successfully fabricated and characterized. To complement hardware development, we implemented an AI-assisted optimization and control pipeline. A deep neural network (DNN) analyzes a single focused STEM probe image to estimate the orientation of two-fold (A2) and three-fold (A3) astigmatism within milliseconds. Integrated into a Python-based graphical interface (GUI), the DNN drives a feedback loop that adjusts the phase-plate biases via a custom PCB controller, enabling automatic aberration correction without switching to conventional Ronchigram or Zemlin procedures.

In summary, this work establishes current-controlled electrostatic phase plates as a compact, low-loss platform for programmable wavefront control in TEM. Combining electrostatic design, AI-assisted optimization, and machine learning–based feedback, this work opens a path toward adaptive beam shaping and real-time aberration correction for next-generation electron microscopy.

Contents

Thesis Overview 1

I Foundations 3

1 Introduction to Electron Microscopy 4

1.1 From Magnifiers to Eyeglasses: The Birth of Optics . . . 5
1.2 The Birth of Optical Microscopy . . . 5
1.3 Theoretical Advances and the Nature of Light . . . 6
1.4 The Quantum Framework . . . 10
1.5 From Theory to Technology: The Birth of Electron Microscopy . . . 11

2 Electron Optics 14

2.1 Introduction to Electron Optics . . . 15
2.1.1 Relativistic electron kinematics . . . 15
2.1.2 Classical Description of Electron Motion . . . 17
2.1.3 Schrödinger equation for charged particles . . . 18
2.2 Electron–Specimen Interactions . . . 20
2.3 Fourier Optics . . . 22
2.4 Point Spread Function (PSF) . . . 24
2.5 Aberrations in Electron Microscopy . . . 27
2.5.1 Defocus . . . 30
2.5.2 Spherical Aberration . . . 31
2.5.3 Chromatic Aberration . . . 32
2.5.4 Twofold and Threefold Astigmatism . . . 32
2.5.5 Combined Effects and Correction . . . 33
2.6 Chapter Summary . . . 33

3 Architecture and Components of a TEM **35**
3.1 Overview of TEM Layout . . . 36
3.2 Electron microscope components . . . 37
3.2.1 Electron Gun . . . 37
3.2.2 Electron Lenses . . . 40
3.2.3 Apertures and Holders . . . 43
3.3 Transmission Electron Microscope Architecture . . . 45
3.3.1 Illumination system . . . 45
3.3.2 Objective System . . . 47
3.4 Diffraction in TEM . . . 48
3.5 Bright-Field and Dark-Field Imaging in TEM . . . 49
3.6 TEM vs. STEM . . . 50
3.7 Aberration Correctors and Stigmators . . . 53
3.7.1 Double hexapole C_s aberration correctors . . . 54
3.7.2 Quadrupole–Octopole C_s Correctors . . . 57
3.8 Aberration Detection and Measurement in TEM . . . 59
3.8.1 Diffraction-Based Aberration Diagnosis (Zemlin tableau) . . . 59
3.8.2 Real-Space Aberration Detection Using Ronchigrams . . . 59

4 Electron Beam Shaping **62**
4.1 Beam shaping using a phase plate . . . 63
4.2 Types of Electron Phase Plates and their principles . . . 65
4.2.1 Thin film phase plates . . . 65
4.2.2 Sculpted thin-film phase plates . . . 73
4.2.3 Electrostatic phase plates . . . 73
4.2.4 Magnetic Phase Plates . . . 80
4.2.5 Photonics Phase Plates . . . 82

II Concepts and Methods **85**

5 Design Philosophy and Methodology **86**
5.1 Phase plates: Why They Have Reached Their Practical Limit? . . . 87
5.2 Motivation for MEMS-based Electrostatic Phase Plates . . . 88
5.3 Current controlled electrostatic phase plate . . . 91
5.4 Sandwich chip design . . . 95

6 Numerical Simulation and Modeling **99**
6.1 COMSOL Simulations . . . 100
6.1.1 Choosing the Physics Modules and Study Type . . . 100
6.1.2 Geometry Import and Model Setup . . . 101
6.1.3 Setting Boundary Conditions . . . 102

6.1.4 Meshing and Computing . . . 103
6.1.5 Optimization and Inverse Design . . . 104
6.2 Phase-Shift Calculation: Projected Potential and Multislice Propagation . . . 106
6.3 Python-Based Simulation, Control, and Data-Analysis Tools . . . 110
6.3.1 Multislice Simulation GUI . . . 110
6.3.2 Electron Microscopy Toolkit GUI . . . 113
6.4 Machine Learning and Neural Network . . . 118

7 Designs **121**
7.1 Overview . . . 122
7.2 Multipoles . . . 123
7.2.1 Electrostatic multipole : Hyperbola phase plate . . . 124
7.2.2 Current-controlled electrostatic multipole : "Helios" phase plate . . . 129
7.3 Spiral phase plate . . . 133
7.4 OAM sorter phase plate . . . 137

III Experimental results **140**

8 Fabrication Process **141**
8.1 Material and Substrate Selection . . . 142
8.2 Lithography and Pattern Transfer . . . 143
8.3 Device Fabrication Results . . . 146
8.4 Electrical Characterization and Breakdown Testing . . . 148

9 Experimental result **149**
9.1 Phase Plate Mounting and Electrical Connections . . . 150
9.1.1 Mounting the MEMS device on the biasing holder . . . 151
9.1.2 External wiring and voltage control . . . 152
9.1.3 Insertion and in-column alignment workflow . . . 152
9.2 Experimental result : Hyperbola and Helios phase plates . . . 153
9.3 Phase Retrieval via Transport of Intensity . . . 158

IV Discussion and Outlook **161**

10 Conclusions and Future Outlook **162**
10.1 Summary of Contributions and Future Directions . . . 163

Acknowledgements **165**

A Code and Simulation Availability **167**

B Python Code: Functions **168**

B.0.1 Center Detection 168

B.0.2 Radius Estimation 169

B.0.3 Angular Edge Detection and Radial Profile Analysis 170

B.0.4 Aberration Detection and Circularity Analysis 172

List of Figures **177**

List of Tables **186**

Bibliography **186**

Thesis Overview

This thesis presents the design, optimization, fabrication, and experimental validation of tunable electrostatic phase plates for electron-beam shaping in transmission electron microscopy (TEM). The manuscript guides the reader from the physical foundations of electron microscopy to the device concepts and their performance in the microscope. The thesis is divided into three main parts: Part I (Foundations) establishes the electron-optics and TEM background needed for beam shaping; Part II (Concepts and Methods) introduces the main design philosophy and the modeling and optimization toolbox used in the thesis, and Part III (Experimental Results) reports fabrication results and experimental characterization.

Part I (Foundations) presents the principles of electron microscopy and wavefront manipulation. Chapter 1 provides a brief historical background from early optics to electron microscopy, highlighting the theoretical and technological milestones that enabled the development of practical instruments. Chapter 2 establishes the electron-optics framework used throughout the thesis, including relativistic electron properties, the de Broglie wavelength, and complementary classical and wave descriptions of electron motion. It then introduces Fourier optics concepts for wave propagation and image formation in TEM and presents the mathematical background of aberrations. Chapter 3 connects this theory to the instrument by reviewing the main components and overall architecture of a TEM, with particular attention to aberration correctors and detectors. Chapter 4 introduces phase plates as electron-optical elements for phase control; it discusses the primary families of phase plates (thin-film, electrostatic, magnetic, and photonic), with emphasis on electrostatic phase plates.

Part II (Concepts and Methods) develops the device design philosophy and the technical toolkit specific to this thesis. Chapter 5 discusses MEMS-based phase plates for TEM and introduces the central concept of this thesis: current-controlled electrostatic phase plates. The following chapter describes the numerical workflow used throughout the thesis. It provides a comprehensive description of COMSOL modeling for phase-plate design, followed by a discussion of Python-based GUI tools developed for wave-optics simulations in electron microscopy. This chapter also outlines how machine-learning methods can support rapid

aberration detection, which is relevant for alignment and on-the-fly control during phase-plate operation. Using these tools, Chapter 7 introduces the phase-plate designs developed and optimized in this work, following a consistent route from physical principles to implementation, including design, modeling, and optimization. The devices include electrostatic multipole elements for compensating three-fold astigmatism, a current-controlled electrostatic vortex-beam generator, and an orbital-angular-momentum (OAM) sorter.

Finally, Part III (Experimental Results) reports the fabrication results and experimental characterization and compares the simulated phase profiles with the experimental measurements. The thesis closes with a discussion of limitations and trade-offs, as well as an outlook on future directions for tunable MEMS-based phase plates for electron-beam shaping in TEM.

Part I

Foundations

1 Introduction to Electron Microscopy

Contents

1.1 From Magnifiers to Eyeglasses: The Birth of Optics 5
1.2 The Birth of Optical Microscopy 5
1.3 Theoretical Advances and the Nature of Light 6
1.4 The Quantum Framework 10
1.5 From Theory to Technology: The Birth of Electron Microscopy 11

1.1. From Magnifiers to Eyeglasses: The Birth of Optics

Throughout human history, we have been curious about the microscopic world. One of the earliest known attempts to explore it dates back to around 700 B.C, when the Neo-Assyrian civilization used a convex, polished rock crystal known as the *Nimrud lens* (Figure 1.1a). They may have used this piece of crystal as a magnifier or to concentrate the sunlight to make fire [1].

The first written evidence of burning glasses, as a tool commonly used in daily life, can be traced back to ancient Greece. In Aristophanes' comedy *The Clouds*, written in 423 BC. for a festival in honor of the god Dionysus, a dialogue refers to the use of a burning glass. The famous story about Archimedes' burning mirrors is part of later tradition. All of these early observations and inventions emerged long before the formal establishment of optics as a scientific discipline [2].

Although philosophers like Euclid (c. 300 BC) and Ptolemy (c. 2nd century AD) described certain geometrical aspects of light, such as reflection and vision, it was Ibn al-Haytham in the 11th century who first transformed the understanding of optics into a structured scientific discipline. Through his seminal work, Kitab al-Manazir (*The Book of Optics*), he introduced a rigorous, experimentally grounded description of light propagation, reflection, and refraction, laying the foundation for modern ray optics (Figure 1.1b). Ibn al-Haytham's detailed study of refraction and lens behavior in The Book of Optics would, centuries later, become fundamental to the design of eyeglasses, telescopes, and microscopes [3].

While the scientific understanding of optics was still evolving, the practical applications of optical principles began to quietly shape daily life. One of the earliest and most transformative examples was the invention of eyeglasses in Florence around 1280 (Figure 1.1c). Although simple in design, these early spectacles marked the first modern application of optical science, translating centuries of theoretical insight — from ancient magnifiers to Ibn al-Haytham's foundational work — into a tangible tool that improved human vision [2]. These milestones established the conceptual and practical roots from which Renaissance and early modern optics would later emerge.

1.2. The Birth of Optical Microscopy

It was at the end of the 16th century that scientists first noticed the magnifying properties of multiple lenses. The credit for the first compound optical microscope is generally given

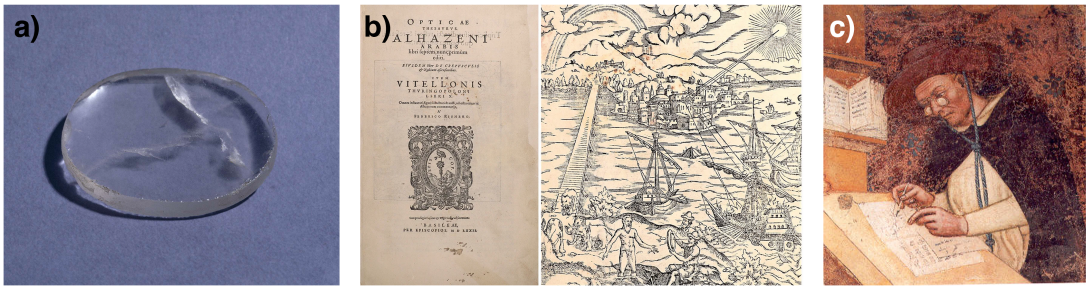


Figure 1.1. Early optics milestones. (a) The Nimrud lens (8th c. BC), a rock-crystal magnifier (Image credit: British Museum). (b) Medieval Latin illustration of a parabolic burning mirror igniting ships (Archimedes legend; Alhazen tradition). (c) Tommaso da Modena's 1352 fresco at San Nicolò, Treviso — earliest known depiction of spectacles on Cardinal Hugh of Saint-Cher.

to Hans Janssen and his son Zacharias in 1595 (Figure 1.2a) [4]. The eyepiece lens was bi-convex, and the objective lens was plano-convex —a very advanced compound design for that period. Focusing on this hand-held microscope was achieved by sliding the tube in or out while observing the sample. The Janssen microscope was capable of magnifying images approximately three times when fully closed and up to ten times when extended to the maximum. The same basic principles underlie both the early compound microscope and Galileo's telescope, which, in a sense, can be regarded as the microscope operated in reverse.

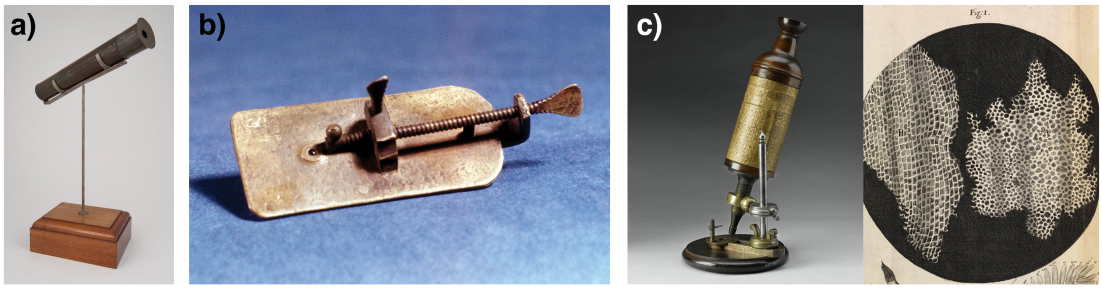


Figure 1.2. Early microscopes. (a) Janssen-type compound microscope (c. 1590s). (b) Antoni van Leeuwenhoek single-lens microscope (1670s). (c) Robert Hooke's compound microscope with an image from his book *Micrographia* (1665). (Images credit: Science Museum Group, London. (website).)

By the late 1600s, designs branched. In England, makers built tripod-style instruments from wood, pasteboard, and leather; in Italy, smaller turned-wood and brass microscopes were popular (see Figure 1.2b–c for examples of early forms). The variety reflects different workshops and use cases, but all shared the same goal: stable focusing and higher usable magnification. It was also during this period that several prominent microscopists conducted fundamental experiments that have defined microscopy throughout the ages. Pioneers such as Antoni van Leeuwenhoek, Robert Hooke, Giuseppe Campani, and Marcello Malpighi made significant discoveries that paved the way for future microscopists.

1.3. Theoretical Advances and the Nature of Light

In our historical journey through the evolution of optics, the time has come to step into the era of modern science — an era marked by experimentation and mathematical reasoning. It was Isaac Newton, one of history's most influential scientific minds, who reshaped the understanding of light itself. With his famous prism experiments, Newton demonstrated that white light is not pure but rather a composition of all visible colors. This groundbreaking discovery, published in *Opticks* in 1704, not only solved long-standing questions about color but also challenged the philosophical views of light that had persisted for centuries [5]. More than just an experimentalist, Newton's work laid the theoretical foundations that would guide optical science for generations, bridging the gap between ancient speculation and the quantitative precision of modern physics.

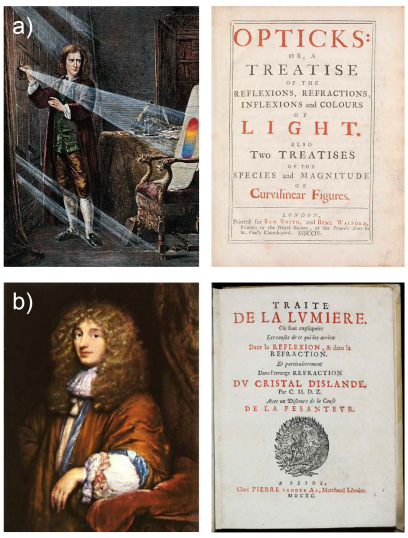


Figure 1.3. Nature of light. (a) Isaac Newton with *Opticks* (1704): light as particles. (b) Christiaan Huygens with *Traité de la lumière* (1690): light as waves. (Image credit: Smithsonian Libraries, Washington, website).

However, even as Newton's work on the composite nature of light laid the foundations of modern optics, the nature of light itself remained an open question. This question would ignite one of the most famous scientific debates of the time. While Newton saw light as tiny particles, his contemporary Christiaan Huygens suggested a very different idea—that light behaves like a wave. Huygens' wave theory, outlined in *Traité de la lumière* (Treatise on Light) in 1690, described light as a wave propagating through an all-pervasive medium. For centuries, these two competing theories coexisted, each partially explaining experimental observations but leaving fundamental questions unanswered. It was not until the advent of quantum mechanics and the experimental work of the 19th and 20th centuries that the true dual nature of light — as both wave and particle — was fully recognized, reshaping the foundations of optical science [6].

One of the earliest experimental demonstrations hinting at the wave nature of light can be traced back to Thomas Young's famous double-slit experiment in 1801 [7]. In 1665, Francesco Maria Grimaldi reported systematic observations of diffraction — the slight spreading and boundary fringes of light passing edges and apertures — which he interpreted as a new property of light and named 'diffraction'. However, he did not frame his observations in terms of interference between two coherent light beams [8]. Years later, inspired by the analogy with overlapping sound beats, Young began his efforts to establish the wave theory of light. He repeated Grimaldi's experiment, but this time sunlight passed through an initial pinhole, which served as the primary source of illumination. The light emerging from this pinhole can be considered spatially coherent, and it was used to illuminate two closely spaced apertures. Figure 1.4 illustrates a schematic of the experiment. Two coherent cones of light then emerged from these apertures and propagated toward the viewing screen. The closer the

apertures were placed, the more their corresponding light disks overlapped on the screen. In the region of overlap, the two waves interfered, producing alternating bright and dark bands, known as interference fringes. Based on the classical understanding of light as particles, one would expect to observe only two bright disks on the viewing screen, corresponding to the two apertures [6].

Although Young's experiment was originally interpreted through classical wave theory, its conceptual power has echoed far beyond the 19th century. In a remarkable modern counterpart, the first single-electron interference experiment in a TEM, conceptually equivalent to a double-slit configuration, was reported by Merli, Missiroli, and Pozzi in 1976, who observed the statistical build-up of the interference fringes from individual electrons in a TEM-based setup [9, 10, 11]. Even more strikingly, these fringes appeared even when electrons were emitted one at a time, seemingly forcing each individual particle to interfere with itself. At first glance, this result seems counterintuitive: electrons are classically regarded as particles. How then can they form a wave-like interference pattern? The resolution to this paradox lies in the quantum framework that emerged in the early twentieth century, a revolution in physics that redefined the nature of both matter and light [12].

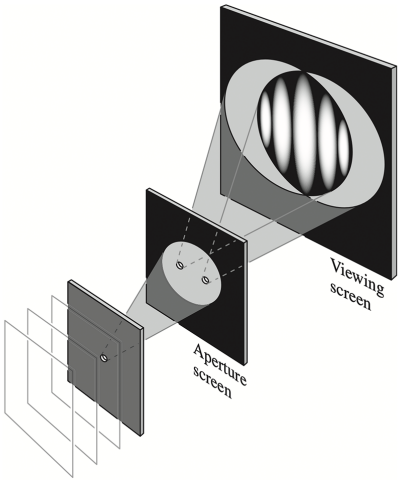


Figure 1.4. Young's double-slit experiment

Following Thomas Young's pivotal double-slit experiment in 1801, the first half of the 19th century witnessed significant advancements that solidified the wave theory of light. Among the most influential contributors was Augustin-Jean Fresnel, who extended Young's work by developing a comprehensive mathematical framework for wave optics. Fresnel's rigorous treatment of diffraction, interference, and polarization — along with his derivation of the famous Fresnel equations — provided strong theoretical support for the wave model. Meanwhile, Joseph von Fraunhofer's discovery and mapping of dark lines in the solar spectrum laid the foundation for modern spectroscopy and deepened the understanding of light-matter interactions. In parallel, Michael Faraday's experimental discoveries in electromagnetism — particularly electromagnetic induction and the Faraday effect — revealed profound connections between light and magnetic fields, suggesting that light itself might be an electromagnetic phenomenon [6, 13]. These cumulative insights set the stage for James Clerk Maxwell, who in the 1860s would unify electricity, magnetism, and optics into a single theoretical framework, marking the birth of classical electromagnetic theory [14, 15].

The laws of classical electrodynamics were discovered in fragments through the pioneering work of scientists such as Benjamin Franklin, Charles-Augustin de Coulomb, André-Marie Ampère, and Michael Faraday. However, it was James Clerk Maxwell who unified and formalized these empirical laws into a consistent theoretical framework. His formulation culminated in the four fundamental equations that govern classical electromagnetism:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \qquad \text{(Gauss's Law for Electricity)} \tag{1.1}$$

$$\nabla \cdot \mathbf{B} = 0 \qquad \text{(Gauss's Law for Magnetism)} \tag{1.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{(Faraday's Law of Induction)} \tag{1.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \qquad \text{(Ampère–Maxwell Law)} \tag{1.4}$$

From these equations, one can derive the wave equation for electromagnetic waves in vacuum. Taking the curl of Faraday's Law and substituting Ampère–Maxwell's Law leads to:

$$\nabla^2 \mathbf{E} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \tag{1.5}$$

and

$$\nabla^2 \mathbf{B} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0 \tag{1.6}$$

These equations reveal that electromagnetic fields propagate as waves through space at a speed

$$v = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \approx 3 \times 10^8 \ \mathrm{m\,s^{-1}},$$

which matches the speed of light. The wave transports energy in the direction of the Poynting vector,

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B},$$

So for a plane wave the propagation direction is set by $\mathbf{E} \times \mathbf{B}$ (with $\mathbf{E} \perp \mathbf{B}$ and both perpendicular to the direction of travel). Thus, Maxwell concluded that light is an electromagnetic wave — a profound realization that united optics and electromagnetism under a single theoretical structure [16, 6].

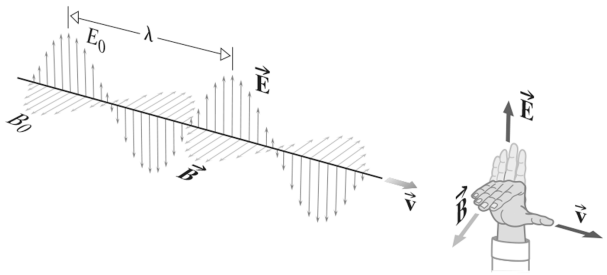


Figure 1.5. Wave propagation in the direction of $\mathbf{E} \times \mathbf{B}$ [6].

1.4. The Quantum Framework

Despite the success of Maxwell's electromagnetic theory in describing the wave nature of light, by the end of the nineteenth century, several results resisted a classical explanation. A central case was blackbody radiation. Classical theory (Rayleigh–Jeans) predicted an "ultraviolet catastrophe," with the emitted energy diverging at short wavelengths. In 1900, Max Planck accounted for the observed spectrum by proposing that electromagnetic energy is exchanged in discrete quanta with $E = h\nu$, where h is Planck's constant and ν is the frequency of the radiation. This step is widely regarded as the start of quantum theory [15, 17, 18].

Building on this concept, Albert Einstein, in 1905, provided a revolutionary explanation of the photoelectric effect — the emission of electrons from a metal surface when illuminated by light. Classical theory predicted that the energy of ejected electrons should depend on light intensity, but experiments showed that it depends instead on the light's frequency. Einstein proposed that light itself is quantized into particles, or photons, each carrying energy $E = h\nu$. This explanation not only confirmed Planck's hypothesis but also provided compelling evidence for the particle-like nature of light [15, 17, 18]. For this contribution, Einstein was awarded the Nobel Prize in Physics in 1921. His work fundamentally challenged the wave-only model of light and opened the path to the wave–particle duality description that underlies quantum mechanics.

Before 1924, the nature of light remained unresolved, as both wave and particle models offered compelling evidence. In 1924, Louis de Broglie introduced the concept of wave-particle duality for matter, proposing that particles such as electrons possess wave-like properties. His well-known relation, $\lambda = h/p$, links the particle's momentum p to its associated wavelength λ. This dual nature implies that moving electrons can exhibit wave behavior at picometer-scale wavelengths [19]. A more detailed discussion of the electron wavelength will be presented in Chapter 2.

Experimental confirmation of de Broglie's theory came in 1927 through two independent studies. Clinton J. Davisson and Lester H. Germer observed electron diffraction by directing an electron beam at a nickel crystal. Initially working with polycrystalline samples, they noticed only smooth angular distributions. However, after an accidental recrystallization created a near-monocrystalline structure, they detected sharp diffraction peaks at specific angles. It was not until Davisson learned of de Broglie's and Schrödinger's wave theories at a 1926 conference in Oxford that they understood the significance of their findings. Their subsequent measurements confirmed that electrons exhibit wave-like behavior consistent with the de Broglie relation $\lambda = h/p$ [20]. Around the same time, G. P. Thomson and Reid independently demonstrated electron diffraction using thin metal foils [21]. For his foundational theoretical insight, de Broglie was awarded the Nobel Prize in Physics in 1929 [22]. Davisson and Thomson were later jointly awarded the 1937 Nobel Prize in Physics for their experimental discovery of electron diffraction [22].

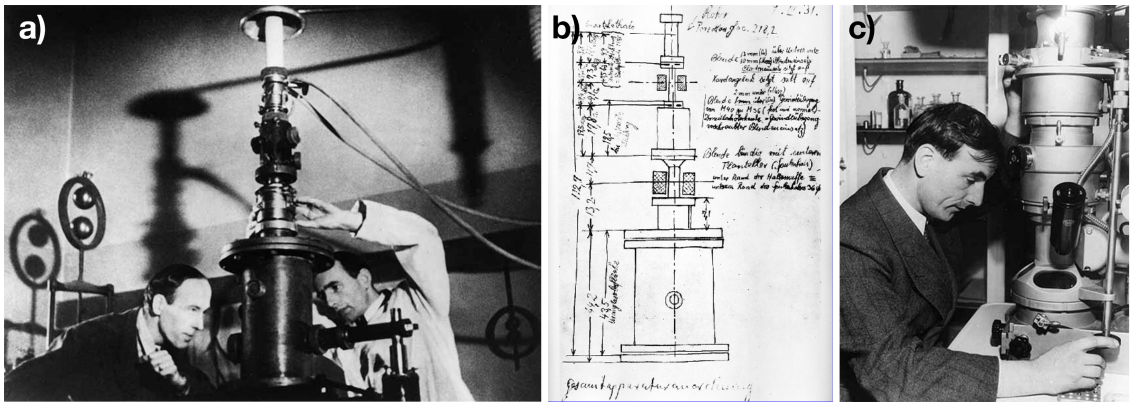


Figure 1.6. Historical milestones in the development of the electron microscope. (a) Ernst Ruska (right) and Max Knoll (left) with the first magnetic electron microscope, Berlin, 1932. (b) Sketch by Ernst Ruska, dated 9 March 1931, depicting the cathode ray tube used to test one-stage and two-stage electron-optical imaging with magnetic electron lenses ("electron microscope"). (c) Ernst Ruska adjusting a Siemens electron microscope, 1949.

1.5. From Theory to Technology: The Birth of Electron Microscopy

Hans Busch, in 1926, theoretically demonstrated that the magnetic field generated by a short coil could act as a convex lens for electrons—an idea that profoundly influenced Ernst Ruska. While reviewing the scientific literature, Ruska encountered Busch's paper and immediately recognized the parallels between magnetic electron optics and conventional light optics. Motivated by this analogy, In April 1931, Ernst Ruska and Max Knoll constructed a two-stage apparatus using short magnetic coils to image aperture grids, achieving a modest magnification of approximately 17×. Although initially unaware of de Broglie's theory, Ruska realized, upon learning of it, that electron wavelengths were five orders of magnitude shorter than those of visible light. Applying Abbe's formula with an accelerating voltage of 75 kV, Knoll and Ruska estimated a theoretical resolution of 0.22 nm—a limit only reached decades later. By 1933, Ruska's design surpassed the resolution of optical microscopes, and throughout the 1930s and 1940s, he collaborated with Siemens to develop commercial electron microscopes [23] (see Figure 1.6). For his contributions to the invention of the electron microscope, Ruska was awarded half of the 1986 Nobel Prize in Physics [22]. Over the same period, electron microscopy gained rapid industrial interest, and early commercial instruments began to appear—most notably from Siemens in Germany, university–industry efforts in Canada, and RCA in the United States (see Figure 1.7). Parallel to Ruska's work, Manfred von Ardenne proposed and built the first scanning transmission electron microscope in Berlin (1937–1938, for Siemens), adding scanning coils to a TEM to raster a fine probe and detect transmitted electrons [24, 25]. Its performance was inferior to TEM, von Ardenne worked on it for only two years, and the instrument was destroyed in a 1944 air raid; he did not return to the project after WWII. Practical STEM was realized later by Albert Crewe in the late 1960s. In 1939, Walther Kossel and Gottfried Möllenstedt demonstrated that projected two-dimensional images and electron diffraction patterns could be recorded, providing information about crystal structure, lattice spacing, and crystallographic symmetry [26].

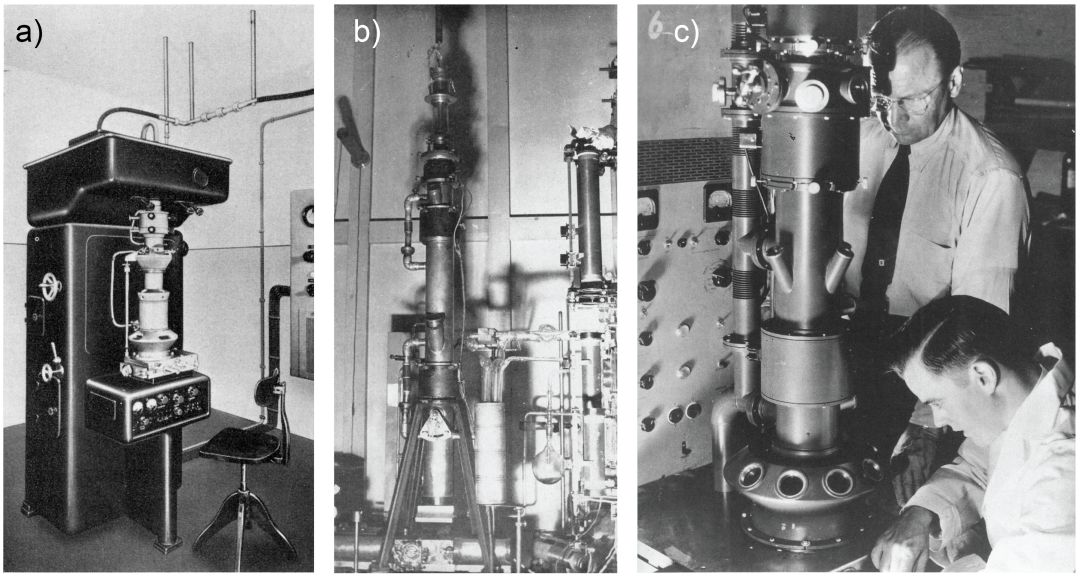


Figure 1.7. Three early commercial electron microscopes. (a) Siemens Übermikroskop, 1939 (Germany); (b) Toronto TEM, 1939 — E. F. Burton, J. Hillier, A. F. Prebus (Canada); (c) RCA Model B, 1940 — V. K. Zworykin, J. Hillier (USA).

After WWII, electron microscopy gained significant attention from scientists worldwide. Richard P. Feynman, in his 1959 lecture to the American Physical Society titled "There's Plenty of Room at the Bottom," envisioned the possibility of constructing and manipulating matter with atomic precision—assembling structures atom by atom. He noted that modern electron microscopes could resolve only around 10 Å. He challenged physicists to improve this by a factor of 100: "The wavelength of the electron ... is only 1/20 of an angstrom. So it should be possible to see the individual atoms. Make the electron microscope 100 times better." Feynman argued that such capability would allow researchers to answer fundamental biological questions simply by "looking at the thing" and connecting atomic arrangements to molecular functions. His visionary remarks—and the now-famous phrase "plenty of room at the bottom"—inspired generations of microscopists and nanotechnologists [27].

A primary barrier to achieving Feynman's vision was lens aberrations. In the 1930s, Otto Scherzer demonstrated that spherical and chromatic aberrations could not be eliminated using rotationally symmetric static electromagnetic fields, establishing a fundamental limit to resolution [28, 29, 30]. In 1990, Harald Rose proposed a practical solution using multipole aberration correctors [31]. Together with Max Haider and Knut Urban, this concept was successfully implemented in the late 1990s [32]. The first aberration-corrected TEM images were published in 1998, and by 2002, sub-ångström scanning TEM (STEM) images were achieved [33]. These aberration-corrected instruments enabled scientists to determine atomic positions with picometer-scale precision and perform spectroscopy on individual atomic columns. For this groundbreaking achievement, Rose, Haider, Urban, and Krivanek were awarded the 2020 Kavli Prize in Nanoscience—a development widely recognized as the pivotal step in transforming Ruska's original TEM into the modern atomic-resolution microscope [34].

Electron microscopes were long considered unsuitable for imaging biological molecules be-

cause of radiation damage. In the early 1980s, Jacques Dubochet and colleagues introduced the vitrification of specimens, allowing macromolecules to be preserved in amorphous ice and observed in near-native conditions [35]. Building on this foundation, Richard Henderson and co-workers demonstrated in 1990 that atomic-resolution structures of proteins, such as bacteriorhodopsin, could be obtained by electron crystallography of two-dimensional crystals embedded in vitrified ice [36, 37]. Joachim Frank developed single-particle analysis to merge noisy two-dimensional images into sharp three-dimensional structures [38]. Cryo-electron microscopy (cryo-EM) now routinely yields atomic-resolution structures of biomolecules. In 2017, Dubochet, Frank, and Henderson were awarded the Nobel Prize in Chemistry "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".

Today, electron microscopes are indispensable tools across materials science, biology, and physics, enabling atomic-resolution imaging, quantitative spectroscopy, *in situ* and operando studies, and cryogenic structure determination. Building on these capabilities, modern electron optics employs aberration correction, phase retrieval, holography, and programmable phase modulation to control the electron wavefront with increasing precision. In the chapters that follow, we develop the foundations of electron microscopy—from basic contrast formation to advanced wave-optical methods—and ultimately focus on the central theme of this thesis: electron beam shaping and its applications to imaging, spectroscopy, and electron-matter interaction engineering.

2 Electron Optics

Contents

2.1 Introduction to Electron Optics 15

2.1.1 Relativistic electron kinematics 15

2.1.2 Classical Description of Electron Motion 17

2.1.3 Schrödinger equation for charged particles 18

2.2 Electron–Specimen Interactions 20

2.3 Fourier Optics 22

2.4 Point Spread Function (PSF) 24

2.5 Aberrations in Electron Microscopy 27

2.5.1 Defocus 30

2.5.2 Spherical Aberration 31

2.5.3 Chromatic Aberration 32

2.5.4 Twofold and Threefold Astigmatism 32

2.5.5 Combined Effects and Correction 33

2.6 Chapter Summary 33

2.1. Introduction to Electron Optics

Electrons are fundamental particles with unique physical and quantum-mechanical properties that make them ideal "probes" for high-resolution imaging. Each electron has a mass of approximately 9.109×10^{-31} kg, a negative charge of 1.602×10^{-19} C, and spin-$\frac{1}{2}$, classifying it as a fermion. Due to their charge, electrons interact strongly with electromagnetic fields and can be accelerated to high velocities using electric potentials, often reaching relativistic speeds. As a result, relativistic effects must be considered in electron microscopy. According to de Broglie's matter-wave duality, moving particles exhibit wave-like behavior. This wave-particle duality serves as a conceptual bridge between classical optics and electron optics. While optics deals with the propagation and focusing of light waves, electron optics applies a similar mathematical framework to describe and control electron beams in imaging systems. The use of electrons instead of photons is primarily motivated by their much shorter wavelength at high energies, which enables spatial resolutions far beyond those achievable with visible light.

This chapter introduces the physical foundations of electron optics and its role in transmission electron microscopy (TEM). It begins with the derivation of the relativistic electron wavelength and explains how this property is key to achieving sub-nanometer resolution. The following sections review electron–specimen interactions, focusing on scattering mechanisms—both elastic and inelastic—that provide contrast and information in TEM imaging. The chapter then introduces mathematical tools such as the point spread function (PSF) and Fourier optics, which describe electron beam propagation and image formation. Finally, it explores the origins and classifications of aberrations in electron lenses, particularly geometrical (monochromatic) aberrations, and discusses their impact on image resolution.

2.1.1. Relativistic electron kinematics

When an electron is accelerated by a potential difference U, its total energy increases due to the gain in kinetic energy. The relativistic expressions below describe how the velocity, momentum, and de Broglie wavelength of the electron are related to the accelerating voltage. As electrons approach a significant fraction of the speed of light, classical mechanics becomes insufficient, and relativistic corrections must be considered [39].

$$E_{\text{total}} = E_{\text{rest}} + eU = m_e c^2 + eU \tag{2.1}$$

Here, m_e is the rest mass of the electron, c is the speed of light, e is the electron charge, and U is the accelerating voltage. The Lorentz factor γ is defined as:

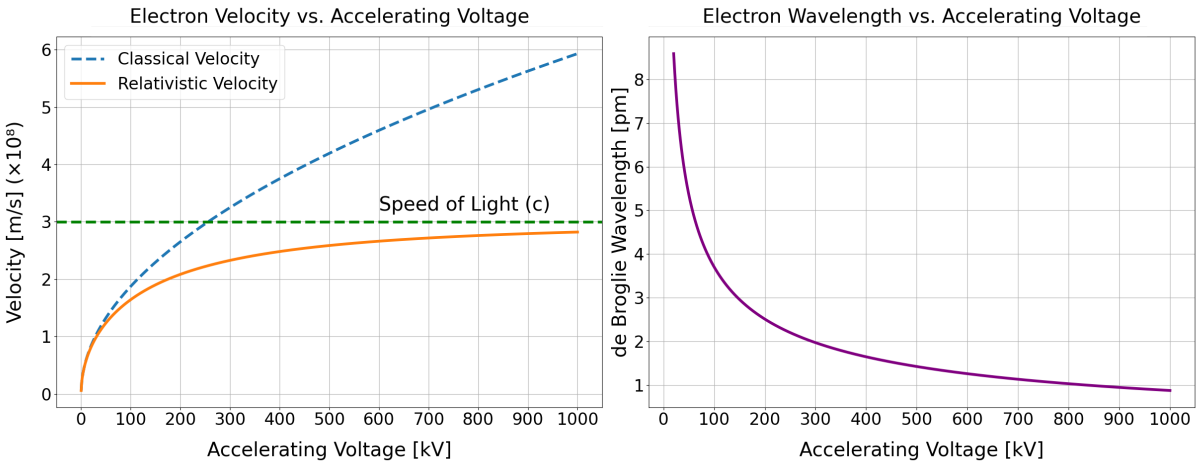


Figure 2.1. Electron velocity and de Broglie wavelength as a function of accelerating voltage. The left plot compares classical and relativistic velocity; the right plot shows the corresponding electron wavelength.

$$\gamma = \frac{E_{\text{total}}}{E_{\text{rest}}} = 1 + \frac{eU}{m_e c^2} \tag{2.2}$$

The relativistic velocity v of the electron is obtained from:

$$\beta = \frac{v}{c} = \sqrt{1 - \frac{1}{\gamma^2}} \tag{2.3}$$

$$v = \beta c = c\sqrt{1 - \frac{1}{\gamma^2}} \tag{2.4}$$

The relativistic momentum of the electron is:

$$p = \gamma m_e v \tag{2.5}$$

Finally, the relativistic de Broglie wavelength is given by:

$$\lambda = \frac{h}{p} = \frac{h}{\gamma m_e v} \tag{2.6}$$

Figure 2.1 compares the classical and relativistic electron velocity v versus accelerating voltage U and shows how the de Broglie wavelength λ decreases as U increases. As shown, relativistic effects become significant at higher voltages. Above several hundred kilovolts, the wavelength approaches or falls below 1 pm; in standard TEM voltages (100–300 kV), it lies in the 2–4 pm range.

This behavior is particularly important in transmission electron microscopy, where the achievable spatial resolution is ultimately bounded by λ, but in practice, it is dominated by lens aberrations and coherence [40]. The calculated parameters for commonly used voltages in electron microscopes are listed in Table 2.1.

Table 2.1. Electron parameters at various accelerating voltages: Lorentz factor γ, velocity ratio β, relativistic and classical velocities, and the de Broglie wavelength λ.

Voltage (kV)	$\boldsymbol{\gamma}$	$\boldsymbol{\beta}$	**Relativistic** v **(m/s)**	**Classical** v **(m/s)**	$\boldsymbol{\lambda}$ **(pm)**
10	1.02	0.195	5.85×10^{7}	5.93×10^{7}	12.21
50	1.10	0.413	1.24×10^{8}	1.33×10^{8}	5.36
100	1.20	0.548	1.64×10^{8}	1.88×10^{8}	3.70
200	1.39	0.695	2.08×10^{8}	2.65×10^{8}	2.51
300	1.59	0.777	2.33×10^{8}	3.25×10^{8}	1.97
400	1.78	0.828	2.48×10^{8}	3.75×10^{8}	1.64
500	1.98	0.863	2.59×10^{8}	4.19×10^{8}	1.42
800	2.57	0.921	2.76×10^{8}	5.30×10^{8}	1.03
1000	2.96	0.941	2.82×10^{8}	5.93×10^{8}	0.87

2.1.2. Classical Description of Electron Motion

In the classical picture, an electron is treated as a point charge $-e$ with mass m_e moving under electric and magnetic fields. Its motion obeys the Lorentz force law [41]:

$$\frac{d\mathbf{p}}{dt} = -e\big(\mathbf{E} + \mathbf{v} \times \mathbf{B}\big), \tag{2.7}$$

where $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, $\mathbf{v}$ is the velocity, and $\mathbf{p}$ is the particle momentum. The electric field changes the magnitude of $\mathbf{p}$ (hence the speed), while the magnetic Lorentz force $-e\,\mathbf{v} \times \mathbf{B}$ is always perpendicular to the instantaneous velocity, so it cannot do work and does not change the speed. Instead, it continuously *turns* the momentum vector. In electrostatic lenses and accelerators, $-e\mathbf{E}$ acts along field lines, modifying the speed and deflecting the path toward regions of higher potential according to Eq. (2.7).

In purely magnetic regions ($\mathbf{E} = 0$), the force is perpendicular to $\mathbf{v}$, producing uniform circular motion in the plane normal to $\mathbf{B}$. It is convenient to express the *radius of curvature* directly in terms of momentum [41]:

$$R = \frac{p_\perp}{eB} = \frac{\gamma m_e v_\perp}{eB}, \tag{2.8}$$

Here $p_\perp$ and $v_\perp$ are the components perpendicular to $\mathbf{B}$. The corresponding *cyclotron frequency* is

$$\omega = \frac{v}{R} = \frac{eBv}{p} = \frac{eB}{\gamma m_e}, \tag{2.9}$$

Where e denotes the magnitude of the electron charge. Equations (2.8)–(2.9) summarize how magnetic fields steer electrons in lenses and deflectors: stronger B or smaller p yield tighter

curvature (stronger focusing/deflection). In electron microscopes, the magnetic field of a lens is not uniform, so this expression represents only the *local radius of curvature* experienced by the electrons. In chapter 3, we discuss in detail how electrostatic and magnetic lenses are employed in electron microscopy to control and focus the electron beam.

2.1.3. Schrödinger equation for charged particles

In the non relativistic approximation, the quantum–mechanical state of a single electron is described by a complex wavefunction $\psi(\mathbf{r},t)$, normalized so that $\int |\psi(\mathbf{r},t)|^2 \mathrm{d}^3\mathbf{r} = 1$. In a static potential $V(\mathbf{r})$, ψ obeys the time-dependent Schrödinger equation [42]

$$i\hbar \frac{\partial \psi}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + eV(\mathbf{r}) \right] \psi(\mathbf{r},t). \tag{2.10}$$

For stationary states $\psi(\mathbf{r},t) = \psi(\mathbf{r})e^{-i\omega t}$ with $E = \hbar\omega$, this becomes

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + eV(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{2.11}$$

In free space ($V = 0$) the solutions reduce to the Helmholtz form

$$(\nabla^2 + k^2)\psi = 0, \qquad k = \sqrt{\frac{2mE}{\hbar^2}} = \frac{p}{\hbar} = \frac{2\pi}{\lambda}, \tag{2.12}$$

The transition to the relativistic form should be done using the proper Dirac equation but for our regime it is sufficient to generalize equations 2.10 and 2.12. Equation 2.12 retains the non-relativistic Helmholtz form, but the relativistic effects of the acceleration voltage are included by replacing the kinetic energy E with an effective, relativistically corrected energy $E^* = \frac{1+\gamma}{2} E$ [43]:

$$\lambda = \frac{h}{\sqrt{2m_0E\left(1 + \dfrac{E}{2E_0}\right)}} = \frac{h}{\sqrt{2m_0E^*}}, \tag{2.13}$$

Moreover, a relativistic correction to the potential in Equation 2.10 must be applied. In free space, Equation 2.12 has simple solutions of two basic types: plane waves and spherical waves. For a paraxial electron beam, it is convenient to factor out the rapid oscillation along z by writing

$$\psi(\mathbf{r}) = u(x,y,z)\, e^{ikz}.$$

Substituting this form into Equation 2.11 and, within the paraxial approximation, neglecting $\partial^2 u/\partial z^2$ yields

$$2ik\frac{\partial u}{\partial z} + \nabla_\perp^2 u - \frac{2me}{\hbar^2}V(\mathbf{r})\, u = 0, \tag{2.14}$$

where $\nabla_\perp^2$ is the transverse Laplacian. This form shows that the potential acts as a spatially varying phase term that modifies the evolution of the electron wave.

Because Eq. (2.10) is linear, electron waves obey the superposition principle. If two partial

waves $\psi_1 = a_1 e^{i\varphi_1}$ and $\psi_2 = a_2 e^{i\varphi_2}$ overlap, the combined intensity is

$$I = |\psi_1 + \psi_2|^2 = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\Delta\varphi, \tag{2.15}$$

with $\Delta\varphi = \varphi_2 - \varphi_1$. Constructive and destructive interference arise directly from this cross term, illustrating that the relative phase carries observable physical information.

Real electrons have a finite energy spread ΔE and an angular divergence θ, which limit the temporal and spatial coherence of the wavefield. The coherence length, $L_c \simeq \lambda E/(2\Delta E)$, characterizes the range over which the phase remains stable in time, while the transverse coherence width, $L_T \sim \lambda/(2\theta)$, specifies the lateral region of mutual phase correlation. Both parameters determine the degree to which interference effects can be observed in practice [43].

This formulation completes the mathematical description of electron waves and their propagation. In the following, these principles are extended to external electromagnetic fields and electron-optical elements that modify the amplitude and phase of the electron wave.

Up to this point, we have considered free-electron propagation and the influence of scalar potentials through the Schrödinger equation. When electromagnetic fields are present, it is convenient to describe the wave evolution in terms of the associated potentials: the scalar potential $V(\mathbf{r})$ and the vector potential $\mathbf{A}(\mathbf{r})$. The time-independent Schrödinger equation then takes the form:

$$\left[\frac{1}{2m}\left(-i\hbar\nabla - q\mathbf{A}(\mathbf{r})\right)^2 + qV(\mathbf{r})\right]\psi(\mathbf{r}) = E\,\psi(\mathbf{r}). \tag{2.16}$$

where q is the particle charge (for electrons, $q = -e$).[1]

The potentials appear explicitly in the Hamiltonian and, therefore, govern the phase evolution of the electron wavefunction. This viewpoint is central to the *Aharonov–Bohm effect*, in which a measurable phase shift can arise in an interference experiment even when the electron propagates through a region where both $\mathbf{E}$ and $\mathbf{B}$ vanish [44].

For an electron traveling along a trajectory $\mathcal{C}$, the total phase accumulated due to electromagnetic potentials is [45]

$$\Delta\varphi = -\frac{e}{\hbar}\int_{\mathcal{C}} \mathbf{A}\cdot d\mathbf{l} + \frac{e}{\hbar v}\int V(\mathbf{r})\,dz. \tag{2.17}$$

The first term corresponds to the magnetic contribution, and the second to the electrostatic one. For magnetic phase shifts, the electron passes through a region enclosing a confined magnetic flux Φ, giving

$$\Delta\varphi = -\frac{e}{\hbar}\oint_{\mathcal{C}} \mathbf{A}\cdot d\mathbf{l} = -2\pi\,\frac{\Phi}{\Phi_0}, \qquad \Phi_0 = \frac{h}{e}. \tag{2.18}$$

[1] For TEM electron energies, a fully relativistic treatment starts from the Dirac equation; in practice, the standard phase-shift expressions used in electron optics are obtained by combining a paraxial wave description with relativistically corrected kinematics, which is reflected in the interaction constant defined below.

Where Φ_0 is the magnetic flux quantum, expressed in units of weber (Wb). Thus, notice that even when $\mathbf{B} = 0$ along the paths, the electron wave accumulates a phase proportional to the enclosed magnetic flux.

In the electrostatic case, where only the scalar potential $V(\mathbf{r})$ is present and the electron moves at a constant velocity v, Eq. 2.17 simplifies to

$$\Delta\varphi(x, y) = \frac{e}{\hbar v} \int V(x, y, z)\, dz = C_E \int V(x, y, z)\, dz, \tag{2.19}$$

with $C_E = e/(\hbar v)$ the interaction constant. Representative values of C_E at common accelerating voltages are summarized in Table 2.2.

Table 2.2. Interaction constant C_E as a function of acceleration voltage.

Acceleration Voltage U (kV)	**λ (pm)**	**C_E [rad/(V µm)]**
60	4.8661	11.36
100	3.7014	9.24
200	2.5079	7.29
300	1.9687	6.53

Equation 2.19 shows how an electrostatic potential modifies the phase of an electron wavefront through the phase accumulated from the electrostatic potential. In the scalar-wave approximation, the exit wave can be expressed as

$$\psi(x, y) = A(x, y)\, e^{i\varphi(x,y)}, \tag{2.20}$$

where the amplitude $A(x, y)$ describes absorption and scattering, and the phase $\varphi(x, y)$ represents the accumulated phase shift. This principle underlies all electrostatic and magnetic methods of phase control and will serve as the foundation for the discussion of phase-shifting devices in the following chapters.

2.2. Electron–Specimen Interactions

Scattering is the fundamental physical principle underlying all electron microscopy techniques. In light optics, when a beam of light interacts with matter, it rarely passes unaffected. The oscillating electric field of the light induces vibrations of the electrons in the material, and these oscillating charges re-emit light waves in various directions — a process known as scattering [6]. The scattering behavior depends strongly on the size of the object relative to the wavelength: particles much smaller than the wavelength scatter isotropically (Rayleigh scattering, responsible for the blue color of the sky); particles comparable in size produce

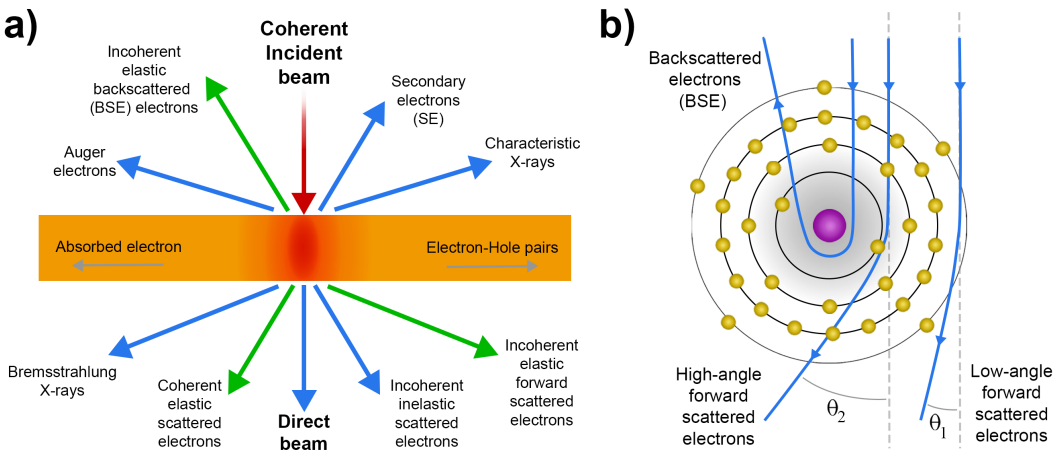


Figure 2.2. (a) Overview of electron–specimen interactions in a thin sample, including elastic and inelastic processes. (b) Elastic-scattering regimes by angle: low-angle forward (coherent diffraction), high-angle forward (large-momentum-transfer), and backscattered (>90°).

more complex angular distributions (Mie scattering, explaining the whiteness of clouds); and objects much larger than the wavelength follow the approximations of geometric optics [6].

A similarly large set of different physical phenomena occurs in the interaction of electrons with matter (e.g., different wavelengths and interaction diagrams), except that the incident beam consists of electrons rather than photons. In this case, scattering refers to the deflection of the electron beam — either elastically or inelastically — from its original trajectory. This deflection occurs due to interactions with the electrostatic potential of the atomic nuclei and the surrounding electron cloud. These interactions redistribute the electron's momentum and, in energy terms, are significant for inelastic processes but typically negligible for elastic scattering at small scattering angles in TEM.

Figure 2.2 provides an overview of the main interaction mechanisms that occur when high-energy electrons pass through a thin specimen. As illustrated in Figure 2.2(a), electrons may be scattered either *elastically*—changing direction without losing kinetic energy—or *inelastically*, transferring part of their energy to the sample. The angular regimes of elastic scattering, ranging from low-angle forward diffraction to high-angle incoherent scattering and backscattering, are summarised in Figure 2.2(b).

The pattern and strength of electron scattering — including the angles, intensities, and energies of the scattered electrons — encode detailed information about the specimen. Elastic scattering primarily reveals structural features such as atomic positions, crystal symmetry, and lattice spacing through diffraction patterns. Inelastic scattering carries complementary information about the specimen's composition, chemical bonding, and electronic structure. Together, these processes provide the fundamental signals used in TEM to form images, record diffraction data, and perform spectroscopic analysis, enabling resolution down to the atomic scale.

A complete treatment of electron–specimen interactions includes topics such as differential scattering cross-sections, dynamical diffraction, multiple scattering, inelastic mean free paths,

and spectroscopic signatures. A detailed discussion of these aspects lies beyond the scope of this thesis, but the reader may find several references for further study, including Reimer and Kohl's *Transmission Electron Microscopy* [43], Egerton's *Electron Energy-Loss Spectroscopy* [46], and Williams and Carter's *Transmission Electron Microscopy: A Textbook for Materials Science* [41].

2.3. Fourier Optics

Fourier optics provides a central framework for describing image formation and analysis in electron microscopy by relating real-space variations in the electron wave to its representation in reciprocal (spatial-frequency) space. Any imaging system records a signal determined by the transmitted electron wave, thereby encoding structural information in the exit-wave amplitude $A(\mathbf{r})$ and phase $\phi(\mathbf{r})$, as described by Eq. 2.20. This information is contained in the spatial variations of both quantities across the field of view: rapid changes in A or ϕ correspond to higher spatial-frequency content. In the local plane-wave (eikonal) approximation, the phase gradient defines a local wavevector, $\mathbf{k}(\mathbf{r}) = \nabla\phi(\mathbf{r})$, and the corresponding spatial-frequency vector (in cycles per unit length) is $\mathbf{f} = \mathbf{k}/(2\pi)$. Regions where the phase varies more rapidly, therefore, contribute more strongly to higher spatial frequencies in the Fourier representation.

Periodic or quasi-periodic structures—such as crystals, gratings, or nanostructured materials—produce discrete Fourier components located at the reciprocal-lattice vectors, giving rise to sharp features in reciprocal space (diffraction spots or rings). The Fourier transform is thus the natural framework for analyzing and interpreting these frequency components: it decomposes the wavefunction $\psi(x, y)$ into its constituent spatial frequencies, revealing both their amplitude (contrast) and phase (relative position)[47].

Under the paraxial approximation, a thin lens acts as a physical Fourier transformer: placing a coherently illuminated object in the front focal plane produces, in the back focal plane, a complex field proportional to the two-dimensional Fourier transform of the object's transmission, with coordinates (u, v) mapping to spatial frequencies as $(f_x, f_y) = (u/\lambda f,\ v/\lambda f)$, where f is the lens focal length. This principle can be extended to a two-lens (4f) imaging system, where the first lens performs a Fourier transform and the second lens performs its inverse, as illustrated in Figure 2.3. In TEM, the Fourier plane of the objective lens corresponds to the back focal plane of the objective lens. Working in this Fourier plane—or equivalently via a numerical Fourier transform—allows one to read, filter, and manipulate structural information (apertures, phase masks) directly in the frequency domain, forming the foundation of diffraction analysis, contrast-transfer theory, and phase-plate imaging in modern microscopy [48].

Let $\Psi(x, y)$ denote the complex field in the object plane. Its two–dimensional Fourier transform and inverse transform are [49]

$$\mathcal{F}\{\Psi\}(f_x, f_y) = \iint_{-\infty}^{+\infty} \Psi(x, y)\, e^{-i2\pi(f_x x + f_y y)}\, dx\, dy, \tag{2.21}$$

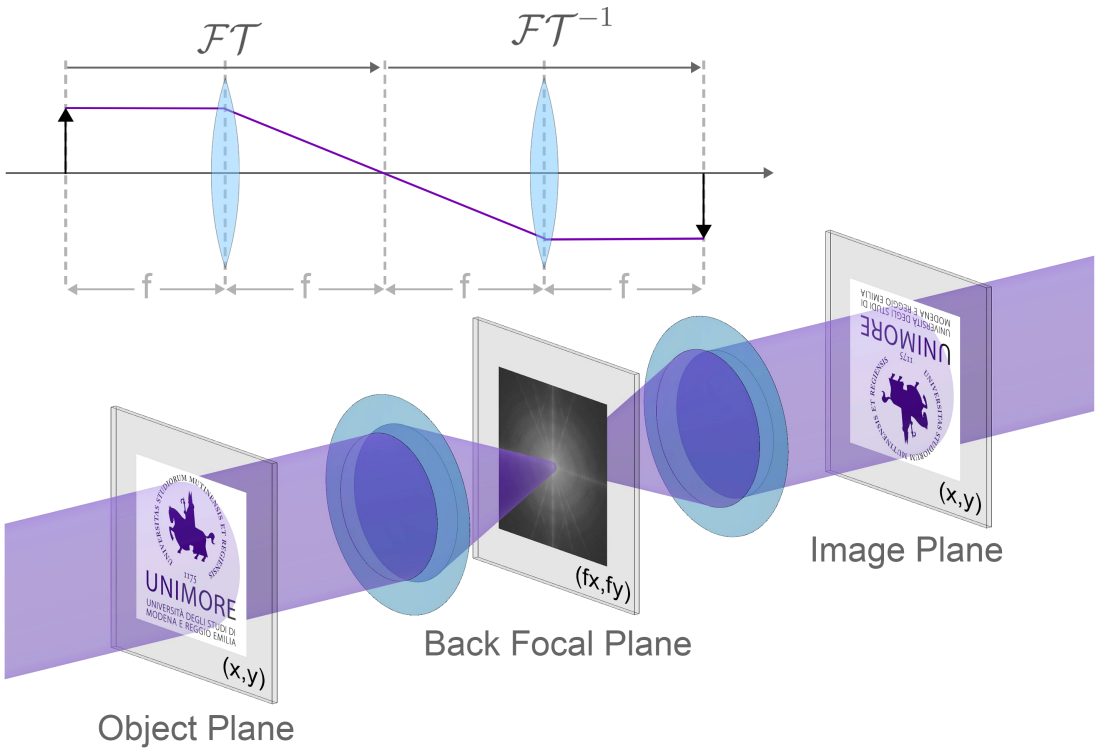


Figure 2.3. Illustration of a two-lens ($4f$) Fourier imaging system. The first lens performs a Fourier transform of the object field, producing its spatial-frequency spectrum in the back focal plane (BFP). The second lens performs an inverse Fourier transform, reconstructing the image in its back focal plane. The intermediate BFP allows direct access to the frequency domain, where filters, apertures, or phase plates can modify selected spatial frequencies. The final image is reproduced with unit magnification ($M = 1$) and a 180° rotation relative to the object, not a mirror inversion.

$$\Psi(x,y) = \iint_{-\infty}^{+\infty} \mathcal{F}\{\Psi\}(f_x, f_y)\, e^{i2\pi(f_x x + f_y y)}\, df_x\, df_y. \tag{2.22}$$

For a linear, shift-invariant imaging system, the image-plane wavefield $\Psi_i(u,v)$ is given by the convolution of the object-plane wavefield $\Psi_o(x,y)$ with the system impulse response $h(u,v)$:

$$\Psi_i(u,v) = \Psi_o(x,y) * h(u,v).$$

By the convolution theorem,

$$\mathcal{F}\{\Psi_i\}(f_x, f_y) = H(f_x, f_y)\, \mathcal{F}\{\Psi_o\}(f_x, f_y),$$

where $H(f_x, f_y)$ is the system transfer function. In coherent imaging, H is directly governed by the pupil function P, so the finite aperture acts as a frequency filter that transmits some spatial frequencies while rejecting others.

This frequency-space viewpoint provides both physical intuition and compact mathematics for image formation. It is the bridge between real-space structures and their spectral content, and it motivates practical workflows in electron microscopy: measure or sculpt the field in the back focal plane to control which spatial frequencies are transmitted; then interpret the

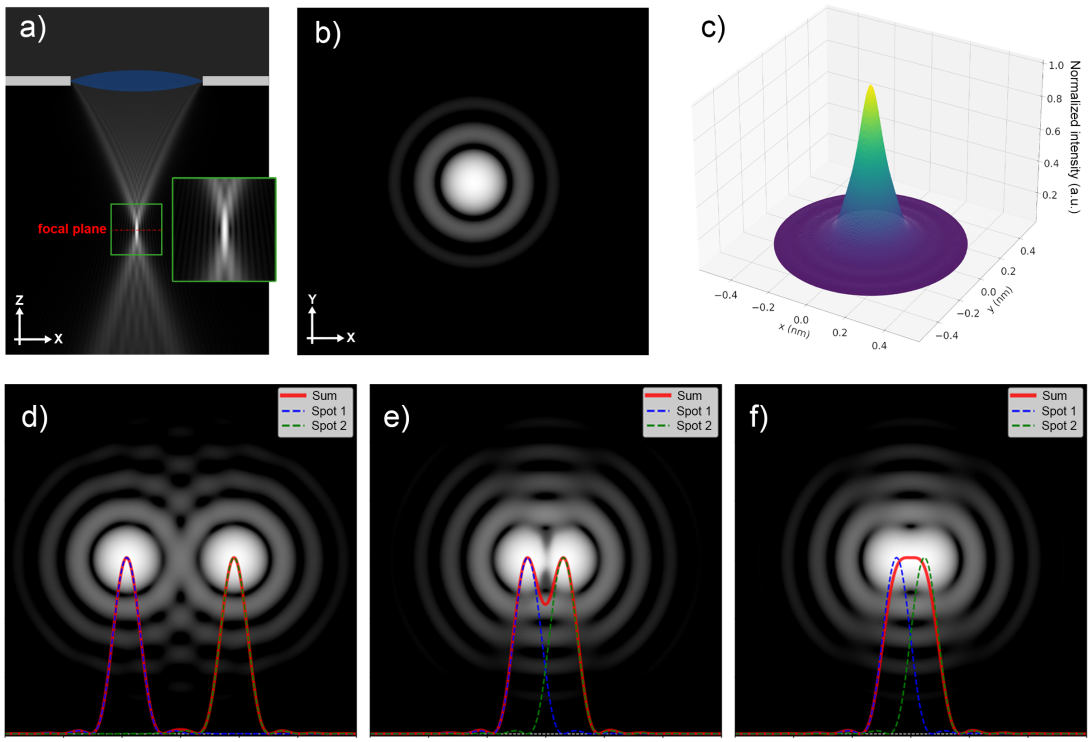


Figure 2.4. Illustration of the diffraction limit set by a finite numerical aperture. (a) Wave propagation in the xz plane showing the probe formation. (b) Airy pattern in the focal xy plane. (c) Three-dimensional representation of the Airy-pattern intensity in the focal plane. (d–f) Summed intensity of two identical Airy patterns as their lateral separation decreases. (f) Diffraction-limited condition (Rayleigh criterion), where the two probes are just resolved and cannot be reliably distinguished in intensity.

resulting image through the corresponding transfer function. In the next section, we connect this framework to the real-space impulse response of the system, the point spread function, which is the Fourier complement of the transfer-function description introduced here.

2.4. Point Spread Function (PSF)

In any imaging system, especially a microscope, the purpose is to map a one-to-one image of an object. The key question is how accurately an optical system can map a single mathematical point in the object plane to exactly one point in the image plane. The answer is governed by diffraction, which prevents any optical system from reproducing a mathematical point exactly; instead, the image of a point source is a three-dimensional intensity distribution. By definition, the *Point Spread Function* (PSF) is the response of an optical system to an idealized point object. The physical origin of the PSF can be explained by the Huygens–Fresnel principle, which states that every point on a wavefront acts as a secondary source of spherical wavelets, whose interference forms the observed intensity pattern. For a perfect circular aperture, the PSF exhibits a bright central peak surrounded by concentric rings known as the Airy pattern (see Figure 2.4). The diameter of the central lobe depends

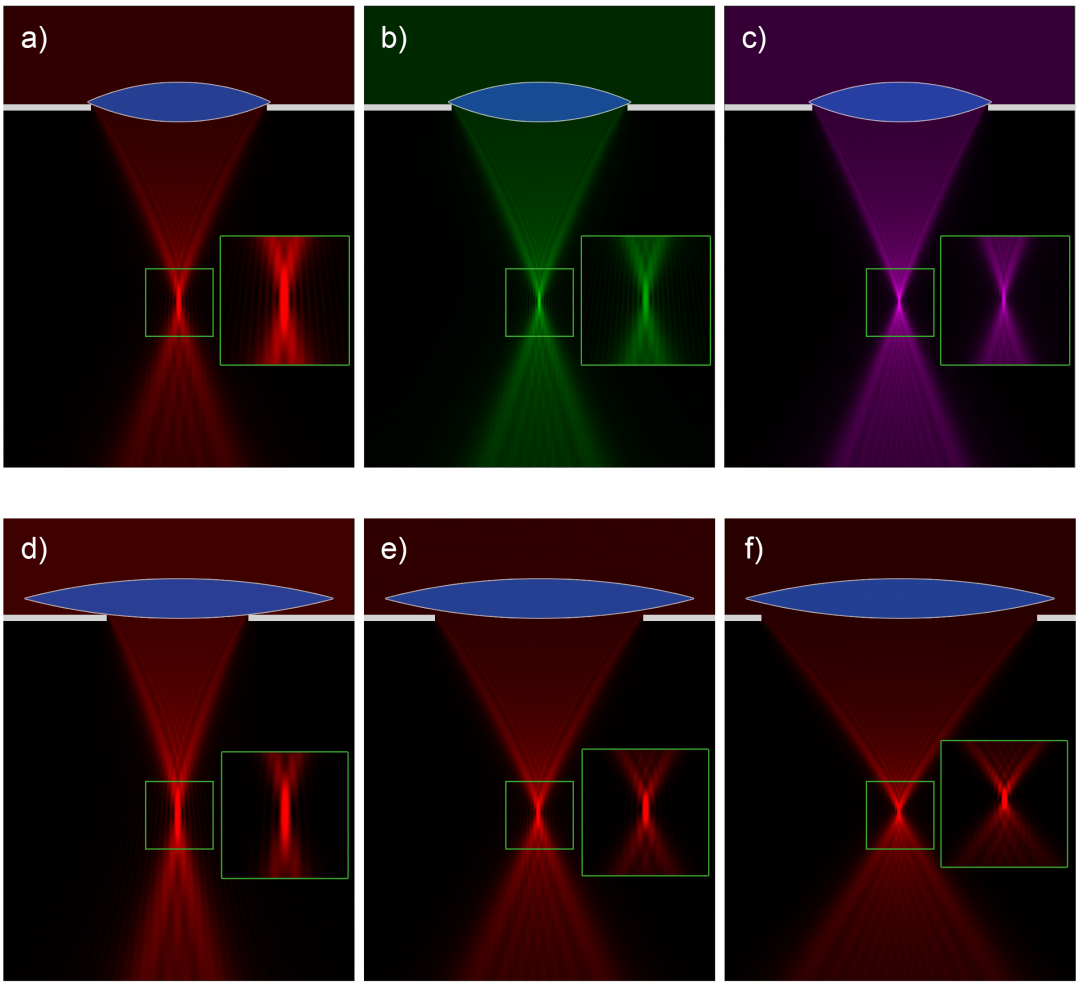


Figure 2.5. Effect of wavelength and numerical aperture on the focused Airy beam. (a–c) Increasing the electron energy (shorter wavelength) reduces the probe size. (d–f) Increasing the numerical aperture (larger aperture angle) further decreases the Airy spot size but also reduces the depth of focus.

on the wavelength and numerical aperture (NA) and is given by:

$$d_{\text{Airy}} \approx \frac{1.22\,\lambda}{\text{NA}},$$

where λ is the wavelength and NA is the numerical aperture. In electron microscopy, the numerical aperture corresponds to the beam convergence semi-angle α.

In microscopes, the lateral resolution is often approximated by the Abbe criterion, $\lambda/(2\,\text{NA})$, which corresponds to half the Airy diameter. Thus, the PSF sets the fundamental resolution limit of optical instruments.

To illustrate these concepts, Figure 2.5 presents optical-range simulations of diffraction-limited focusing by an ideal circular lens for different wavelengths (Figure 2.5 a,b, and c) and numerical apertures (Figure 2.5 d,e, and f). As expected from the Airy formula, shorter wavelengths or larger NAs produce a smaller central spot and narrower PSF, thereby increasing resolution. In electron microscopy, the wavelength is vastly shorter: for 200 keV

electrons, $\lambda \approx 2.5\,\mathrm{pm}$, about 2×10^5 times smaller than visible light. Even with a modest α of 10 mrad, typical in TEM, the diffraction-limited resolution is on the order of 0.1 nm, far beyond the optical limit. These simulations highlight the enormous theoretical resolving power of electron optics, and also that in TEM, the ultimate resolution is rarely set by diffraction, but rather by lens aberrations and other imperfections.

In addition to lateral resolution in the xy-plane, an imaging system is also characterized by its *depth of focus* (DOF), which defines the axial range along z within which the image remains acceptably sharp. In electron microscopy, DOF is inversely proportional to the beam convergence semi-angle α, meaning that modern aberration-corrected systems with large convergence angles achieve exceptional lateral resolution but reduced DOF. This effect is illustrated in Figure 2.6. At exact focus, increasing the convergence semi–angle α reduces diffraction blur and produces a smaller Airy core. However, the same increase in α also makes the probe more sensitive to defocus: even a small Δf causes the point–spread function to broaden and its rings to strengthen. As a result, a larger numerical aperture improves lateral resolution but shortens the depth of focus, consistent with the behaviour discussed above.

The PSF is defined under the assumption that image formation is linear and shift-invariant. The linearity conditions are not always fulfilled in microscopy and electron microscopy because the intensity is the square of the wavefunction that is linearly transferred. In SEM and STEM with ADF detector (see chapter 3) the linearity is induced by the incoherent nature of the imaging. In phase imaging in TEM, this is a useful approximation for these objects, neglecting non-linear terms [43].

Under these conditions, the observed image $I(\mathbf{r})$ is the convolution of the object $O(\mathbf{r})$ with the PSF $h(\mathbf{r})$:

$$I(\mathbf{r}) = (O * h)(\mathbf{r}) = \int O(\mathbf{r}')\, h(\mathbf{r} - \mathbf{r}')\, d\mathbf{r}'.$$

In the spatial frequency domain, this convolution becomes a product: $\tilde{I}(\nu) = \tilde{O}(\nu)\, H(\nu)$, where $H(\nu)$ is the optical transfer function (OTF), the Fourier transform of the PSF. High spatial frequencies above $2\,\mathrm{NA}/\lambda$ are strongly attenuated. Real optical instruments also introduce aberrations and refractive-index mismatches, so the measured PSF may deviate from the ideal Airy pattern; experimentally, it is often obtained using sub-resolution fluorescent beads or point emitters.

In scanning electron microscopy (SEM) or in incoherent forms of scanning transmission electron microscopy (STEM), the PSF describes the spatial distribution of electrons in the probe intensity, determined by diffraction, coherence, and aberrations. In coherent imaging (in the thin object linear approximation) in TEM or STEM with a Bright Field detector (see chapter 3), the PSF is a complex quantity and can be seen as the complex wave function of a diffracted blank beam transmitted through the lens system. Accurate measurement of the electron PSF allows instrument designers to identify lens imperfections and achieve near-diffraction-limited performance. Modern aberration-corrected microscopes use larger aperture angles, which broaden the pupil function in reciprocal space and narrow the real-space PSF, improving lateral resolution at the expense of depth of field.

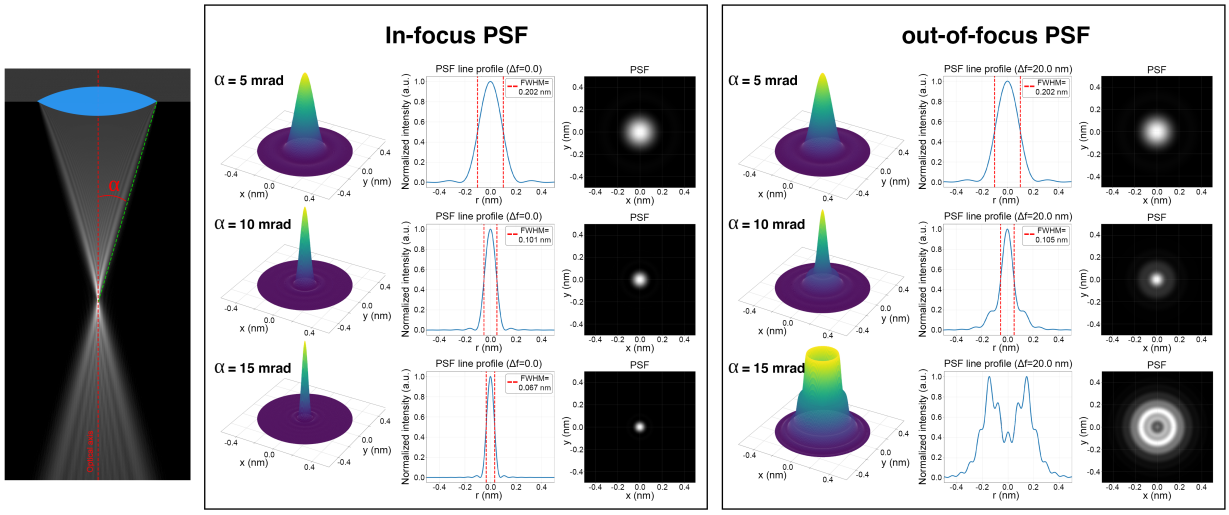


Figure 2.6. Effect of convergence semi–angle α on the point–spread function (PSF) at focus and defocus $\Delta f = 20\ nm$. At exact focus, increasing α (i.e., larger NA $\simeq \sin\alpha \approx \alpha$) reduces diffraction blur, shrinking the Airy core (first zero $r_1 \approx 0.61\,\lambda/\mathrm{NA}$). Under defocus (either sign), larger α makes the PSF broaden and rings strengthen. Thus, a wider aperture yields a smaller in-focus spot but greater sensitivity to Δf.

In the following section, we examine these aberrations in detail and discuss how they limit performance and how modern correction techniques overcome them.

2.5. Aberrations in Electron Microscopy

Similar to any optical system, an electron microscope is not a perfect imaging device. The symmetry and spatial distribution of magnetic fields within electromagnetic lenses prevent these lenses from behaving like the idealized elements known from classical light optics. Nevertheless, a close analogy exists between light and electron optics. In light optics, a lens focuses light by introducing a spatial variation in the refractive index, thereby modifying the optical path length and bending rays toward a focal point. Likewise, in electron optics, magnetic lenses focus electron beams by deflecting their trajectories through the Lorentz force. In the wave-optical description, an ideal thin magnetic lens imparts an approximately quadratic phase to the electron wavefront, analogous to the phase delay introduced by a refractive lens in light optics. However, as in conventional optics, imperfections and physical limitations give rise to aberrations such as spherical and chromatic distortions. These aberrations, originating from fundamental electromagnetic constraints and practical design limits, distort the electron wavefront and consequently degrade both resolution and image quality. The following section examines the physical origins and main characteristics of these aberrations in electron microscopes.

In high-resolution imaging, only a small portion of the sample located near the optical axis is imaged. This approach, known as the *paraxial approximation*, simplifies the analysis of lens behavior and aberrations [41]. As illustrated in Figure 2.7, consider an object plane, a lens, and an image plane. A point P_0 in the object plane is conjugate to the point P_i in

the image plane. For an ideal lens, all rays emerging from P_0 and passing through the lens converge to form an image at P_i. From a wave optics perspective, the same focusing behavior is described by a spherical wavefront, often referred to as the *Gaussian wavefront*. In the case of a real, imperfect lens, however, the wavefront exiting the lens deviates from the ideal spherical surface.

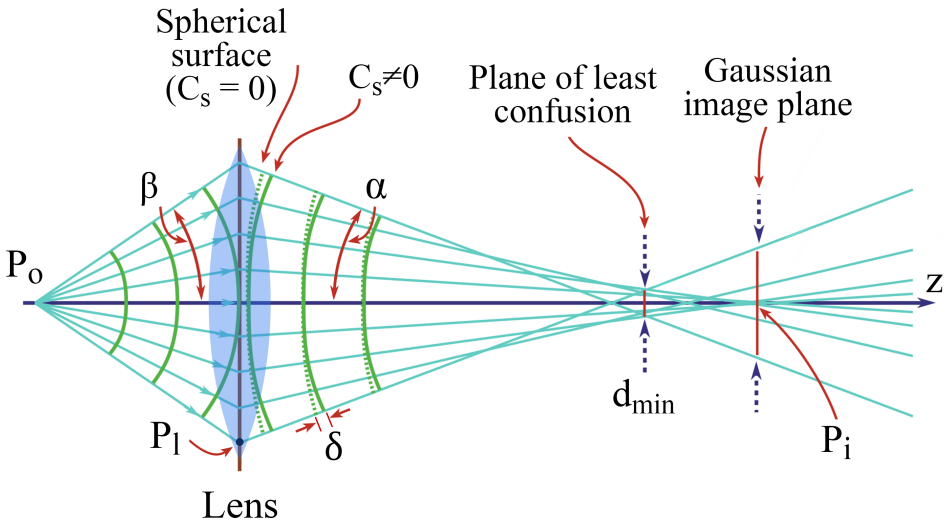


Figure 2.7. Schematic illustration of aberration in an electron optical system (Image credit: [41], modified by the author).

The magnitude of this deviation depends on the object and image positions, as well as on the ray angle β and the location where the electron passes through the lens plane, denoted as point P_l in the figure 2.7. These deviations, collectively referred to as *aberrations*, distort the electron wavefront. The wavefront's slight deviation from a perfect sphere is described by the path-length deviation δ, which produces the corresponding phase error χ [50]:

$$\chi = \frac{2\pi}{\lambda}\delta \tag{2.23}$$

where λ is the electron wavelength.

Physically, aberrations cause rays from a single object point to no longer converge at a single image point. As a result, the image becomes blurred, and high spatial frequencies are distorted by aberration. In Fourier-space terms, aberrations act as phase distortions that modulate the transfer of spatial frequencies through the objective lens transfer function, thereby limiting image contrast and resolution.

The error term δ describing the deviation of the real wavefront from the ideal Gaussian wavefront, can be mathematically expressed using various basis functions. A common and intuitive approach is to expand δ as a power series in terms of positional deviations (x, y) and angular deviations (α_x, α_y) from the optical axis. This provides a systematic and physically meaningful framework for describing different types of aberration, classified according to their

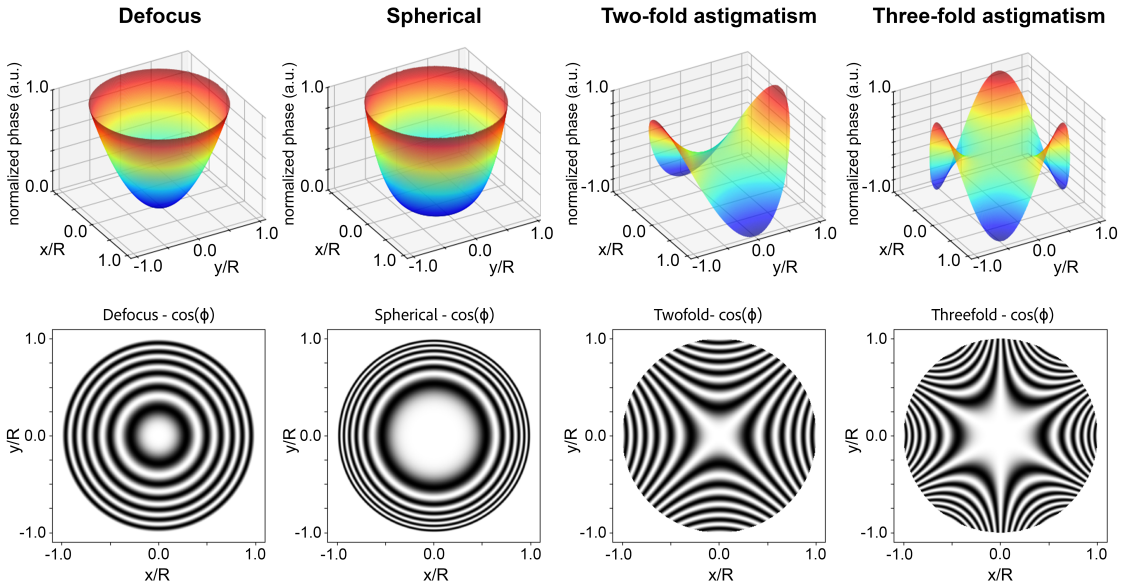


Figure 2.8. Phase profiles of defocus, spherical aberration, and twofold and threefold astigmatism shown in 2D (wrapped phase) and corresponding 3D surface representations.

order and symmetry (limiting to the rotation symmetric case) :

$$\chi(\alpha) = \frac{2\pi}{\lambda}\delta = \frac{2\pi}{\lambda}\left(\frac{1}{2}C_1\alpha^2 + \frac{1}{4}C_3\alpha^4 + \frac{1}{6}C_5\alpha^6 + \cdots\right) \tag{2.24}$$

where the coefficients C_n have units of length. $\Delta f = -C_1$ is the defocus and $C_3 = C_{s3} = C_s$ is the third-order spherical aberration. $C_5 = C_{s5}$ represents the fifth-order spherical aberration, and higher-order terms continue similarly.

The semi-angle into the objective lens, α, which is also the angle between the incident and scattered rays in TEM, is related to the spatial frequency k in the image plane by multiplication with the electron wavelength λ [50]:

$$\alpha = \lambda k \tag{2.25}$$

Thus, the aberration function (in this case considering also non symmetric terms) can also be written in spatial-frequency form [50]:

$$\begin{aligned}\chi(k) = &\frac{\pi}{2}C_s\lambda^3k^4 - \pi\Delta f\lambda k^2 + \pi f_{a2}\lambda k^2\sin\left[2(\phi-\phi_{a2})\right] \\ &+ \frac{2\pi}{3}f_{a3}\lambda^2k^3\sin\left[3(\phi-\phi_{a3})\right] + \frac{2\pi}{3}f_{c3}\lambda^2k^3\sin\left(\phi-\phi_{c3}\right) + \cdots\end{aligned} \tag{2.26}$$

where f_{a2} represents twofold astigmatism, f_{a3} threefold astigmatism, and f_{c3} coma. The angles ϕ_{a2}, ϕ_{a3}, and ϕ_{c3} denote their azimuthal orientations. The corresponding 2D wrapped phase maps and 3D phase profiles of these four main aberrations are shown in Figure 2.8.

The aberration function described above can be generalised to include all higher-order terms

with arbitrary radial order and angular symmetry. Several notation systems exist for this purpose, including the Krivanek notation widely used in aberration-corrector design [51] and the Haider notation commonly employed in high-resolution TEM analysis [52]. In Krivanek notation, the general aberration function is written as

$$\chi(\alpha,\phi) = \frac{2\pi}{\lambda}\sum_{n=0}^{\infty}\sum_{m}\frac{\alpha^{n+1}}{n+1}\left[C_{nma}\cos(m\phi) + C_{nmb}\sin(m\phi)\right], \tag{2.27}$$

with the allowed angular symmetries

$$m = \begin{cases} 0,2,4,\ldots,n+1, & n\,\text{odd}, \\ 1,3,5,\ldots,n+1, & n\,\text{even}. \end{cases}$$

Here, α is the aperture semi-angle, ϕ is the azimuthal angle in the aperture plane, and λ is the electron wavelength. The integer n specifies the radial order of the aberration, while m controls its angular symmetry. This compact representation unifies all commonly encountered aberrations within a single formalism, and is the standard framework used in the design and alignment of modern multipole aberration correctors.

The objective lens modifies the electron wave in Fourier space by adding an aberration-dependent phase shift to each spatial frequency (or, equivalently, to each scattering angle). This effect is described by the objective-lens transfer function

$$H_O(k) = \exp[-i\chi(k)] = \cos[\chi(k)] - i\,\sin[\chi(k)], \tag{2.28}$$

which shows that every spatial frequency is multiplied by a complex factor determined by the aberration phase function $\chi(k)$. In terms of TEM imaging, the cosine term represents the amplitude-contrast component, while the sine term defines the phase-contrast response. The latter forms the basis of the contrast transfer function (CTF), which describes how phase information from the specimen is converted into observable image contrast in TEM [50]. The detailed formulation of the CTF and its implications for imaging performance will be discussed in the following chapter.

2.5.1. Defocus

Defocusing the lens produces a deviation of the electron trajectory that departs from the ideal Gaussian image plane. Defocus can be adjusted either by moving the specimen or by changing the excitation of the objective lens (i.e., the current through its coil). The amount of defocus, Δf, is defined as the axial displacement of the actual image plane from the Gaussian focus. In the wave-optical description, defocus introduces a quadratic phase error in the electron wavefront, proportional to α^2:

$$\chi_{\text{def}}(\alpha) = -\pi\frac{\Delta f}{\lambda}\,\alpha^2. \tag{2.29}$$

Positive Δf corresponds to underfocus (a weaker lens), while negative Δf represents overfocus. Small controlled defocus values are often used to optimize phase contrast in TEM imaging.

2.5.2. Spherical Aberration

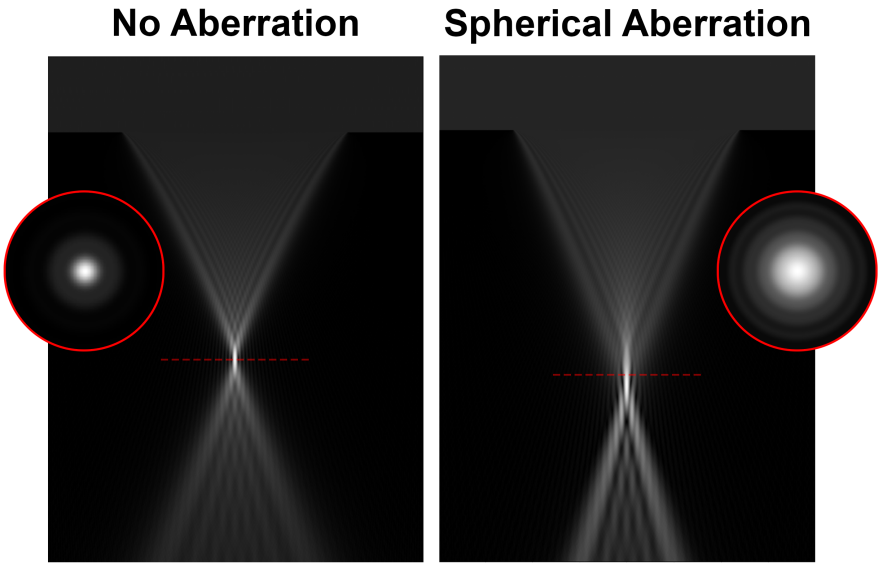


Figure 2.9. Effect of spherical aberration on the point-spread function (PSF). The transverse PSF is taken along the red dashed line.

The most fundamental intrinsic limitation of a rotationally symmetric magnetic lens is *spherical aberration*, denoted C_s. Because the magnetic field increases with radial distance from the optical axis, electrons traveling at larger angles α are overfocused, forming image points closer to the lens. This results in a wavefront phase error proportional to α^4:

$$\chi_{\text{sph}}(\alpha) = \frac{\pi}{2} C_s \frac{\alpha^4}{\lambda} \tag{2.30}$$

According to Scherzer's theorem, a static, rotationally symmetric magnetic lens always has $C_s > 0$, meaning that spherical aberration cannot be eliminated without using multipole correctors [28, 53] . The combined phase error from defocus and spherical aberration is then

$$\chi(\alpha) = \frac{2\pi}{\lambda} \left(\underbrace{-\frac{1}{2}\Delta f \alpha^2}_{\text{Defocus}} + \underbrace{\frac{1}{4} C_s \alpha^4}_{\text{Spherical Ab.}} \right) \tag{2.31}$$

which defines the characteristic shape of the contrast transfer function (CTF) in conventional TEM.

2.5.3. Chromatic Aberration

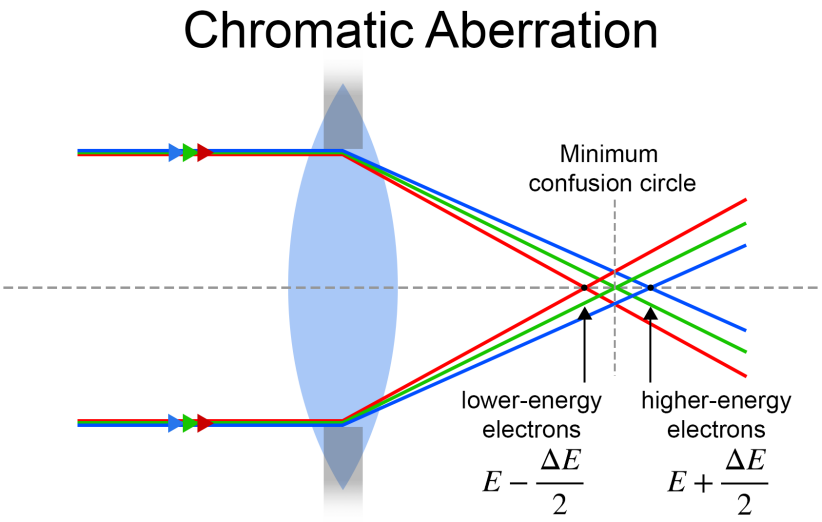


Figure 2.10. Chromatic aberration

Chromatic aberration arises from the finite energy spread of the electron beam. Electrons with slightly higher energy are focused less strongly, while those with lower energy are over-focused, leading to an image blur proportional to $(\Delta E/E)C_c$, where C_c is the chromatic aberration coefficient. This effect reduces image sharpness and contrast, particularly in thick specimens or when energy filtering (monochromator) is not employed [54]. Chromatic aberration is therefore a key factor limiting resolution in both TEM and STEM at high accelerating voltages.

2.5.4. Twofold and Threefold Astigmatism

Deviations from perfect rotational symmetry in the magnetic field produce angularly dependent aberrations. The most common is the *twofold astigmatism* (A_2), which causes the focal length to differ along two perpendicular directions. Consequently, the image appears sharp along one axis and blurred along the other. Twofold astigmatism usually originates from slight imperfections in the machining or alignment of the objective lens pole pieces and can be compensated using stigmators.

Threefold astigmatism (A_3) is a higher-order aberration with a characteristic threefold angular dependence. It distorts the probe or image into a three-lobed pattern and becomes increasingly significant at sub-ångström resolutions. This aberration typically originates from subtle asymmetries in the lens field or from higher-order multipole components, and it cannot be corrected using conventional stigmators. Instead, dedicated multipole correctors—most commonly hexapoles—are required to compensate for A_3.

Because a major part of this thesis is devoted to the analysis, design, and implementation of threefold-astigmatism correctors based on hexapole elements, a more detailed discussion of

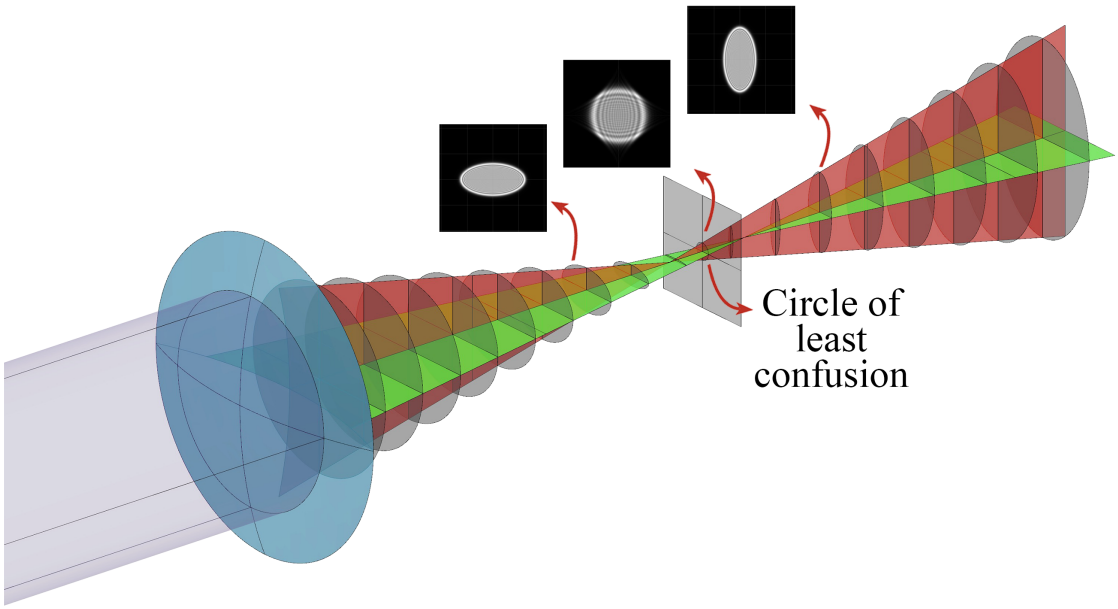


Figure 2.11. Schematic illustration of two-fold astigmatism. Anisotropic focusing produces two orthogonal line foci, indicated by the x-plane (green) and y-plane (red). Intensity cross-sections are shown at underfocus, at the circle of least confusion, and at overfocus, illustrating the spot evolution along the optical axis.

the physics and correction strategies associated with A_3 is deferred to later chapters.

2.5.5. Combined Effects and Correction

In practice, the combined effects of defocus, spherical aberration, and astigmatism determine the microscope's overall resolution and contrast transfer. While slight defocus adjustments can partially compensate for spherical aberration over a limited range of spatial frequencies (the so-called Scherzer defocus), complete correction requires multipole aberration correctors that introduce negative spherical aberration to counteract the positive C_s of conventional round lenses. In the next chapter, we will discuss aberration correctors in electron microscopy.

2.6. Chapter Summary

In this chapter, we established the physical foundations of electron optics as they apply to transmission electron microscopy. We began by reviewing the relativistic motion of high-energy electrons and showed how their short de Broglie wavelength enables sub-nanometer resolution. We then developed both the classical and wave-mechanical descriptions of electron propagation, covering the Lorentz force, the paraxial Schrödinger equation, and the influence of electromagnetic potentials on phase evolution, including the Aharonov–Bohm effect.
Next, we introduced the basic principles of Fourier optics and explained how electron lenses act as Fourier transformers that control the amplitude and phase of the wave. Concepts such as the point spread function, diffraction-limited resolution, and depth of focus were used to link real-space imaging behaviour with spatial-frequency descriptions. Finally, we examined

the origin and mathematical form of lens aberrations—from defocus and spherical aberration to higher-order terms such as astigmatism and coma—using their standard expansion in aberration theory.

Together, these elements form the theoretical framework needed to understand how electron beams propagate, accumulate phase, and form images in a TEM. In the next chapter, we will turn to the architecture of the transmission electron microscope and discuss how its optical components work together to form a practical imaging system.

3 Architecture and Components of a TEM

Contents

3.1 Overview of TEM Layout 36

3.2 Electron microscope components 37

3.2.1 Electron Gun 37

3.2.2 Electron Lenses 40

3.2.3 Apertures and Holders 43

3.3 Transmission Electron Microscope Architecture 45

3.3.1 Illumination system 45

3.3.2 Objective System 47

3.4 Diffraction in TEM 48

3.5 Bright-Field and Dark-Field Imaging in TEM 49

3.6 TEM vs. STEM 50

3.7 Aberration Correctors and Stigmators 53

3.7.1 Double hexapole C_s aberration correctors 54

3.7.2 Quadrupole–Octopole C_s Correctors 57

3.8 Aberration Detection and Measurement in TEM 59

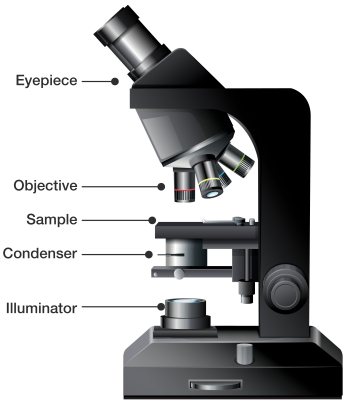
3.8.1 Diffraction-Based Aberration Diagnosis (Zemlin tableau) 59

3.8.2 Real-Space Aberration Detection Using Ronchigrams 59

3.1. Overview of TEM Layout

In Chapter 2, we introduced the basic principles of electron optics. In this chapter, we focus on the main structural components of the electron microscope, with particular emphasis on the Transmission Electron Microscope (TEM). Similar to any optical imaging system, a TEM consists of several functional subsystems that work together to generate, manipulate, and detect an electron beam.

To better understand the structure of a TEM, it is helpful first to consider a conventional optical microscope. In an optical microscope, a light source illuminates the sample by passing through condenser lenses that shape the beam, either by focusing or collimating it. Once the light interacts with the specimen, it is transmitted, absorbed, or scattered. The beam is then collected by the objective lens, which forms a magnified intermediate image. The eyepiece lens further enlarges the image for direct detection by the human eye or for digital recording using a CCD or CMOS camera.



TEM follows a conceptually similar architecture. As illustrated in Figure 3.1, a TEM can be divided into three major subsystems:

- **Illumination system** — Generates and conditions the beam. Includes the electron gun (thermionic or FEG), extractor/anode, optional monochromator, and the condenser lenses and apertures that set coherence and convergence.
- **Objective system** — Positions and manipulates the sample (stage/holder, *in situ* options). The objective lens forms the first image and diffraction pattern, while the objective and selected-area apertures set contrast and collection.
- **Imaging system** — Intermediate and projector lenses provide magnification or camera length for diffraction. Post-projection optics (an energy filter for EELS/EFTEM) and detectors (a fluorescent screen, a CCD/CMOS, and a direct electron detector) record the signal.

The following sections examine each subsystem in detail to provide a clear view of the TEM architecture. Before describing the microscope from top to bottom, we introduce its main components, including the electron source, magnetic lenses, apertures, and sample holders.

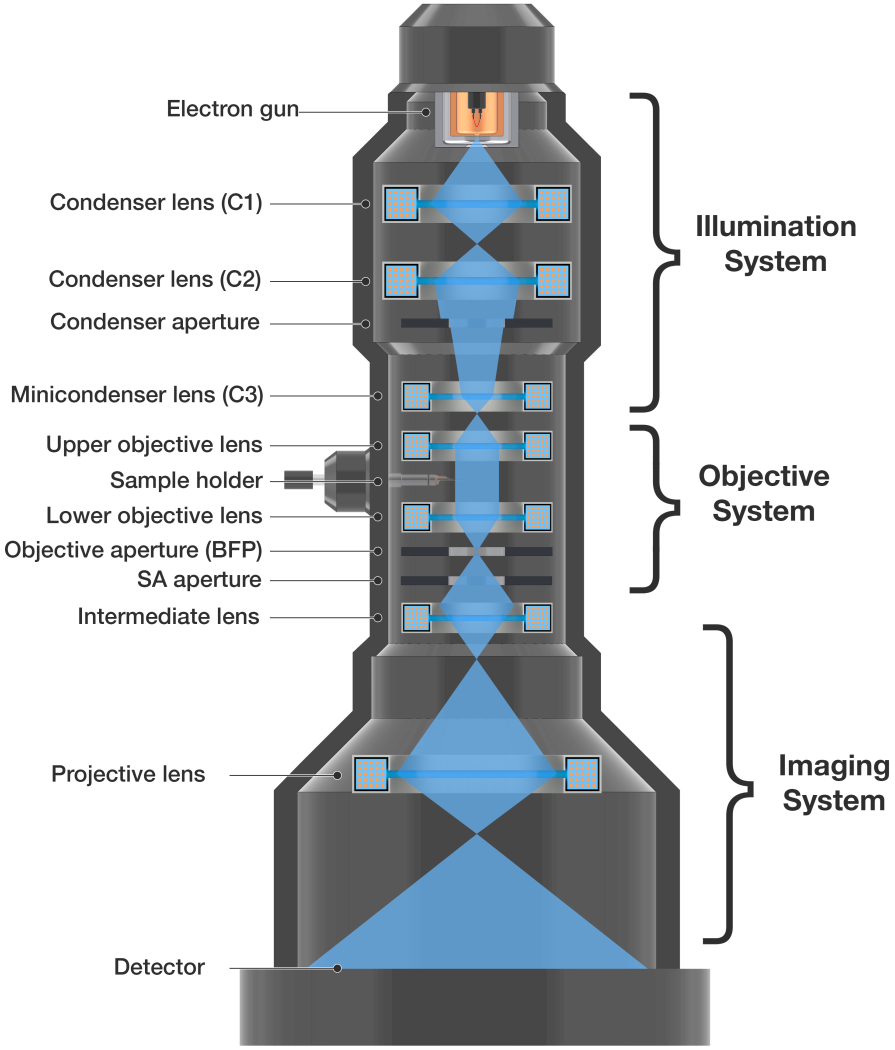


Figure 3.1. Overview of Transmission Electron Microscope (TEM)

After establishing these building blocks, we outline how they are assembled into a complete instrument. The chapter then provides an overview of the main operational modes — TEM, STEM, and diffraction — that define how the microscope is configured for different imaging and analytical tasks. Finally, the last section discusses the aberration correctors and diagnosis.

3.2. Electron microscope components

3.2.1. Electron Gun

In transmission electron microscopy (TEM), electrons serve as the primary illumination source. To achieve a coherent and energetic electron beam (i.e., with a sufficiently short wavelength), electrons must first be generated and then accelerated to the desired energy. This process is carried out by the electron gun, one of the most critical components of the microscope. The electron gun must meet two essential criteria [41]:

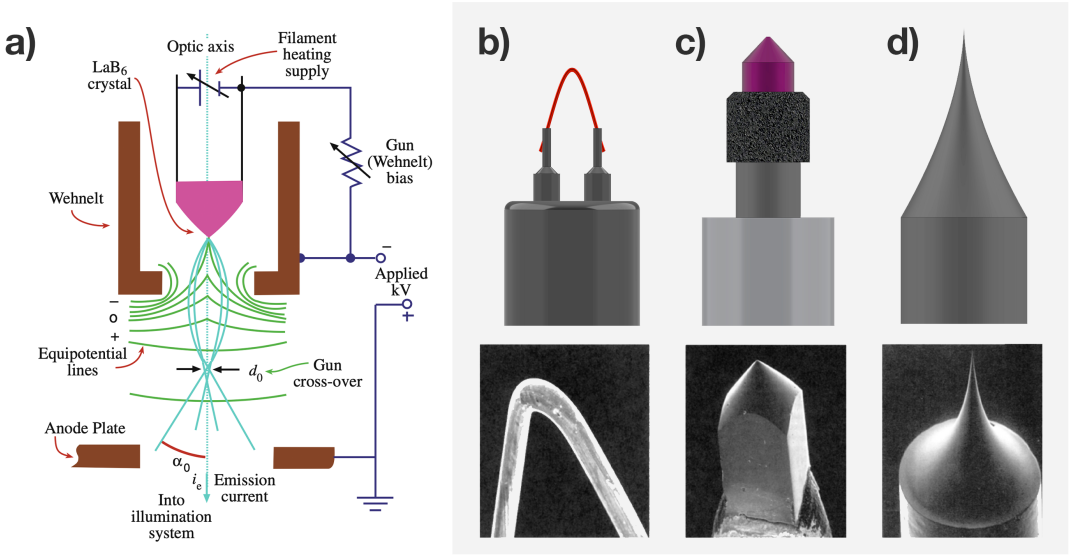


Figure 3.2. Electron gun and common electron sources. (a) Electron gun with a Wehnelt cylinder. (b) Tungsten filament. (c) LaB_6 cathode. (d) Field-emission (FEG) tip [41].

1. **Spatial Coherence and Source Size** Spatial coherence refers to the phase correlation across the wavefront at various points in space. For electron guns, it is predominantly determined by the effective source size and the angular spread of emitted electrons. Smaller sources and narrower angular distributions result in higher spatial coherence, enabling high-resolution imaging and well-defined diffraction.

2. **Temporal coherence (energy spread)** — Temporal coherence in TEM is determined by the energy spread of the source (ΔE), which creates a distribution of electron wavelengths that degrades the phase contrast. A smaller energy spread gives a longer temporal coherence length and reduces chromatic effects. This benefits phase-sensitive imaging and determines the achievable energy resolution in EELS.

To achieve both energy and spatial coherence, electrons must be accelerated uniformly along the optical axis.
Two main types of electron sources are commonly used in TEM[41] :

- **Thermionic sources** emit electrons through thermal excitation when a material (typically tungsten or LaB_6) is heated.

- **Field emission sources** generate electrons by applying a strong electric field between the emitter and an anode, leading to the quantum tunneling of electrons from the tip of a sharp emitter.

Thermionic emission occurs when electrons in a heated metal gain enough energy to overcome the work function and escape into vacuum [55]. The emitted current density at the cathode

surface is commonly described by the Richardson–Laue–Dushman (RLD) equation [56]:

$$J = AT^2 \exp\left(-\frac{W}{k_B T}\right), \tag{3.1}$$

where J is the current density, A is the Richardson constant (material dependent, and typically of the order of $10^6\,\mathrm{Am^{-2}K^{-2}}$), W is the work function of the material, k_B is the Boltzmann constant, and T is the absolute temperature. In practice, thermionic sources use refractory materials with high melting points, and low effective work functions are preferred to reach useful emission at moderate temperatures. Tungsten filaments were historically common, while many modern instruments use LaB_6 emitters because they provide higher brightness and improved stability compared to tungsten [41].

A thermionic gun typically includes a *Wehnelt cylinder* that surrounds the emitter and shapes the extraction field (Figure 3.2 (a)). By biasing the Wehnelt more negative than the cathode, the emission is confined to a small effective area near the aperture, which improves beam formation and helps control the emitted current [57].

Field emission guns (FEGs) rely on a strong electric field at a sharp tip, which lowers the surface barrier and allows electrons to tunnel into vacuum [58]. Cold field emission provides very high brightness and a narrow energy spread, but it requires a very clean emitter tip and a high-quality vacuum to maintain stable operation [58]. To improve stability in routine operation, many TEMs use *Schottky* (thermal field-emission) sources, in which a ZrO_2-coated tungsten tip is operated at elevated temperature to reduce the effective work function and provide more stable emission under less demanding conditions [41].

Table 3.1 summarizes the most commonly used electron sources in modern electron microscopes [41].

Table 3.1. Comparison of electron sources [41].

Property	**Unit**	**Tungsten**	**LaB_6**	**Schottky FEG**	**Cold FEG**
Work function, ϕ	eV	4.5	2.4	3.0	4.5
Richardson constant, A_{eff}	A/m^2K^2	6×10^9	4×10^9	-	-
Operating temperature	K	2700	1700	1700	300
Current density	A/m^2	5	10^2	10^5	10^6
Crossover size	nm	$>10^5$	10^4	15	3
Brightness* (at 100 kV)	A/m^2sr	10^{10}	5×10^{11}	5×10^{12}	10^{13}
Energy spread (at 100 kV)	eV	3	1.5	0.7	0.3
Vacuum requirement	Pa	10^{-2}	10^{-4}	10^{-6}	10^{-9}
Lifetime	hr	100	1000	>5000	>5000

* Brightness is beam current per unit area per unit solid angle (not total current). High brightness enables smaller probe sizes and improved spatial resolution, which is particularly critical in high-resolution transmission electron microscopy (HRTEM) and scanning transmission electron microscopy (STEM) applications.

3.2.2. Electron Lenses

Before introducing magnetic lenses, it is helpful to first review the basic principles of light optics. In visible light microscopes, lenses are used to control the wavefront of light. Optical lenses are made from materials with a refractive index different from the surrounding medium (typically air). Their focusing behavior can be understood in two complementary ways.
From the *ray optics* perspective, based on Snell's law, light rays are refracted at different points on the lens surface and bent toward a common focal point. Alternatively, in the *wave optics* framework, the lens imposes a spatially varying phase delay on the wavefront. This delay transforms a planar wavefront into a spherical one, resulting in focusing.
Convex lenses are most commonly used in optical systems. In ray optics, each ray represents the path of a photon, and ideal lenses obey several key principles [6]:

1. All rays emerging from a point on the object are focused to a corresponding point in the image plane.
2. All rays parallel to the optical axis are focused to a single point, called the focal point.
3. Rays originating from the focal point emerge as parallel rays after passing through the lens.

Another important property is *reversibility*: light rays can be traced in the opposite direction (interchanging object and image), and the system's optical behavior remains unchanged.
In practice, however, perfect lenses do not exist. Limitations such as the finite size of the entrance aperture (also referred to as the entrance pupil) restrict the amount of light collected. Additionally, manufacturing imperfections prevent the fabrication of an ideal spherical surface, introducing optical aberrations.
Electron lenses are conceptually similar to optical lenses, but with important differences. Unlike photons, electrons are charged particles, thereby interacting more strongly with matter. Therefore, electron lenses must operate in vacuum and are typically aperture-like structures rather than bulk material devices such as a shaped glass. The focusing of electron beams relies on electromagnetic forces, exploiting the charge of the electrons (see also Section 2.1.2).
Electron lenses can be classified into electrostatic, magnetic, and hybrid types, depending on the nature of the fields used to manipulate the electron beam. In transmission electron microscopy (TEM), magnetic lenses are predominantly employed due to their stronger focusing capabilities and lower chromatic aberration [59].
Electrostatic lenses, which utilize electric fields generated by electrodes, can also focus electron beams. Simple configurations, such as Einzel lenses, as illustrated in figure 3.3, achieve axial focusing through radial electrostatic forces. However, their application is limited by high-voltage requirements and relatively strong aberrations. As a result, electrostatic lenses are mainly used in electron guns and low-energy beam systems, rather than for imaging in TEM [57]. Therefore, the rest of this section will focus on the design and function of magnetic lenses, which are essential components in modern TEMs.

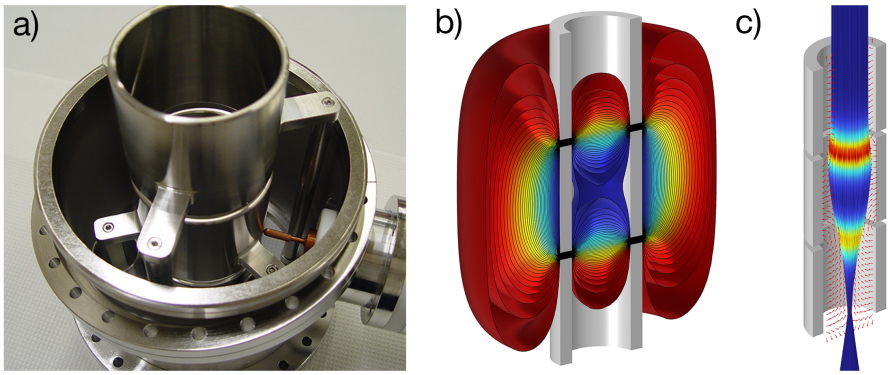


Figure 3.3. Electrostatic electron lens configurations and simulations. (a) Image of a three-electrode Einzel lens device. (b) Electrostatic simulation showing the electric potential distribution generated by the three cylindrical electrodes. (c) Simulation of electron trajectories through the lens; red arrows represent the local electric field vectors.

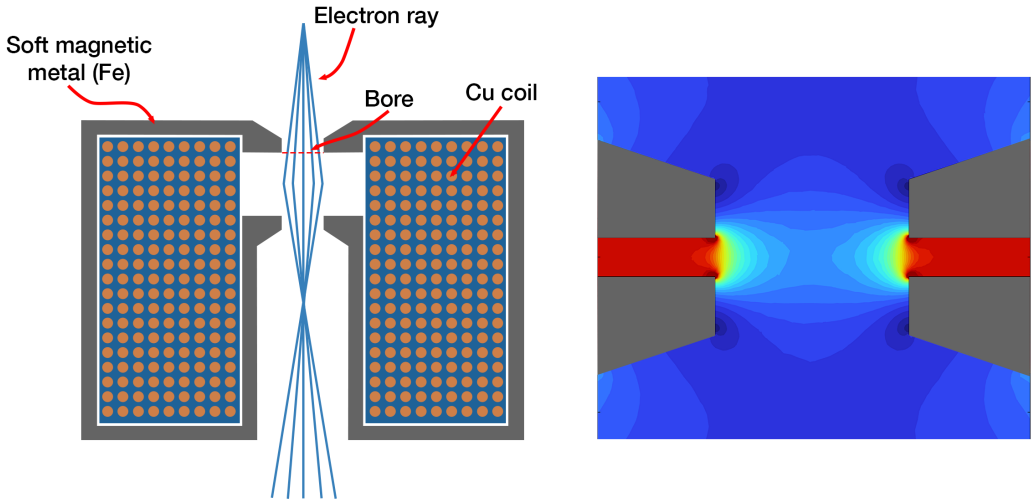


Figure 3.4. Left: simplified schematic showing the main structural components of a magnetic lens. Right: inhomogeneous magnetic field distribution generated near the pole pieces, illustrating the field confinement within the lens gap.

To focus an electron beam, magnetic lenses are commonly used. These lenses generate a magnetic field by passing current through a coil of wire wound around a ferromagnetic core. For effective focusing, the lens must have cylindrical symmetry to produce a symmetric magnetic field.

As illustrated in Figure 3.4, a magnetic lens consists of two main parts. The first is a cylindrically symmetric core made from a soft magnetic material, such as iron, with a drilled hole in the middle. The core is referred to as the polepiece, and the hole within it is called the bore of the polepiece. The second part of the lens is a coil of copper. When current passes through the coil, a magnetic field is created in the bore. The field is inhomogeneous along the electron path but axially symmetric [41]. Due to the high current in the coil, the lens must be cooled down by water circulation.

Unlike optical lenses, where focal length and magnification are fixed by the lens geometry

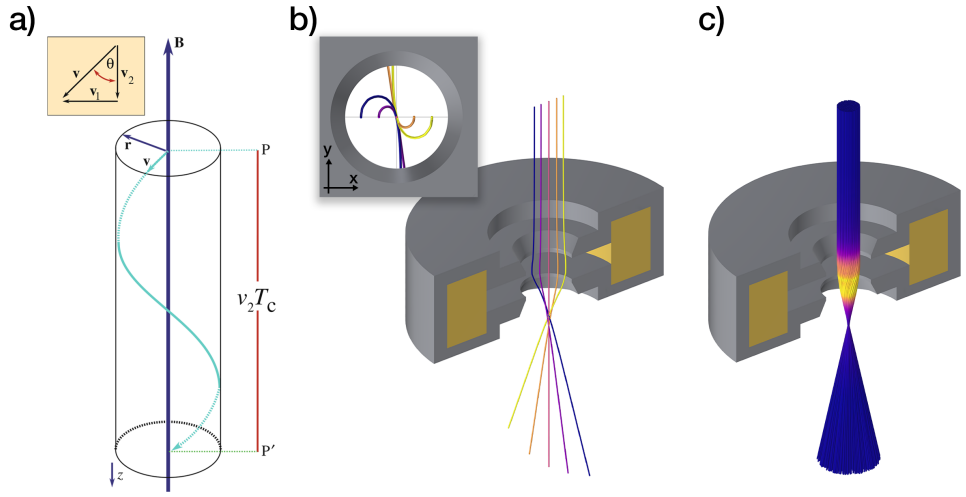


Figure 3.5. (a) Spiral motion of an electron in a magnetic field. (b) Simulation of five individual electron trajectories showing stronger bending for off-axis rays due to higher magnetic field strength (spherical aberration). (c) Simulation of 10 000 electron trajectories illustrating overall beam focusing behavior.

and refractive index, the focal properties of magnetic lenses can be dynamically adjusted. This is achieved by varying the current through the coil, which alters the lens strength. In TEM, the focusing, illumination intensity, and magnification are all controlled by tuning the strength of the magnetic lenses.

As discused earlier in section 2.1.2, electrons in an inhomogeneous magnetic field experience the Lorentz force

$$\mathbf{F} = -e\,(\mathbf{v} \times \mathbf{B}) \tag{3.2}$$

leading to a spiraling motion as depicted in Figure 3.5 (a). This is because in a static magnetic field this force is perpendicular to both $\mathbf{v}$ and $\mathbf{B}$. Thus, the electron speed remains constant while its direction changes. Inside a magnetic lens, both the axial and radial field components vary along z, so $\mathbf{v} \times \mathbf{B}$ changes continuously and bends trajectories toward the optical axis, resulting in a focusing effect.

For electrons traveling exactly along the optical axis, the net Lorentz force is zero, allowing them to pass through the magnetic lens without deflection. In contrast, electrons with off-axis trajectories experience more complex motion (see figure 3.5 (b)). Under the paraxial approximation, when electrons enter the magnetic field of a magnetic lens, they follow a spiral trajectory. As they move forward, the angle between their velocity vector and the magnetic field changes, causing the spiral radius to gradually decrease. This convergence of trajectories toward the optical axis constitutes the focusing effect of the lens.

The rate of change in radius along the z-axis is governed by the following expressions [39]:

$$\frac{d^2r}{dz^2} + \frac{\eta^2 B^2 r}{2\sqrt{U}} = 0, \tag{3.3}$$

$$\frac{d\theta}{dz} = \frac{\eta B}{2\sqrt{U}}, \tag{3.4}$$

where U is the accelerating voltage of the microscope and $\eta = \sqrt{\frac{e}{2m_e c^2}} = 9.88 \times 10^{-4}\,\mathrm{m\,A/\sqrt{kg}}$ is a constant from fundamental physical parameters. These relations show that the radial rate strengthens with the magnetic field B and weakens with the accelerating voltage U. Any change in U alters the electron velocity and thus the focal length of the magnetic lens. Therefore, the magnetic field must be adjusted to maintain proper focus.

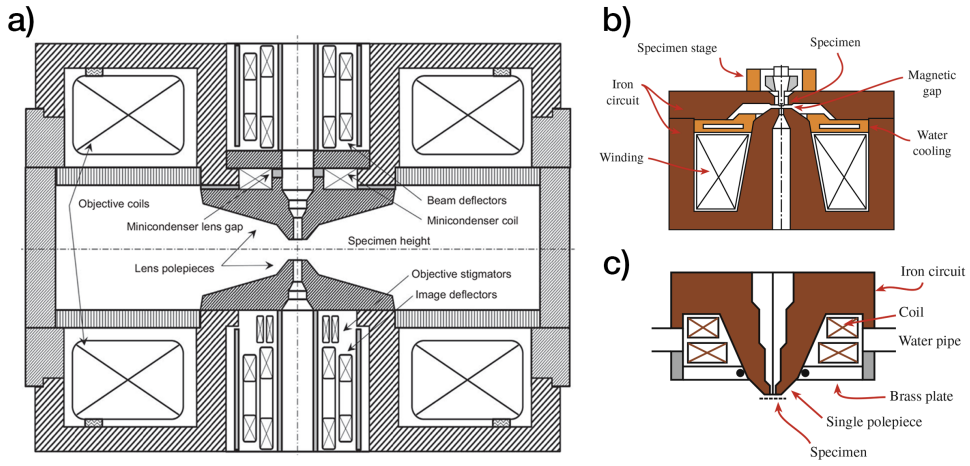


Figure 3.6. Examples of magnetic lenses used in electron microscopy. (a) Objective lens region in a modern TEM (FEI Talos). (b) Top-entry immersion lens, providing strong axial field at the specimen. (c) Snorkel lens with extended gap [41].

A transmission electron microscope employs several types of magnetic lenses, each designed and optimized for a specific function within the column. Condenser lenses control the beam convergence and illumination angle, the objective lens provides the main image formation and determines the ultimate resolution, while intermediate and projector lenses magnify and transfer the image to the detector plane (also see figure 3.1). When high spatial resolution is required, the focal length of the objective lens must be minimized, which demands a strong magnetic field and, consequently, a more powerful lens design. As illustrated in Figure 3.6, real magnetic lenses are far more complex than the simplified schematics, incorporating optimized geometries and materials to reduce aberrations and improve field uniformity.

3.2.3. Apertures and Holders

Apertures

Another essential component in electron microscopes is the aperture: a small, precisely machined hole in a metallic disk placed along the beam path to control the angular distribution of electrons. By blocking high-angle electrons, apertures limit beam current, improve spatial coherence, and define the illuminated area. In TEM, the objective aperture removes high-angle scattered electrons from the image, thereby enhancing contrast. In STEM, condenser apertures set the probe current and the convergence semi-angle.

The precision of aperture fabrication is critical. Commercial apertures are drilled or etched with smooth edges to minimize unwanted scattering. During operation, contamination can

accumulate at the aperture rim, typically as a carbonaceous film formed from residual hydrocarbons under electron irradiation [60]. This buildup can distort the effective aperture shape, introduce charging, and contribute to astigmatic effects [41]. For this reason, apertures are often heated to promote desorption and maintain stable beam conditions.
Aperture materials and geometries depend on their position and function in the column. Common materials include platinum–iridium alloys and molybdenum. Typical hole diameters range from a few micrometers to a few hundred micrometers. Specialized designs include strip apertures selected by mechanical translation and thin-film apertures (e.g., gold-foil types), which can reach micrometer-scale openings but are mechanically fragile [61]. Apertures are placed at several planes along the microscope column, each serving a distinct role. The most common TEM apertures are the condenser aperture, objective aperture, and selected-area aperture, while other apertures (e.g., Wehnelt or phase-plate apertures) serve more specialized purposes.

Sample holders

A sample holder is the mechanical interface between the specimen (e.g., a grid or a chip) and the microscope column. It positions the sample at the eucentric height and allows controlled tilt and translation. In many cases, the holder also enables in-situ conditions such as heating, electrical biasing, or gas and liquid environments.
TEM holders are typically side-entry rods constructed from non-magnetic materials such as stainless steel or titanium, chosen for mechanical stability, thermal isolation, and compatibility with high-vacuum operation [41]. The available tilt range and operating conditions depend on the microscope geometry, in particular the objective-lens pole-piece gap.
For this thesis, *electrical biasing holders* are of special interest because they provide multiple electrical feedthroughs to apply controlled voltages and currents to MEMS-based phase plates inside the TEM. These holders enable direct tuning of the electrostatic potential distribution on the chip during imaging, which is essential for programmable wavefront shaping and in-situ device characterization.
When selecting a TEM holder, several practical factors must be considered [41]:

1. **Tilt range and pole-piece gap** — The achievable tilt range is limited by the pole-piece geometry. High-tilt designs require narrow tips to avoid contact with the objective lens, which can impose constraints on specimen size and holder layout.

2. **Thermal stability and drift** — For heating or cryogenic experiments, thermal stability is critical. Poor thermal control can increase drift and degrade resolution during long acquisitions.

3. **Safety and sample compatibility** — The holder must match the sample's mechanical, chemical, and electrical properties. This is especially important for biasing and environmental holders to avoid damage to the specimen and the microscope.

3.3. Transmission Electron Microscope Architecture

Up to this point, we have discussed the main components of a TEM individually — the electron gun, magnetic lenses, apertures, and specimen stage — and described how each works in principle. We are now ready to place these elements together to see how they form a complete microscope.

Although specific designs vary depending on manufacturer and application, most modern TEMs share a similar arrangement of components along a vertical column. Figure 3.1 shows a schematic of a typical TEM, with its major functional sections arranged from the top of the column to the bottom.

At the very top is the **electron gun**, which generates a coherent beam of electrons and accelerates them to the desired energy. This beam then enters the **illumination system**, which uses condenser lenses and apertures to shape the beam according to the required imaging or analytical mode. The illumination can be adjusted to produce a nearly parallel beam, ideal for conventional TEM imaging and selected-area diffraction, or a convergent beam, required for STEM imaging, X-ray microanalysis, and convergent-beam electron diffraction (CBED).

The electrons next encounter the **objective system**, located around the specimen stage. This is the most critical part of the TEM: here, the electron beam interacts with the sample, and the **objective lens** forms both the initial image (in its image plane) and the diffraction pattern (in its back focal plane). An **objective aperture**, placed in a plane conjugate to the back focal plane, can be inserted to enhance contrast by limiting the range of scattering angles that contribute to image formation.

The next subsystem is the **imaging system**, which consists of intermediate and projector lenses. In TEM mode, by changing their excitation, the operator can project either the intermediate image or the diffraction pattern onto the viewing screen or detectors. This flexibility allows rapid switching between imaging mode and diffraction mode without altering the sample or electron optics upstream.

Finally, at the base of the column, the magnified image or diffraction pattern is recorded by the detector system, which may include a fluorescent viewing screen, CCD/CMOS camera, direct electron detector, or specialized STEM/analytical detectors.

This top-to-bottom arrangement enables a TEM to operate in multiple modes — imaging, diffraction, and STEM — within the same instrument. The following sections will examine each of these components in detail, explaining their design, function, and role in the performance of the microscope.

3.3.1. Illumination system

The illumination system comprises the electron gun and a set of magnetic lenses that prepare the beam at the specimen (sample) plane in two modes (see Figure 3.7):

1. **TEM:** quasi-parallel illumination,
2. **STEM:** convergent illumination.

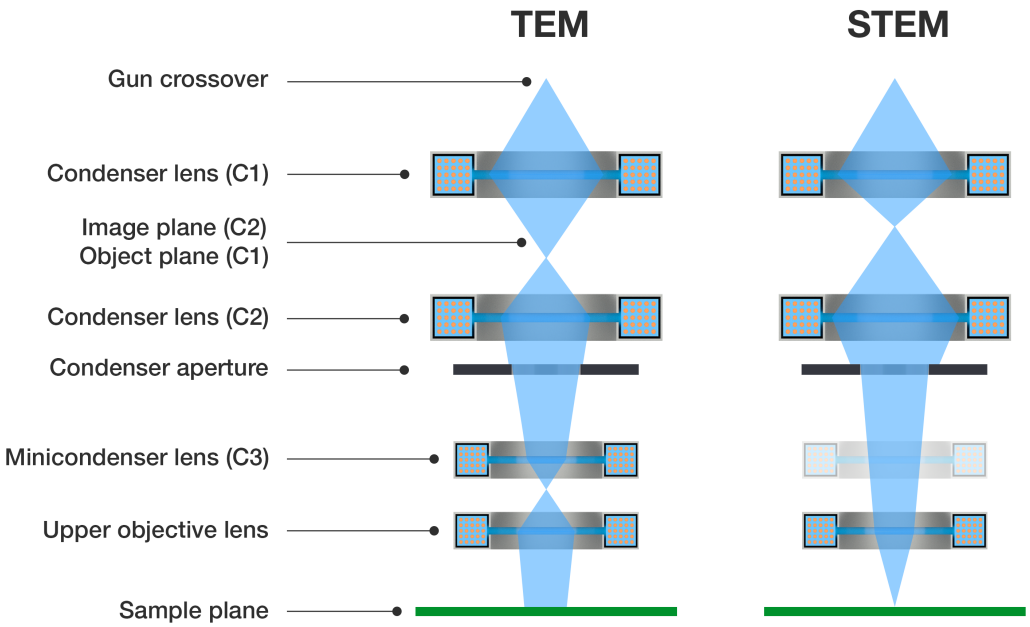


Figure 3.7. Schematics comparing TEM and STEM imaging modes.

Column layouts vary by instrument. Here, we describe a common arrangement with two condenser lenses (C1, C2), a mini-condenser (objective pre-field), and the upper part of the objective lens before the specimen.

C1 is a strong lens with a focal length of a few millimeters ($f_{C1} \sim 1-5\,\text{mm}$). It images the virtual source (gun crossover) located far upstream ($p \sim 100-300\,\text{mm}$, so $p \gg f$). From the thin-lens relation $1/f = 1/p + 1/q$, we obtain $q \approx f$ and a small magnification $M = q/p \approx f/p \ll 1$. For example, with $p = 200\,\text{mm}$ and $f = 2\,\text{mm}$, $q \approx 2.02\,\text{mm}$ and $M \approx 0.010$. Thus, C1 strongly demagnifies the source image and sets the downstream current density (while brightness is conserved).

The second condenser lens (C2) is relatively weaker, with a focal length on the order of centimeters ($f_{C2} \sim 10$–$50\,\text{mm}$), and typically operates near unity magnification. It conditions the beam for the desired illumination. The condenser (C2) aperture, with a diameter of $D \sim 10-200\,\mu\text{m}$, is placed in or near the C2 region and blocks high-angle rays. By doing so, it sets the convergence semi-angle α at the specimen: small α ($\sim 0.1-5\,\text{mrad}$) for TEM quasi-parallel illumination, larger α ($\sim 5-30\,\text{mrad}$, up to $\sim 60-100\,\text{mrad}$ with correctors) for STEM.

A final lens element above the objective—the mini-condenser (objective pre-field)—refines the beam and directs it to the upper objective field, which defines the illumination at the specimen plane. This stage provides fine control of the illumination angle and spot size before interaction with the sample.

In addition to the main condenser lenses, the illumination system includes stigmators. Modern transmission electron microscopes are equipped with stigmators, and in advanced models, aberration correctors (e.g., C_s and C_c correctors) ensure optimal beam and image quality.

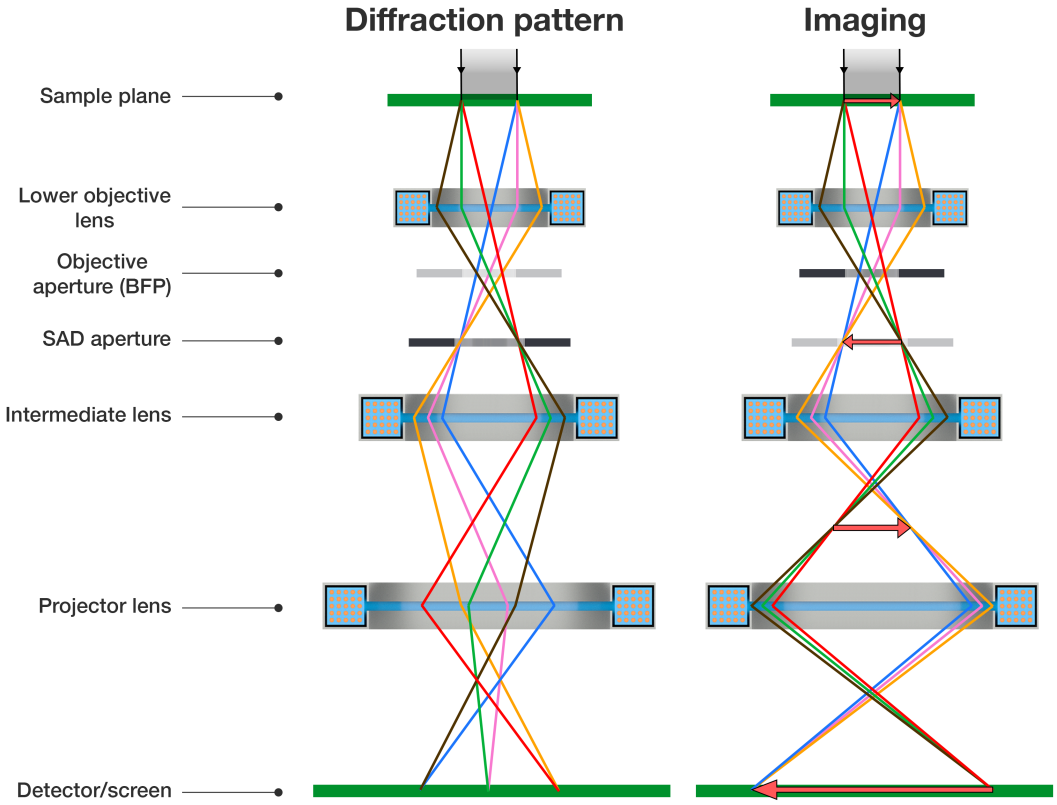


Figure 3.8. Imaging vs. Diffraction modes.

Stigmators, typically positioned near the condenser and objective lens regions, compensate for residual two-fold astigmatism and maintain circular beam symmetry (see Section 3.7). In more sophisticated instruments, multipole aberration correctors—placed below the condenser or objective lens depending on the configuration—further minimize higher-order aberrations such as spherical and chromatic, allowing sub-ångström resolution and improved optical performance. Proper stigmation is essential in both TEM and STEM modes to achieve symmetric beam profiles and optimal probe formation.

3.3.2. Objective System

The objective lens is the heart of the transmission electron microscope (TEM). It is responsible for forming both the first intermediate image of the specimen and the diffraction pattern (DP). These two outcomes occur in distinct conjugate planes of the objective lens: the *objective image plane*, where the real-space image forms, and the *back focal plane* (BFP), where the diffraction pattern is located. Figure 3.8 illustrates the ray diagrams for the two principal modes in TEM.

In practice, modern TEMs allow the operator to switch seamlessly between these two modes via software. The intermediate lens system determines which plane—image plane or BFP—is projected onto the detector:

- **Image mode**: The intermediate lens images the objective *image plane*, producing a magnified real-space view of the specimen. Contrast can be modified by inserting

an *objective aperture* in a plane conjugate to the BFP. This limits the collection of high-angle scattered electrons and can be used to isolate specific scattering angles for *bright-field* (BF) or *dark-field* (DF) imaging, as discussed in the next section.

- **Diffraction mode**: The intermediate lens images the *BFP* of the objective lens, projecting the diffraction pattern onto the detector. This pattern encodes reciprocal-space information about the specimen, including lattice geometry, symmetry, and orientation.

In conventional diffraction, the recorded pattern contains contributions from the entire illuminated area of the specimen. To restrict the diffracting volume, *selected-area diffraction* (SAD) is employed. A *selected-area diaphragm* (SAD aperture) is inserted not at the specimen plane—where the specimen already resides—but in the *objective image plane*, which is conjugate to the specimen plane. This creates a *virtual aperture* at the specimen location, limiting the region that contributes to the diffraction pattern.
Only electrons passing through this virtual aperture proceed to the BFP and form part of the recorded DP, while those outside are blocked by the SAD aperture. In practice, the smallest SAD apertures are typically $\sim 10\,\mu$m in diameter, and, with a demagnification of $\sim 1/25$ to the specimen plane, the minimum selected area is on the order of $0.4\,\mu$m. This size is suitable for many crystallographic studies but may be too large for certain nanoscale analyzes.
Another key parameter in diffraction mode is the *camera length* (L), analogous to the sample-to-detector distance in X-ray diffraction. It determines the scaling between the scattering angle and the spot position on the detector. The camera length can be adjusted by acting on the projector lenses, but it must be calibrated for accurate measurement of lattice spacings.

The ability to switch between imaging and diffraction modes—combined with apertures that control contrast, resolution, and sampling area—makes the objective lens system the central hub for both real-space imaging and reciprocal-space analysis in TEM.

3.4. Diffraction in TEM

When a thin specimen is illuminated in TEM, the transmitted wave leaving the specimen contains a superposition of a direct (central) beam, surrounded by elastically and inelastically scattered electrons. The objective lens converts the scattering angles into positions in its back focal plane (BFP), where the *diffraction pattern* forms. In the small-angle (paraxial) approximation, each diffracted beam corresponds to a reciprocal-lattice vector $\mathbf{g}$, and the spot positions scale with the camera length and the scattering angle. Therefore, the diffraction pattern provides a direct and highly sensitive fingerprint of the specimen periodicity and orientation (also see Figure 3.9). Spot separations are inversely related to real-space interplanar spacings d_{hkl}, while spot symmetries reflect the projected crystal symmetry along the zone axis [41].
A useful interpretation, valid for periodic and non-periodic objects alike, is that the diffraction

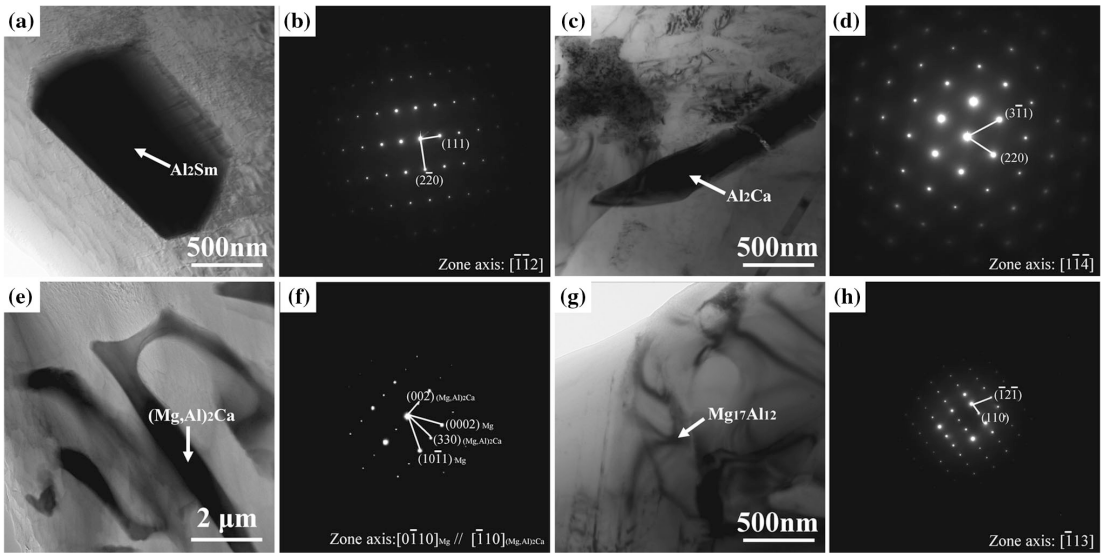


Figure 3.9. TEM images and selected-area electron diffraction (SAED) patterns of (a,b) Al_2Sm, (c,d) Al_2Ca, (e,f) lamellar $(Mg,Al)_2Ca$, and (g,h) $Mg_{17}Al_{12}$ [62].

pattern represents the squared modulus of the Fourier transform of the specimen transmission, or, more precisely, of the exit wave emerging from the specimen.

3.5. Bright-Field and Dark-Field Imaging in TEM

An *objective aperture* placed in the BFP selects which scattered beams contribute to image formation. This beam selection defines the basic amplitude-contrast modes: *bright-field* (BF) and *dark-field* (DF) imaging [63]. In BF imaging, the objective aperture is centered on the direct beam so that the image is formed mainly from electrons that remain close to the optic axis. Contrast then arises from processes that reduce the transmitted intensity, including mass–thickness variations and elastic scattering into diffracted beams (and, to a lesser extent, inelastic losses), so thicker regions, higher-Z regions, or strongly diffracting areas appear darker. In DF imaging, the objective aperture is shifted to select a chosen diffracted beam while excluding the transmitted beam. Only regions that satisfy the Bragg condition for that reflection contribute strongly, appearing bright on a dark background, which provides crystallographic selectivity and is useful for highlighting specific grains, domains, phases, or defects. In practice, BF is often used first as a rapid survey mode to locate regions of interest and assess overall morphology and thickness variations, while DF is used to isolate selected crystallographic information by choosing an appropriate reflection in the BFP. Since this thesis focuses on electron beam shaping and phase control, BF and DF are introduced here only as standard imaging configurations that are defined by beam selection in the objective BFP.

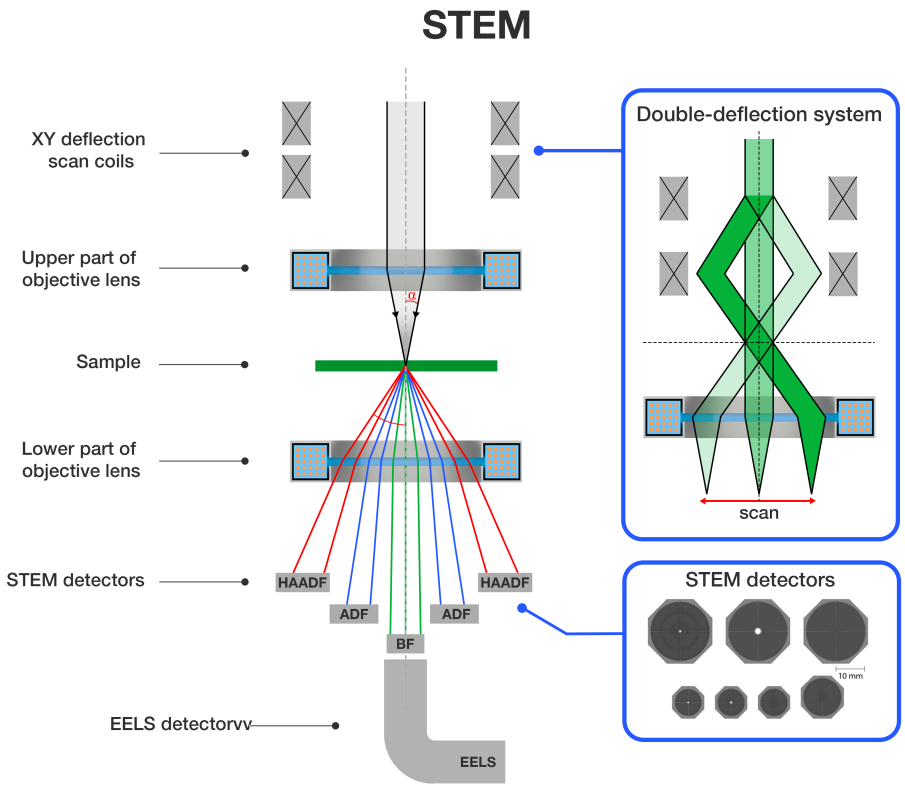


Figure 3.10. Schematic layout of STEM imaging. A focused, raster-scanned electron probe is formed by the condenser–objective system and steered by the scan (deflector) coils across the specimen. Transmitted electrons are collected by the bright-field (BF), annular dark-field (ADF), and high-angle annular dark-field (HAADF) detectors.

3.6. TEM vs. STEM

In conventional TEM, a broad (nearly parallel) beam illuminates the specimen and the objective lens forms a magnified image from the exit wave. In scanning transmission electron microscopy (STEM), the microscope instead forms a highly focused *convergent* probe and scans it across the specimen in a raster. At each probe position, the transmitted and scattered electrons are measured by post-specimen detectors, and the recorded signal is assigned to the corresponding pixel. Therefore, the STEM image is built point-by-point from a synchronized pair: probe position and detector signal [43, 64] (see Figure 3.10).

Figure 3.10 summarizes the STEM layout. The gun and condenser system demagnify the source and set the probe convergence semi-angle α using condenser apertures. Scan coils (typically above the objective lens) deflect the probe to successive positions on the specimen. The objective lens is the probe-forming element and thus largely defines the attainable spatial resolution in STEM. Below the specimen, detectors are placed in (or conjugate to) the diffraction plane to collect electrons within selected angular ranges. Because STEM resolution is set by the *probe size* rather than by post-specimen imaging lenses, probe formation is the central performance bottleneck for both imaging and spectroscopy.

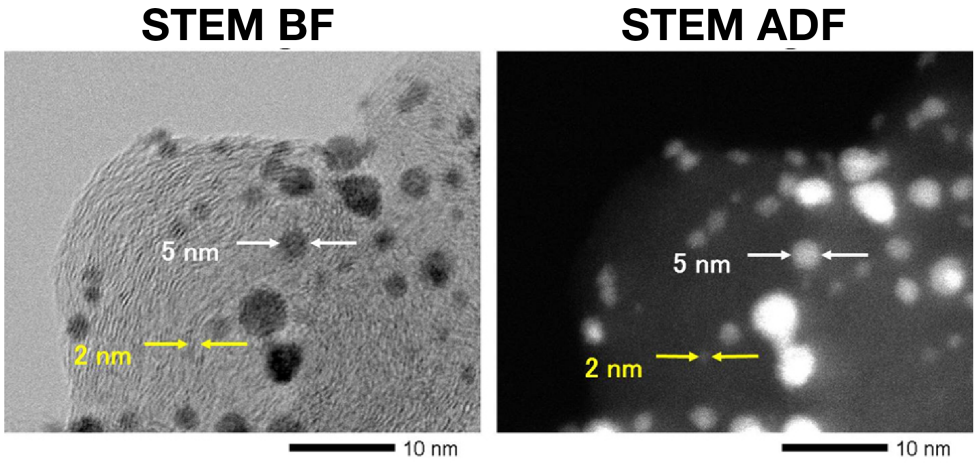


Figure 3.11. Bright-field (left) and annular dark-field (right) STEM images of Pt catalyst particles on a graphite support, showing dark contrast in BF and bright (reversed) contrast in ADF due to high-angle scattering collected by the annular detector.

Electron detection in STEM: BF, ADF, and HAADF. STEM contrast is controlled mainly by the detector geometry, i.e., by the range of scattering angles that are integrated at each pixel [43, 64]. A bright-field (BF) detector collects the central, low-angle part of the transmitted beam, making the signal sensitive to coherent effects such as phase and diffraction contrast (also see Figure 3.10). An annular dark-field (ADF) detector excludes the central beam and collects electrons scattered to higher angles. When the inner collection angle is sufficiently large (HAADF), the signal is dominated by incoherent, high-angle scattering, and the image becomes easier to interpret because the intensity varies monotonically with scattering strength. This is the basis of so-called "Z-contrast" imaging, which is widely used for atomic-resolution structural imaging and for guiding spectroscopic mapping[64].

A key advantage of STEM is that spectroscopy can be recorded in a spatially resolved way by synchronizing a spectrum acquisition with the scan. In electron energy-loss spectroscopy (EELS), a post-specimen spectrometer disperses the transmitted electrons by their energy loss. Core-loss edges provide elemental fingerprints, while the near-edge fine structure (ELNES) reports local bonding and valence information [66]. In spectrum-imaging, a full EELS spectrum is recorded at every probe position, yielding a data cube that can be processed into elemental and chemical-state maps (also see Figure 3.12). Energy-dispersive X-ray spectroscopy (EDX) can be acquired in parallel and is often complementary, especially for robust identification and quantification of heavier elements [43, 41].

In STEM, lens aberrations directly broaden the focused probe and reduce peak intensity, which degrades both spatial resolution and signal-to-noise in imaging and in spectrum imaging. The probe size is determined by diffraction (set by the convergence angle α) and by residual aberrations of the probe-forming optics, most importantly spherical aberration and lower-order asymmetries. Without correction, increasing α eventually worsens the probe because aberrated rays contribute more significantly; with aberration correctors, a larger usable α becomes possible, enabling smaller probes and higher current densities at the speci-

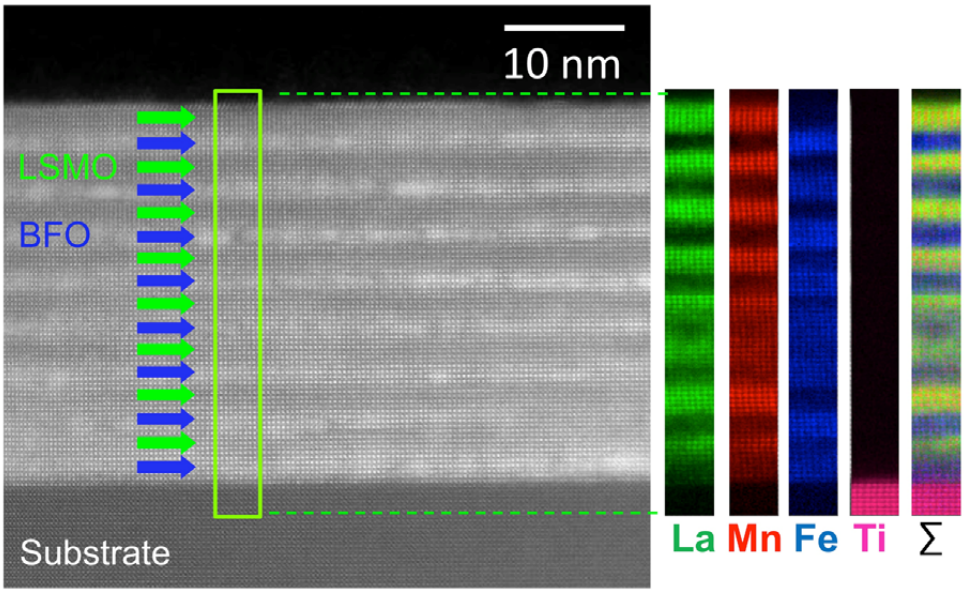


Figure 3.12. Atomic-resolution HAADF-STEM image of a BFO/LSMO superlattice grown on an STO substrate, with EELS elemental maps (right) of La $M_{4,5}$, Mn $L_{2,3}$, Fe $L_{2,3}$, and Ti $L_{2,3}$, and a composite overlay [65].

men [64]. This link between residual aberrations, convergence angle, and probe quality is the main reason why aberration correction is essential for modern high-resolution STEM and for atomic-scale EELS mapping.

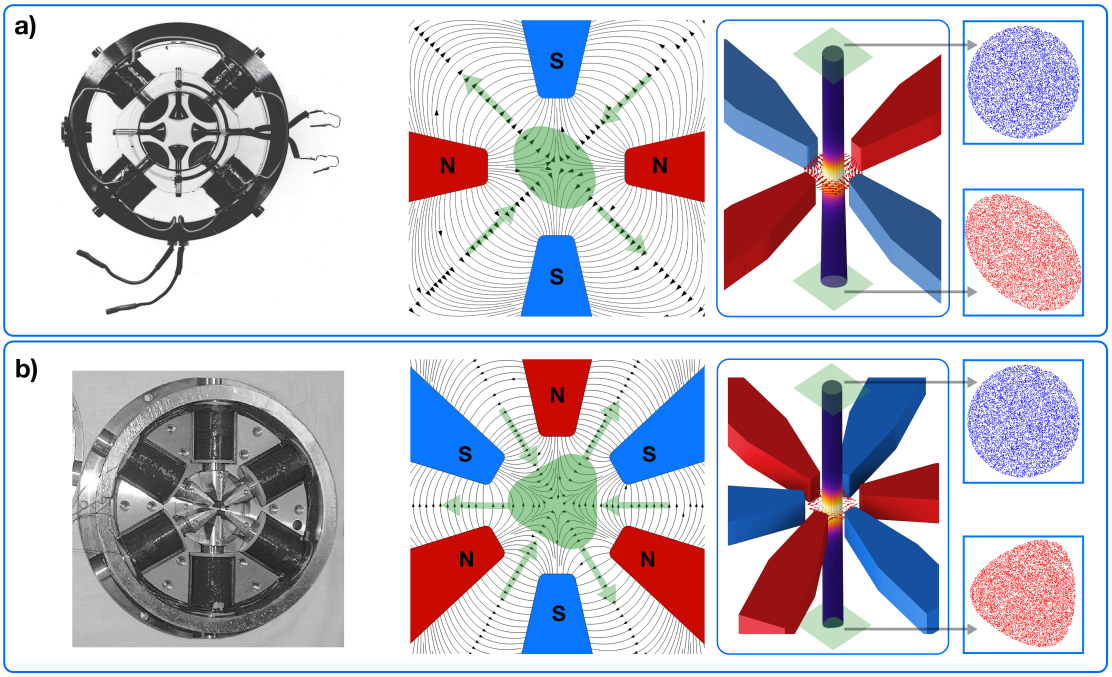


Figure 3.13. Device, magnetic field, and electron beam trajectory simulations for (a) a quadrupole and (b) a hexapole [67].

3.7. Aberration Correctors and Stigmators

Another essential type of magnetic element used in TEM for aberration correction is the multipole lens. Unlike conventional round magnetic lenses (solenoids), whose magnetic fields are mainly aligned with the optical axis, multipole lenses generate transverse magnetic fields that act perpendicular to the electron-beam direction. These lenses contain several pole pieces—typically an even number arranged symmetrically around the axis—with alternating north and south poles. When current flows through the coils, a controlled transverse magnetic field forms between the poles [57]. The symmetry of the pole arrangement determines the angular dependence of the field and, therefore, the type of aberration the element can generate or correct. Figure 3.13 shows two types of multipoles, quadrupole and hexapole, commonly used in TEM aberration correctors, together with simulations of their magnetic fields and the resulting charged-particle trajectories. These simulations illustrate how the Lorentz force modifies electron paths, leading to characteristic elliptical distortions for quadrupoles and three-lobed (triangular-like) distortions for hexapoles [67].

In modern transmission electron microscopes, the quadrupole lens is the primary magnetic multipole used to control and correct twofold astigmatism, denoted f_{a2} in Eq. 2.26 (also written as C_{12} in Krivanek notation and A_1 in Haider notation). As discussed in Section 2.5, two-fold astigmatism is a second-order aberration that stretches the probe along one axis while compressing it along the perpendicular one. A magnetic quadrupole naturally produces this type of anisotropic focusing: its four-pole field increases linearly with radial distance and exerts opposite focusing actions in the two transverse directions [68]. When a rotationally symmetric beam passes through such a field, the resulting aberration is pure

twofold astigmatism, whose orientation is set by the physical rotation of the quadrupole. This one-to-one correspondence between field symmetry and aberration symmetry is the basis for using quadrupoles as dedicated astigmatism correctors (stigmators).

A single quadrupole cannot fully correct astigmatism because it provides only one fixed orientation of the ϕ_{a2} term. For complete control, modern stigmators use two quadrupoles rotated with respect to one another. By adjusting their excitation strengths and relative angles, the system can generate an astigmatic field with arbitrary magnitude and direction, allowing the operator to cancel both components [68]. When placed in the objective lens region (or condenser region), this quadrupole pair compensates for residual imperfections such as mechanical misalignments, pole-piece asymmetries, and small magnetic-field inhomogeneities—factors that historically limited resolution before spherical aberration became the dominant concern. Scherzer had already emphasized the need for a tunable quadrupole field to obtain a true point focus [28], and by the late 1940s, quadrupole-based stigmators were established as essential components of high-resolution TEM instruments.

Although stigmators effectively remove twofold astigmatism, they do not address the intrinsic positive third-order spherical aberration C_s that arises from all rotationally symmetric magnetic lenses. Scherzer's theorem showed that for round, static magnetic lenses operating in a vacuum, C_s must remain positive [53]; therefore, spherical aberration could never be eliminated simply by improving lens design or manufacturing tolerances. Early attempts to overcome this limitation—such as the electrostatic correctors of Seeliger in the 1950s [69, 70] and the multipole concepts explored by Moellenstedt [71], Scherzer, and later the Chicago and European Molecular Biology Laboratory (EMBL) groups demonstrated the feasibility of generating negative spherical aberration fields [72]. However, practical implementations were hindered by magnetic hysteresis, mechanical instability, and the lack of precise, quantitative alignment methods [70, 29].

The theoretical foundations for modern correctors emerged in the 1970s and 1980s, when Harald Rose and colleagues established how axially separated multipole elements could generate "combination aberrations" capable of canceling the spherical aberration of a round lens. However, it was not until the late 1990s when stable power supplies, improved magnetic materials, CCD-based diagnostics, and computer-controlled alignment systems became available that practical spherical aberration correction was achieved. Two successful approaches emerged: the double-hexapole corrector demonstrated by Haider and coworkers for TEM [67, 72], and the quadrupole-octopole architecture realized by Krivanek and collaborators in STEM [73]. These designs form the basis of the correctors used in today's aberration-corrected microscopes.

3.7.1. Double hexapole C_s aberration correctors

A hexapole lens consists of six pole pieces with alternating polarity arranged around the optical axis. Close to the axis, its scalar magnetic potential in cylindrical coordinates (r, ϕ, z)

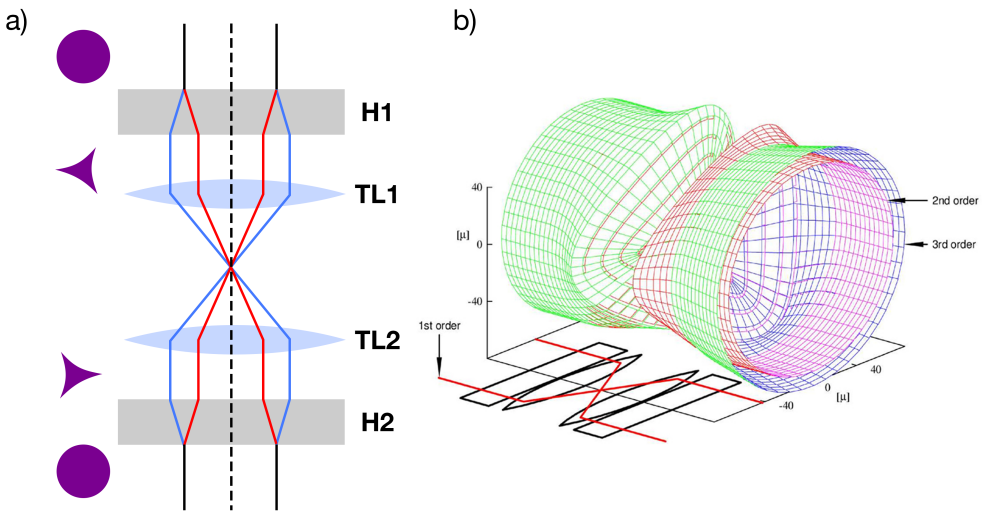


Figure 3.14. Hexapole Cs correctors [50].

is dominated by a threefold term

$$\Phi_m(r,\phi) \propto r^3 \cos(3(\phi - \phi_0)), \tag{3.5}$$

which produces a transverse field that grows in amplitude as r^2. In ray optics, a thin hexapole gives a transverse kick proportional to the square modulus of the complex conjugate position [74],

$$\Delta w' \propto \kappa\, \bar{w}^2, \tag{3.6}$$

where $\bar{w} = x - iy$ is the complex coordinate of the electron trajectory. The parameter κ represents the hexapole excitation strength. This angular dependence is the signature of threefold astigmatism ($A_2/C_{2,3}$): a single hexapole produces a strong triangular distortion but no rotationally symmetric C_s term.

The key insight is that, in a thick or multi-element configuration, the first-order deflection caused by the hexapole modifies the electron trajectory as it propagates through the field. The resulting second-order interaction between the trajectory and the r^2 field produces a third-order aberration. For a suitable double-hexapole configuration, this third-order contribution is rotationally symmetric and can be written as an effective spherical aberration [75, 76]

$$C_{s,\mathrm{corr}} \approx -\xi L \kappa^2, \tag{3.7}$$

where ξ is a geometric factor determined by the transfer lens optics, and L is the effective interaction length of the hexapole field. The minus sign indicates that the corrector generates negative C_s that can be tuned to cancel the positive C_s of the objective lens.

Operating principle

A thick hexapole can therefore generate negative C_s, but it also introduces a large threefold astigmatism, which would make the image unusable. The classic Rose–Haider corrector overcomes this by using two hexapoles (H1 and H2) separated by a transfer lens (TL) system. The transfer lens images the principal plane of H1 into that of H2 with a magnification of $M = -1$. This corresponds to a rotation of the azimuthal coordinate by π:

$$\phi \to \phi + \pi.$$

Since the hexapole potential scales as $\cos(3\phi)$,

$$\cos[3(\phi + \pi)] = \cos(3\phi + 3\pi) = -\cos(3\phi),$$

the threefold astigmatism from H2 has the same magnitude but opposite sign of H1. In other words, the double-hexapole plus the $M = -1$ transfer lens satisfies two simultaneous conditions:

- **Generation of negative C_s:** the spherical aberration terms of H1 and H2 add.
- **Cancellation of threefold astigmatism:** the A_2 contributions of H1 and H2 cancel.

In practice, the transfer lens acts as a symmetric doublet (a telescopic "$4f$" system) that keeps the global Gaussian optics of the column unchanged while providing the required $M = -1$ between the two hexapoles.
Modern correctors use dodecapole (12-pole) hardware rather than pure hexapoles. By applying suitable current combinations, the same element can generate hexapole, quadrupole, and dipole fields, which allow the corrector to provide both the main C_s correction and the fine alignments (beam tilt, stigmation, centering) needed for stable operation [77].

Practical criteria and residual aberrations

The correction process subtracts two large numbers: the positive C_s of the objective lens and the negative C_s of the corrector. The residual is small and, therefore, very sensitive to fluctuations. Three main criteria must be satisfied in any practical hexapole corrector:

1. **Power-supply stability.** The currents driving the objective lens, adaptor lenses, and multipoles must be stable. Small current drifts translate into changes of focal length and effective C_s, which blur high spatial frequencies and lower the information limit.
2. **Environmental stability.** Mechanical vibrations, acoustic noise, temperature drifts, and Unintended AC magnetic fields can all move the optical axis or change lens strengths. Corrected microscopes, therefore, require vibration-isolated foundations, acoustic enclosures, temperature-controlled rooms, and active magnetic field compensation.

3. **Automated alignment.** A hexapole corrector contains many coupled multipole elements. Alignment is performed via automated procedures (often based on Zemlin tableaux [78]) that tilt the beam, record a series of images or Ronchigrams, and iteratively adjust multipole strengths to minimize coma and residual astigmatism.

Once the low-order aberrations (defocus, A_1, B_2, A_2, C_3) are compensated, the performance becomes limited by higher-order terms. For hexapole correctors, the most important are the fifth-order sixfold astigmatism A_5 and the fifth-order spherical aberration C_5. The intrinsic A_5 produces a six-pointed star pattern in the Ronchigram at large angles and sets the maximum usable convergence semi-angle $\alpha_{\max}$. This, in turn, constrains the achievable point resolution and the depth of field in aberration-corrected TEM and STEM.

Commercial hexapole corrector families

The Rose–Haider concept has been implemented and commercialized mainly by CEOS GmbH as modular correctors that are integrated into microscopes from major manufacturers. The main hexapole-type product lines can be summarized as follows [79]:

- **CETCOR (TEM C_s corrector).** An Image-side double-hexapole corrector is placed after the objective lens in TEM. It compensates for the objective C_s and low-order aberrations up to the third order, allowing for near-delocalized phase-contrast imaging and direct interpretation of atomic-resolution lattice images.
- **CESCOR (STEM C_s corrector).** A Probe-forming corrector mounted in the condenser/objective system of STEM. It reduces C_s and associated aberrations of the probe-forming lens, enabling sub-ångström electron probes with high current, which is essential for analytical STEM techniques such as EELS and EDS.
- **DCOR and ASCOR (advanced hexapole correctors).** Next-generation STEM correctors that extend the hexapole concept with additional multipole degrees of freedom. They are designed to minimize the intrinsic sixfold astigmatism A_5 and to control fifth-order spherical aberration C_5, thereby permitting large convergence semi-angles (typically > 40 mrad) and improved depth sectioning and 3D imaging.

Together, these systems have transformed high-resolution TEM and STEM from C_s-limited instruments into tools where the dominant limits are now higher-order aberrations, chromatic effects, and sample stability, rather than the fundamental properties of the objective lens itself.

3.7.2. Quadrupole–Octopole C_s Correctors

A second major class of spherical aberration correctors is based on combinations of quadrupoles and octopoles. In contrast to the hexapole doublet—which relies on symmetry to cancel second-order terms and generate a net negative third-order spherical aberration quadrupole

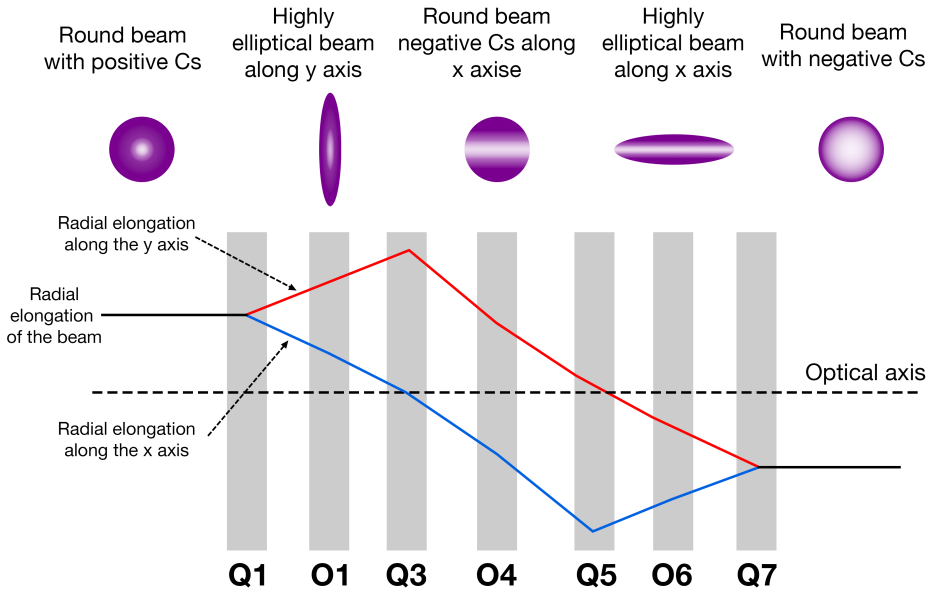


Figure 3.15. Quadrupole-Octopole C_s corrector. The corrector consists of three octupoles (O2, O4 and O6) and four quadrupoles (Q1, Q3, Q5 and Q7). The radial elongation of the beam through the corrector is indicated along two orthogonal axes x and y (Adapted from Ref. [80] and redesigned by the author).

octopole correctors exploit the interaction of multipole fields at *different axial planes* to create so-called *combination aberrations*.

An octopole (eight-pole) magnetic lens generates a scalar potential of the form

$$\Phi_m(r, \phi) \propto r^4 \cos(4\phi). \tag{3.8}$$

with a corresponding transverse field scaling as $B \propto r^3$. Acting alone on a round beam, an octopole therefore produces a *fourfold astigmatism* term, denoted A_3 in Haider's notation. This is a third-order angular aberration with multiplicity $m = 4$. Like quadrupoles and hexapoles, an isolated octopole does *not* generate spherical aberration; its field symmetry does not match the rotationally symmetric $C_{3,0}$ term.

When successive multipole elements are placed at different planes along the optical axis, their aberration contributions combine in a nontrivial way. The operation of a quadrupole–octopole corrector can be described in four main stages [80]:

Stage 1: X-plane correction. A set of quadrupoles stretches the beam along x and compresses it along y. With $y \approx 0$, the octopole potential reduces to

$$\Phi_{m4}(x, y \approx 0) \propto x^4,$$

This generates a negative spherical aberration term in the x-plane.

Stage 2: Beam rotation. A central quadrupole group exchanges the roles of the two axes by focusing in x and defocusing in y, effectively rotating the line focus by 90°.

Stage 3: Y-plane correction. The beam then enters the second octopole as a line extended

along y, giving

$$\Phi_{m4}(x \approx 0, y) \propto y^4,$$

which provides the matching negative spherical aberration in the y-plane.

Stage 4: Symmetry restoration. Exit quadrupoles restore the beam to a circular cross-section. The combined corrections in the two axes form a net rotationally symmetric negative spherical aberration that compensates the positive spherical aberration of the objective lens. The overall correction strength is the sum of the contributions from each octopole, weighted by the fourth power of the magnification at their respective planes:

$$C_{s,\text{corr}} \propto \sum_i k_i M_{x,i}^4 \; + \; \sum_j k_j M_{y,j}^4.$$

By tuning the octopole excitations, the total negative spherical aberration can be adjusted to match $-C_{s,\text{obj}}$.

3.8. Aberration Detection and Measurement in TEM

Aberration detection and correction are essential requirements for achieving sub-Ångström resolution in modern transmission electron microscopy (TEM). Even though modern instruments include advanced multipole-based aberration correctors, their performance still depends on how precisely the residual aberrations can be measured. Accurate aberration detection is therefore not only a diagnostic step but also a central part of optical alignment, directly influencing the probe size, information limit, coherence, and image stability. Over the past decades, several complementary strategies have been developed to extract aberration information from diffraction patterns and Ronchigram. This section summarizes the main approaches and highlights their physical principles and practical relevance.

3.8.1. Diffraction-Based Aberration Diagnosis (Zemlin tableau)

One of the earliest and still influential techniques is the Zemlin tableau [78]. In this method, a series of diffractograms is recorded at a constant beam-tilt magnitude while the tilt azimuth is varied. When arranged into a tableau, azimuth-dependent changes reveal residual asymmetric aberrations such as axial coma and astigmatism. Zemlin and co-workers showed that even well-aligned microscopes display strong asymmetries when examined with this method [78]. Although time-consuming, the Zemlin tableau remains a benchmark for validating alignment and is still used on modern aberration-corrected microscopes [82, 81]. Its characteristic tilt-dependent patterns, illustrated in Figure 3.16, clearly reveal the effect of uncorrected aberrations and the improvement achieved after correction.

3.8.2. Real-Space Aberration Detection Using Ronchigrams

A Ronchigram is a convergent-beam electron diffraction (CBED) pattern (commonly recorded in STEM) that contains both real-space-like (direct/shadow) and reciprocal-space (diffrac-

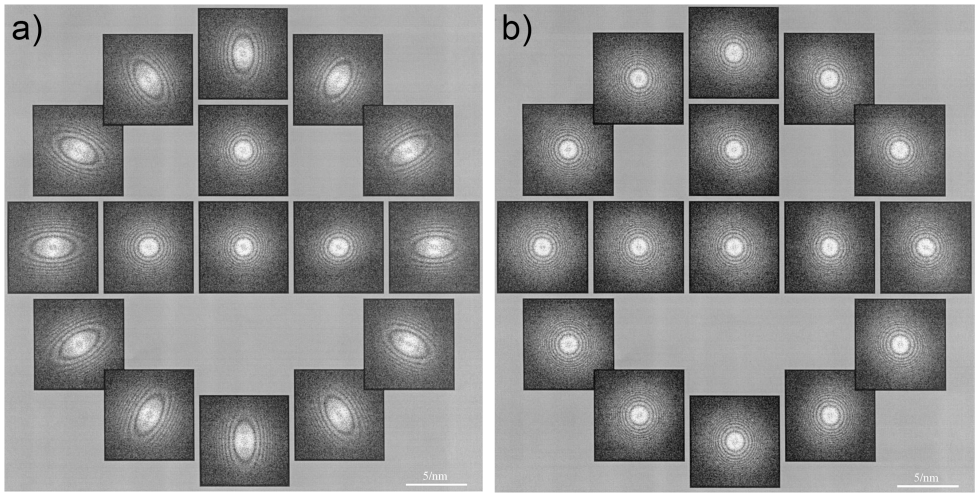


Figure 3.16. Zemlin tableau recorded at a beam tilt of 10.8 mrad. (a) Uncorrected but aligned microscope showing tilt-dependent variations dominated by residual spherical aberration. (b) Corrected microscope showing minimal changes in defocus and astigmatism [81].

tion) information in a single image. It is obtained by focusing a strongly convergent probe at or near a thin amorphous specimen and recording the transmitted intensity across the diffraction plane. Real-space diagnosis using Ronchigrams is one of the most powerful alignment tools in STEM. Ronchigram contrast reflects the second derivatives of the aberration phase and is therefore highly sensitive to both low-order and high-order axial aberrations [83]. A key qualitative indicator in Ronchigram analysis is the presence of regions with nearly constant phase. These areas grow larger and become more symmetric as spherical aberration and other parasitic terms are reduced. The shape and boundaries of these constant-phase regions provide immediate visual feedback on the quality of the probe correction. Figure 3.17 shows simulated Ronchigrams illustrating how different aberrations modify the central constant-phase region and introduce characteristic symmetry changes [84].

Several quantitative Ronchigram-based methods have also been developed. The Segmental Ronchigram Autocorrelation Method (SRAM) introduced by Sawada *et al.* [85] divides the Ronchigram into angular sectors and computes their autocorrelation. The resulting matrix encodes aberrations up to the fifth order and can be obtained from a single Ronchigram, making the method fast and suitable for automated corrector alignment.

A complementary approach was proposed by Lupini *et al.* [86], who analysed Ronchigrams under the assumption that small patches are locally isoplanatic. By computing the local Fourier transform and fitting a contrast transfer function (CTF), local defocus, astigmatism, and higher-order derivatives can be retrieved. This method can also separate pre-sample and post-sample aberrations, providing deeper insight into the illumination system.

Another practical aspect of Ronchigram analysis is identifying the correct aberration center. Mis-centering introduces artificial aberration terms and reduces the accuracy of any measurement. Sannomiya *et al.* proposed an automated procedure to determine the optical axis directly from a single Ronchigram [87]. This step is crucial for stable, automated correction routines.

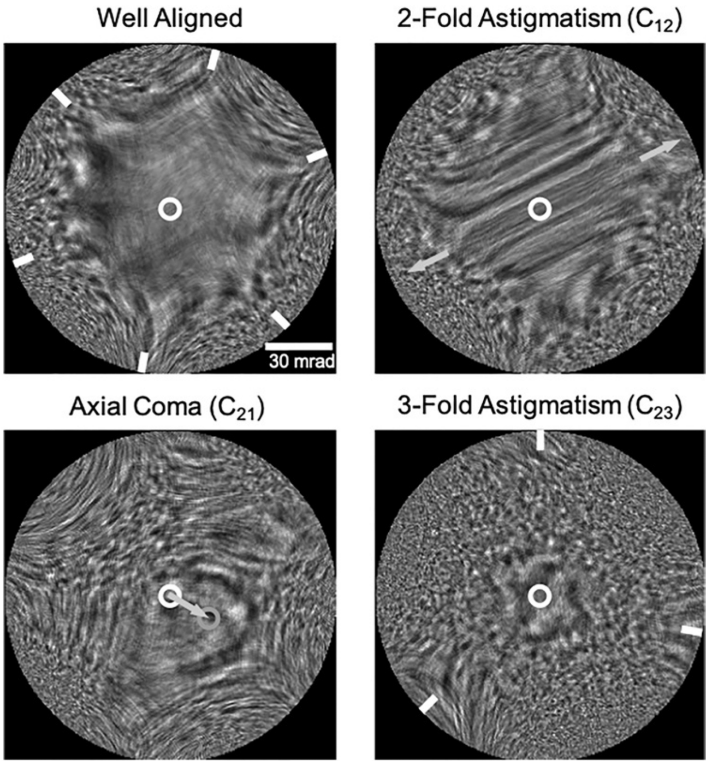


Figure 3.17. Simulated Ronchigrams illustrating the effect of common aberrations in STEM. The reference Ronchigram shows a large central constant-phase region, while added twofold astigmatism, coma, and threefold astigmatism introduce characteristic distortions and symmetry changes [84].

More recently, AI-based aberration detection has emerged as a fast and robust solution. Bertoni *et al.* from our group demonstrated that a neural network can infer several aberration coefficients directly from one Ronchigram in near-real time [82]. This approach is especially useful for systems using MEMS-based phase plates or tunable electrostatic lenses.

4 Electron Beam Shaping

Contents

4.1 Beam shaping using a phase plate 63

4.2 Types of Electron Phase Plates and their principles 65

4.2.1 Thin film phase plates 65

Defocus phase contrast and Scherzer defocus (no phase plate) . . . 68

Zernike and Hole-free phase plate (Symmetric Phase Contrast) . . 69

Hilbert and Schlieren Phase Plates (Directional Contrast) 72

4.2.2 Sculpted thin-film phase plates 73

4.2.3 Electrostatic phase plates 73

Vortex Beams and Orbital Angular Momentum 75

Vortex beam generation with holographic fork holograms 77

Generation of electron vortex beams using electrostatic phase plates ("chopsticks") 77

OAM Spectroscopy of electron vortex beams using electrostatic phase plates ("OAM sorter") 78

Vortex beams from programmable MEMS phase plates 80

4.2.4 Magnetic Phase Plates 80

4.2.5 Photonics Phase Plates 82

Continuous-wave laser phase plates 83

Pulsed-laser phase plates 84

4.1. Beam shaping using a phase plate

In TEM, beam formation in both the illumination and projection parts of the microscope is performed by magnetic lenses—supplemented, when required, by electrostatic lenses—and refined by apertures according to the selected imaging mode. For small angles (paraxial electrons), a magnetic lens behaves like a positive, convex lens: it forms crossovers and images. Because a static magnetic field performs no work, electrons passing through it retain their kinetic energy—and thus their de Broglie wavelength—while their trajectories are only redirected. In wave-optics terms, an ideal thin lens imparts a radially varying phase that gives the wavefront spherical curvature converging to a focus at distance f. In this sense, magnetic lenses act as elementary beam shapers, converting a plane wave into a spherical wave.

However, as in light optics, more refined control of the beam shape is often required. In optics, diffractive elements and computer-generated holograms can impose an arbitrary phase distribution to generate Bessel, Airy, or vortex beams [88, 89]. Spatial light modulators (SLM) and deformable mirrors provide programmable phase control for aberration correction and custom beam modes [90, 91]. More recently, metasurfaces have enabled compact and highly efficient beam shaping [92]. The same concept applies to electrons: by imprinting a designed phase profile, one can sculpt the electron beam into the desired form. In the following sections, we translate this optical toolbox into the context of TEM phase plates.

Phase plates are electron-optical elements strategically positioned along the electron beam path to modulate the phase and, in some cases, the amplitude of the electron wavefront [93]. Historically, their main purpose was to enhance image contrast in transmission electron microscopy (TEM), particularly for weak-phase objects (WPO)—specimens that produce negligible amplitude variations and therefore exhibit low contrast under conventional imaging conditions. A typical example is cryogenic electron microscopy (cryo-EM) of biological samples, where unstained proteins or nucleic acids yield poor contrast because of their weak scattering potential [94]. By introducing a controlled phase shift between the unscattered (central) and scattered electrons in the diffraction plane, phase plates convert phase variations into measurable intensity differences, greatly improving the quality and interpretability of TEM images [95]. Over the past two decades, designs such as Zernike, hole-free phase plates (also referred to as the Volta phase plate), and electrostatic phase plates have driven major advances in high-contrast imaging across both structural biology and materials science [93].

The principle of phase-contrast imaging was first introduced by Frits Zernike in 1942, who showed that applying a controlled phase shift between scattered and unscattered waves at the Back Focal Plane (BFP) of the microscope could render transparent specimens visible—an innovation that revolutionized optical microscopy [96]. Only a few years later, in 1947, Hans Boersch extended this concept to electrons by proposing the first electron-optical method for phase contrast in TEM, establishing the theoretical foundation for phase plate development [97].

Beyond contrast enhancement, phase plates have now evolved into powerful tools for beam shaping and probing quantum phenomena in the electron microscope. Holographic and electrostatic vortex phase plates enable the generation of electron beams carrying orbital angular momentum (OAM), opening new possibilities for studying angular momentum transfer and magnetic interactions at the nanoscale [98, 99]. Laser-based phase plates, exploiting ponderomotive potentials, further extend the coherent control of electron waves through light–matter interaction, suggesting routes toward quantum-coherent electron microscopy [100]. Thus, modern phase plates serve not only as practical instruments for high-contrast imaging but also as experimental platforms for exploring the quantum nature of the electron.

Despite the conceptual clarity of Zernike's and Boersch's proposals, implementing phase plates in TEM posed severe technical challenges for decades. The main difficulties arose from the need for nanometer-scale precision to fabricate structures that impart a controlled phase shift without introducing unwanted artifacts. Producing mechanically stable and precisely aligned phase-shifting elements, often only a few micrometers in size,proved technologically demanding. Moreover, the high-vacuum environment and intense electron beam exacerbated problems such as contamination and electrostatic charging, both of which degraded device stability and performance [101, 102, 103]. Consequently, although the theoretical advantages of phase contrast were recognized early, practical demonstrations in TEM were delayed for more than sixty years after Boersch's proposal.

It was only in the early 2000s that thin-film phase plates—typically amorphous carbon membranes with micrometer-scale central holes—were successfully developed, marking a major breakthrough. These devices finally enabled in-focus phase-contrast imaging in TEM, significantly enhancing the signal-to-noise ratio for WPO specimens. Unlike defocus-based contrast methods, phase plates achieve high contrast without reducing resolution. Their ability to reveal fine structural details at low electron doses has made them essential for high-resolution cryo-EM and other dose-sensitive applications [104].

In this chapter, we review the main types of phase plates developed for TEM, focusing on their physical principles, fabrication methods, advantages, and limitations. For clarity, the existing phase-plate technologies are grouped as follows [105, 93]:

- Thin-Film Phase Plates
- Sculpted Thin-Film Phase Plates (STPPs)
- Electrostatic Phase Plates

- Magnetic Phase Plates
- Laser-Based Phase Plates

4.2. Types of Electron Phase Plates and their principles

4.2.1. Thin film phase plates

In general, the back focal plane (BFP) of a transmission electron microscope contains the spatial frequency information of the specimen, representing how image intensity varies across different spatial scales. Each region of the BFP corresponds to a specific range of spatial frequencies that encodes different structural features of the object.

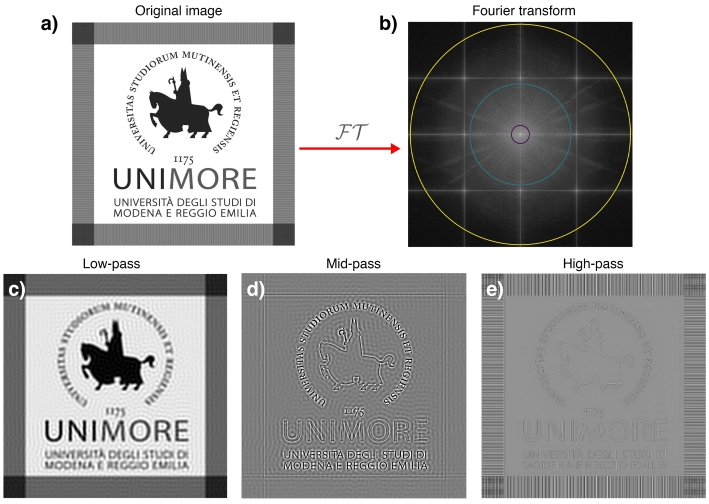


Figure 4.1. Spatial-frequency decomposition of an image. (a) Original image. (b) Fourier-magnitude spectrum (log scale) with three annular masks marking low, mid, and high frequency bands. (c) Low-pass frequencies: slowly varying components that set overall contrast and large-scale structure. (d) Mid-band frequencies: edges and mesoscale textures. (e) High-pass frequencies: finest details and high-frequency texture (also most sensitive to noise).

- **Low spatial frequencies** — Slowly varying components that encode the overall shape, large-scale structure, and broad intensity variations. They set the overall contrast. In in-focus weak-phase imaging without a phase plate, the CTF vanishes at $q = 0$, so these components are poorly transferred, and the image looks flat.
- **Mid spatial frequencies** — Mesoscale features such as edges, boundaries, and textures at intermediate length scales; these dominate the perceived "structure" of many specimens.
- **High spatial frequencies** — Rapid variations arising from sharp edges and fine details; they determine image sharpness and the visibility of the smallest features.

Conventional phase-contrast imaging (Fresnel and Foucault modes) in TEM is often inefficient for WPO because phase modulations are only weakly converted into intensity, leading to

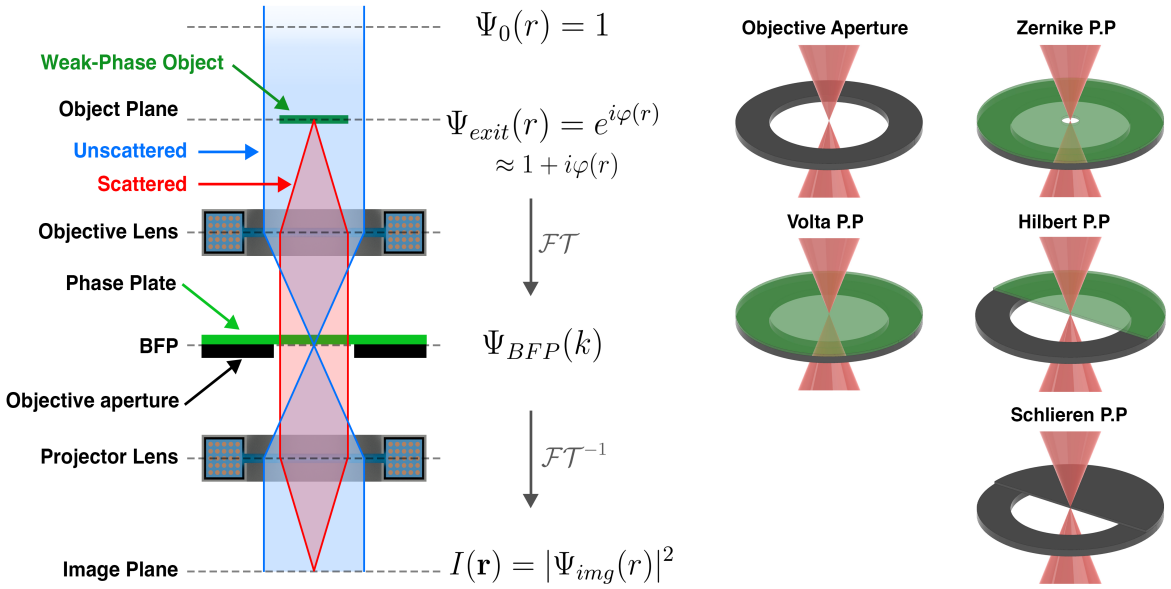


Figure 4.2. Schematic of image formation pathway of Weak Phase Object (WPO)

poor low-spatial-frequency contrast. Thin-film phase plates are placed in the back focal plane (BFP) of the objective lens to overcome the poor transfer of low spatial frequencies by selectively shifting the phase of scattered electrons relative to the unscattered central beam. This enhances the transfer of low- and mid spatial frequencies, improving the visibility of the specimen without altering its appearance.

Figure 4.2 schematically illustrates the image formation pathway: an incoming plane wave interacts with the object, producing a modulated exit wave. The objective lens performs a Fourier transform, producing a diffraction pattern in the BFP, where a phase plate can selectively modify specific components. Finally, the inverse Fourier transform performed by the imaging system reconstructs the real-space image at the detector. In the following, we derive this process mathematically; from input wave to final image.

A unit-amplitude plane wave $\Psi_0(\mathbf{r}) = 1$ illuminates a thin specimen, and finite source size and illumination-aperture effects are neglected. The specimen is described by the transmission function

$$t(\mathbf{r}) = \exp\big[i\,\varphi(\mathbf{r})\big],$$

where $\varphi(\mathbf{r})$ denotes the object-induced phase shift. For a WPO with $|\varphi| \ll 2\pi$, a first-order expansion gives

$$t(\mathbf{r}) \approx 1 + i\,\varphi(\mathbf{r}).$$

The exit wave immediately after the specimen is

$$\Psi_{\text{exit}}(\mathbf{r}) = \Psi_0(\mathbf{r})\,t(\mathbf{r}) \approx 1 + i\,\varphi(\mathbf{r}). \tag{4.1}$$

At the back focal plane (BFP) of the objective lens, the wave function equals the Fourier

transform of the exit wave:

$$\Psi_{\mathrm{BFP}}(\mathbf{k}) = \mathcal{F}\{\Psi_{\mathrm{exit}}\}(\mathbf{k}) = \delta(\mathbf{k}) + i\,\Phi(\mathbf{k}), \tag{4.2}$$

with

$$\Phi(\mathbf{k}) = \mathcal{F}\{\varphi\}(\mathbf{k}) = \iint \varphi(\mathbf{r})\, e^{-i2\pi\mathbf{k}\cdot\mathbf{r}}\, d^2\mathbf{r}, \tag{4.3}$$

where $\mathbf{r} = (x, y)$, $\mathbf{k} = (k_x, k_y)$, and $\delta(\mathbf{k})$ represents the unscattered central beam.

Image formation by an ideal thin objective lens is described by the circularly symmetric pupil function

$$P(k) = A(k)\, e^{i\chi(k)}, \tag{4.4}$$

where $A(k) = 1$ inside the aperture and $A(k) = 0$ outside of it, and the aberration function is

$$\chi(k) = -\pi\lambda\,\Delta f\, k^2 + \frac{\pi}{2}\, C_s\, \lambda^3\, k^4, \qquad k = \|\mathbf{k}\|. \tag{4.5}$$

Here Δf denotes defocus (negative for underfocus), C_s denotes the spherical aberration coefficient, and λ is the relativistic electron wavelength.

Thus, in the BFP,

$$\Psi_{\mathrm{BFP}}(\mathbf{k}) = [\delta(\mathbf{k}) + i\,\Phi(\mathbf{k})]\, A(k)\, e^{i\chi(k)}. \tag{4.6}$$

Since $A(0) = 1$ and $\chi(0) = 0$, the central beam is unchanged, and

$$\Psi_{\mathrm{BFP}}(\mathbf{k}) = \delta(\mathbf{k}) + i\,\Phi(\mathbf{k})\, A(k)\, e^{i\chi(k)}. \tag{4.7}$$

Propagation to the image plane through an ideal projector lens at unit magnification gives

$$\Psi_{\mathrm{img}}(\mathbf{r}) = \mathcal{F}^{-1}\{\Psi_{\mathrm{BFP}}\}(\mathbf{r}) = 1 + i\,\mathcal{F}^{-1}\big\{\Phi(\mathbf{k})\, A(k)\, e^{i\chi(k)}\big\}. \tag{4.8}$$

Defining

$$F(\mathbf{r}) \equiv \mathcal{F}^{-1}\big\{\Phi(\mathbf{k})\, A(k)\, e^{i\chi(k)}\big\}, \qquad \Rightarrow \qquad \Psi_{\mathrm{img}}(\mathbf{r}) = 1 + i\,F(\mathbf{r}). \tag{4.9}$$

The detected intensity for a WPO is

$$I(\mathbf{r}) = \big|\Psi_{\mathrm{img}}(\mathbf{r})\big|^2 = |1 + iF|^2 \approx 1 + 2\,\mathrm{Re}\{iF(\mathbf{r})\} = 1 - 2\,\mathrm{Im}\{F(\mathbf{r})\}. \tag{4.10}$$

Using $\mathrm{Re}\{e^{i\chi}\} = \sin\chi$ and the linearity of $\mathcal{F}^{-1}$ yields the real-space relation

$$I(\mathbf{r}) \approx 1 - 2\,\mathcal{F}^{-1}\Big\{\Phi(\mathbf{k})\, A(k)\, \sin\chi(k)\Big\}. \tag{4.11}$$

Equivalently, in Fourier space, the image contrast is

$$\Delta I(\mathbf{k}) \equiv \mathcal{F}\{I - 1\}(\mathbf{k}) \approx \Phi(\mathbf{k})\, \underbrace{\big[-2\, A(k)\, \sin\chi(k)\big]}_{\mathrm{CTF}(\mathbf{k})}. \tag{4.12}$$

Accordingly, the phase-contrast transfer function for a weak-phase object (without a phase plate) is

$$\boxed{\mathrm{CTF}(\mathbf{k}) = -\,2\,A(k)\,\sin\chi(k)}\,. \tag{4.13}$$

At exact focus—neglecting spherical aberration—we have $\chi = 0$; hence $\sin\chi = 0$, and a purely phase object produces vanishing first-order contrast.

In practice, phase contrast can be recovered either by applying defocus—which uses the $\sin\chi(k)$ term to convert phase into intensity—or by inserting a phase plate that introduces controlled phase shifts to selected spatial frequencies, thereby enhancing low-k transfer and improving image contrast.

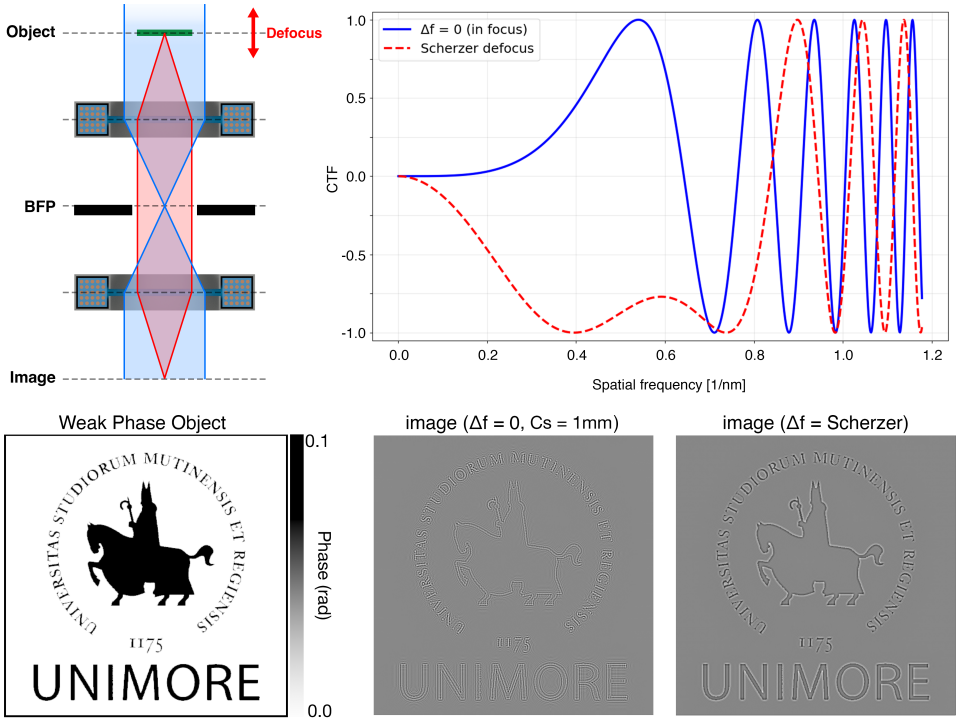


Figure 4.3. Defocus phase contrast in TEM: in-focus ($\Delta f = 0$) versus Scherzer defocus ($\Delta f = \Delta f_S$), with pathway schematic, CTF, and example phase object/images illustrating contrast changes.

Defocus phase contrast and Scherzer defocus (no phase plate)

In the absence of a phase plate, phase objects produce image contrast only through deliberate defocusing of the objective lens. Applying a controlled defocus introduces a nonzero $\sin\chi(k)$, enabling the objective lens to convert specimen phase into measurable intensity. The resulting CTF oscillates as a function of spatial frequency because the aberration phase (Equation 4.5) contains both quadratic (defocus) and quartic (spherical aberration) contributions (blue line in Figure 4.3). These oscillations generate alternating passbands and stopbands, causing contrast reversals that complicate image interpretation and ultimately limit resolution [43, 41, 101].

For a round, spherical-aberration–limited objective, the classical *Scherzer defocus* [53]

$$\Delta f_S \approx -1.2\sqrt{C_s\,\lambda}$$

maximizes the positive transfer of low and intermediate spatial frequencies up to the first CTF zero (the Scherzer resolution), providing a widely used operating point for conventional defocus phase contrast [53]. Figure 4.3 shows the CTFs for in-focus imaging ($\Delta f = 0$) and for Scherzer defocus, together with their corresponding effects on image contrast.

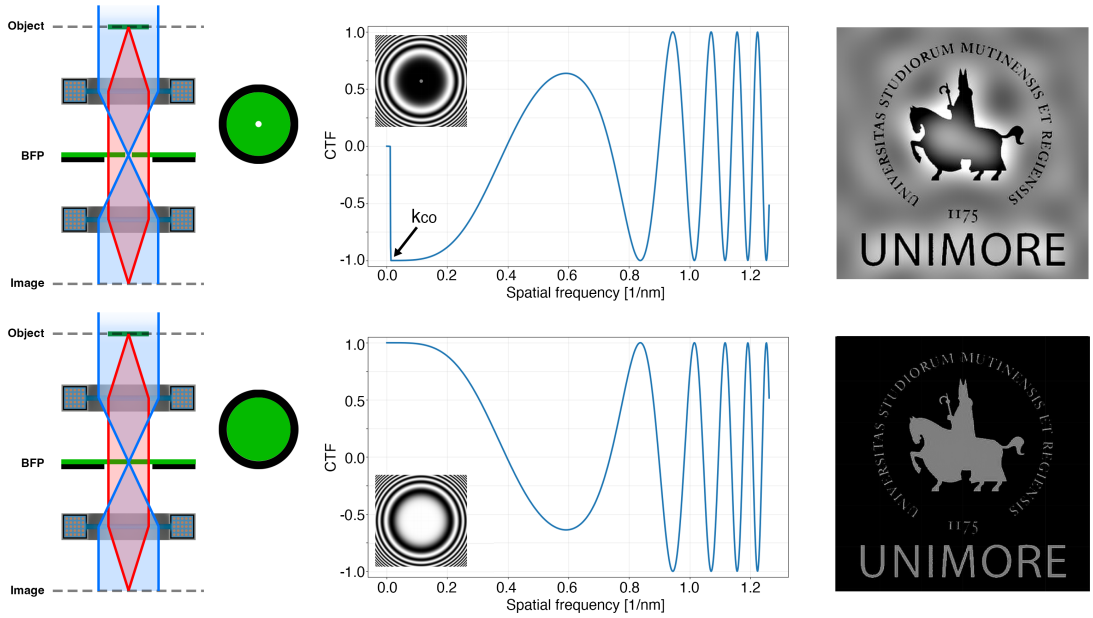


Figure 4.4. Zernike and HFPP phase plates.

Zernike and Hole-free phase plate (Symmetric Phase Contrast)

The Zernike phase plate (ZPP) consists of a uniform thin-film membrane with a small central hole and is positioned in the BFP of the objective lens. It introduces a phase shift of approximately $\pi/2$ to the scattered electron waves while leaving the intense unscattered central beam unaffected. This enhances low- and mid-frequency contrast for weak-phase objects [106].

The phase shift introduced by a thin film of thickness t and mean inner potential V_0 is given by (also see Section 2.1.3):

$$\varphi = C_E \int V(z)\,dz = C_E V_0 t \tag{4.14}$$

Table 4.1 lists the materials, their mean inner potentials, and the corresponding thickness required to produce a $\pi/2$ phase shift [107].

Table 4.1. Mean inner potential V_0 and corresponding $\pi/2$ phase-shift thickness $t_{\pi/2}$ for selected elements at 100 kV and 300 kV accelerating voltages [107].

Element	V_0 **(V)**	$t_{\pi/2}$ **(nm) @ 100 kV**	$t_{\pi/2}$ **(nm) @ 300 kV**
Be	7.8	21.8	30.9
C	7.8	21.8	30.9
Al	12.1	14.0	19.9
Si	11.5	14.8	20.9
Cu	20.1	8.5	12.0
Ge	15.6	10.9	15.4
Ag	20.7	8.2	11.6
W	23.4	7.3	10.3
Au	21.1	8.1	11.4

The central hole of radius r_h leaves low-frequency components unmodified. Thus, the cut-on frequency is [107]

$$k_{\text{edge}} = \frac{r_h}{f\,\lambda}, \tag{4.15}$$

where f is the focal length of the objective lens. If the phase plate is located in a conjugate diffraction plane, the camera length takes on the role of f.
The phase shift applied by the phase plate is

$$\varphi_{\text{PP}}(k) = \begin{cases} 0, & k < k_{\text{edge}}, \\ -\dfrac{\pi}{2}, & k \geq k_{\text{edge}}. \end{cases} \tag{4.16}$$

The wave in the BFP is

$$\Psi_{\text{BFP}}(\mathbf{k}) = \left[\delta(\mathbf{k}) + i\,\Phi(\mathbf{k})\right] A(\mathbf{k})\, e^{i\chi(k)}\, e^{i\varphi_{\text{PP}}(k)}. \tag{4.17}$$

Transforming back to real space and expanding to first order yields the final image intensity:

$$I(\mathbf{k}) \propto \Phi(\mathbf{k})\, A(\mathbf{k}) \underbrace{\sin[\chi(k) + \varphi_{\text{PP}}(k)]}_{\text{CTF}(k)}. \tag{4.18}$$

Thus, the CTF becomes:

$$\text{CTF}(k) = \begin{cases} \sin[\chi(k)], & k < k_{\text{edge}}, \\ -\cos[\chi(k)], & k \geq k_{\text{edge}}. \end{cases} \tag{4.19}$$

Figure 4.4 shows the CTF obtained with a Zernike phase plate and the resulting contrast enhancement for a weak-phase object (Unimore logo).

A *hole-free phase plate* (HFPP) is a continuous amorphous-carbon film (without a central

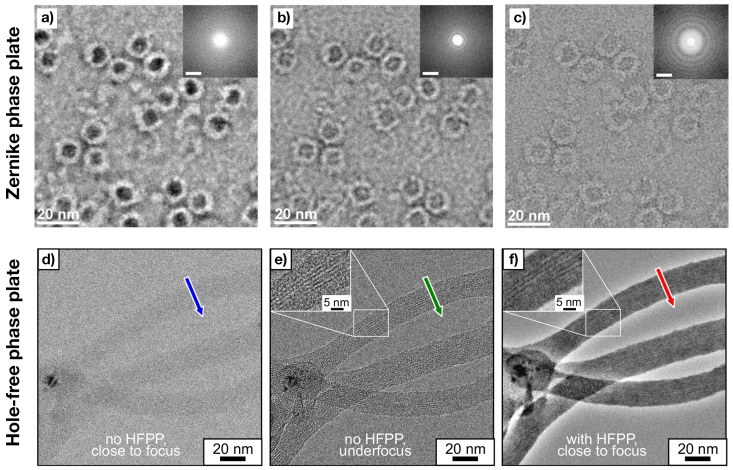


Figure 4.5. Phase-contrast TEM using thin-film phase plates. (a–c) Negatively stained horse spleen ferritin: (a) phase TEM image acquired with a Zernike phase plate, and (b,c) conventional TEM images recorded at defocus values of 2550 nm and 540 nm, respectively [107]. (d–f) Carbon nanotube sample: (d) conventional TEM near focus, (e) underfocus, and (f) with a Hole-free phase plate (HFPP) providing a central phase shift of $\phi_{\mathrm{PP}}(0) = -0.8\pi$ [108].

hole) or a metallic thin film placed in the objective back focal plane. The film is often heated to $\sim 200^{\circ}$C to reduce contamination [108]. In many publications this approach is referred to as the *Volta phase plate*, where *Volta* is the commercial name introduced by FEI (now Thermo Fisher Scientific). When the central (unscattered) beam irradiates the film, a localized beam-induced surface potential (the so-called Volta potential) forms around the illuminated region. A simple model treats this potential as a screened Coulomb-like field [109],

$$V(r,z) = \frac{N_e e}{4\pi\varepsilon_0}\,\frac{1}{\sqrt{r^2+z^2}}, \tag{4.20}$$

Here, r is the radial distance from the illuminated spot in the phase-plate plane, z is the axial coordinate along the electron-beam direction, and N is an effective number of elementary charges describing the net beam-induced charging (total charge $Q = Ne$) which produces an additional phase shift for electrons passing through the illuminated spot relative to the scattered electrons. In practice, this advances the phase of the unscattered reference wave by approximately $\pi/2$ compared with the scattered waves [110].

The phase-shifting region forms *in situ* during exposure and stabilizes after a short pre-irradiation on a fresh area. Image formation can then be described by the standard phase-contrast model with an added constant phase term. In practice, the phase shift is close to $\pi/2$ [110].

Operational benefits include in-focus phase contrast with strong low-spatial-frequency transfer, simple alignment (no hole centering), and a long service life. Limitations include the need for heating, brief preconditioning, and a higher-than-expected signal loss. The HFPP is widely used for near-focus electron cryotomography, where it improves the visibility of weak features in intact cellular specimens [111, 112, 105]. Real examples of contrast enhancement using the Zernike phase plate and HFPP are shown in Figure 4.5.

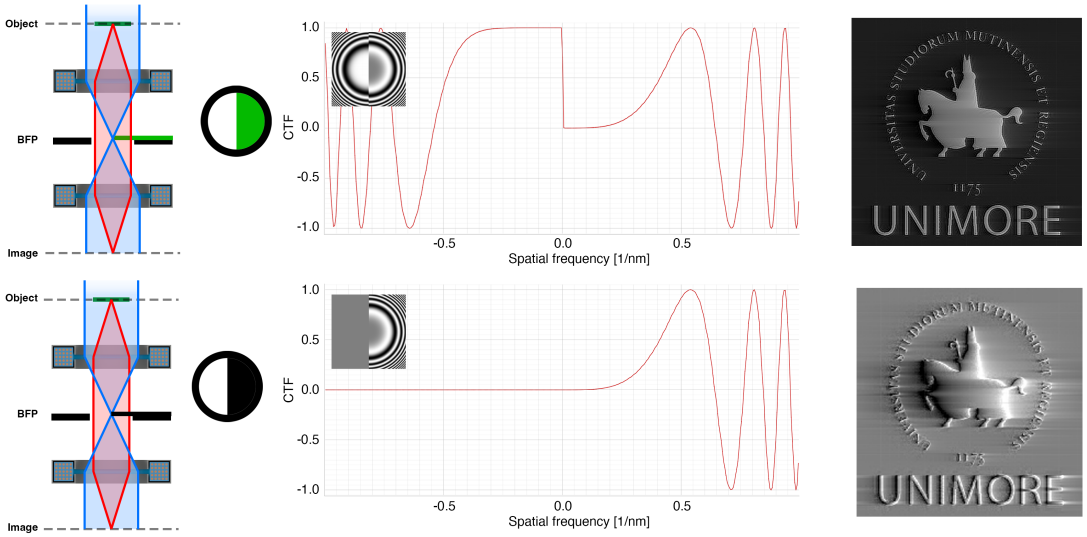


Figure 4.6. Hilbert (top) and Schlieren (bottom) phase plates.

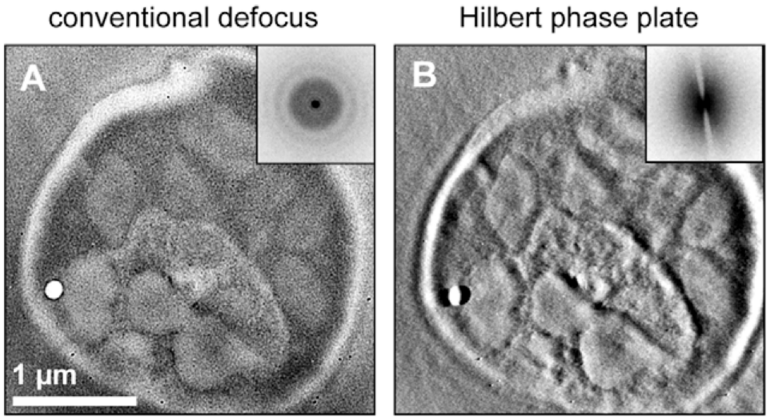


Figure 4.7. Hilbert phase plate on a biological sample.

Hilbert and Schlieren Phase Plates (Directional Contrast)

So far, we have talked about phase plates that symmetrically manipulate these frequencies. Hilbert and Schlieren (Foucault knife edge) are different kinds of phase plates (or imaging methods) that asymmetrically alter the frequencies only for the semicircle.

The Hilbert phase plate (Hilbert differential contrast, HDC) introduces a half–back-focal-plane phase step. A thin amorphous-carbon edge is placed in the objective back focal plane so that one half of the diffraction pattern experiences a phase shift, typically close to π, while the other half remains unshifted (see also Figure 4.6). At near Gaussian focus, it converts weak phase into amplitude, producing an antisymmetric point-spread function and a directional, differential-like image contrast [113]. In practice, HDC provides strong in-focus low-spatial-frequency transfer with a high signal-to-noise ratio, which is useful for tomography [114].

The Schlieren (single–sideband, Foucault knife edge) phase plate implements an *amplitude*

mask at the objective back focal plane: a knife edge blocks one half-plane of spatial frequencies so that only a single diffraction sideband is transmitted (also see Figure 4.6). In the weak-phase limit, this yields a directional, differential-like contrast with an antisymmetric point-spread function and an *odd, single-sideband* transfer, distinct from Hilbert's half-plane *phase* step (odd cosine) and Zernike's even cosine transfer [94].

4.2.2. Sculpted thin-film phase plates

Sculpted thin-film phase plates (STPPs) are passive, nanofabricated membranes that use patterned thickness variations—typically in Si_3N_4, amorphous carbon, or metal-coated films—to impose a designed phase shift on the electron wave. Advances in Focued ion beam (FIB) milling and electron-beam lithography enable nanometre-scale thickness control across micrometre patterns, allowing complex phase maps to be encoded directly into the membrane [115]. When positioned in a condenser or objective aperture plane, these sculpted films act as refractive phase masks that shape the outgoing beam without relying on interference-based holographic encoding [99]. Their straightforward fabrication and compatibility with standard TEM hardware make them an attractive alternative to field-based phase plates [116]. STPPs have been demonstrated for beam shaping and aberration engineering, including a sculpted Si_3N_4 membrane that introduces the opposite-signed phase curvature needed to compensate for spherical aberration in STEM [117, 118]. They have also been used to generate structured beams such as vortex modes, OAM sorters, and tailored illumination profiles [115].
Despite their advantages, STPPs involve several limitations and design trade-offs. Because phase modulation arises from electron–matter interaction, it inevitably introduces elastic and inelastic scattering that reduces coherence and attenuates high-frequency information [116]. Their performance depends on precise thickness control; however, FIB milling can induce surface roughness, ion implantation, and charging that distort the intended phase. STPPs are static devices that offer optimal correction only for a specific convergence semi-angle and accelerating voltage [117]. As a result, they always require balancing the desired phase modulation against scattering-induced signal loss, which sets a fundamental constraint on their overall efficiency.

4.2.3. Electrostatic phase plates

The operation principle of electrostatic phase plates is rooted in the Aharonov–Bohm (A–B) effect, a quantum mechanical phenomenon demonstrating that the phase of an electron wave changes when it passes through a region with an electromagnetic potential (see Section 2.1.3). The phase shift introduced by an electrostatic potential is generally expressed as

$$\Delta\varphi(x,y) = C_E \int V(x,y,z)\, dz, \tag{4.21}$$

where $\Delta\varphi(x,y)$ is the local phase shift, C_E is the electron–energy–dependent interaction constant, and $V(x,y,z)$ is the electrostatic potential generated by the phase plate. By analogy

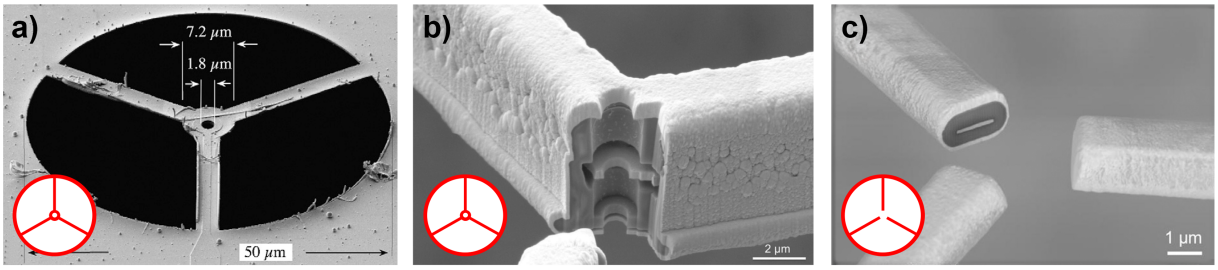


Figure 4.8. Three examples of electrostatic Boersch phase plates: (a,b) einzel-lens configurations formed by five alternating layers of Au and either Al_2O_3 [121] (a) or Si_3N_4 [120] (b); (c) a modified design in which the central electrodes are removed by FIB milling [120].

with light optics, the projected potential acts as an effective refractive index that locally alters the electron wavefront phase. Representative values of C_E at common accelerating voltages are summarized in Table 2.2.

The idea of an electrostatic phase plate was first proposed by Boersch in 1947 [97]. In this design, a miniature einzel lens positioned at the back focal plane (BFP) applies a controlled phase shift to the unscattered electrons. The device consists of a biased central electrode surrounded by two grounded electrodes, forming a highly localized electrostatic potential. The active region of the lens is typically on the order of $\sim 1\,\mu$m in diameter. This potential imposes a tunable phase shift on the central beam, while scattered electrons passing outside the electrode region are unaffected. Under the negligible-lens approximation, the integrated potential in the channel is laterally uniform, and the primary beam acquires a uniform phase shift. The imposed phase is approximately linear with the applied voltage and is calibrated for each device [119, 120].

Although proposed in 1947, only recent advances in nanofabrication have enabled practical realizations. For example, Figure 4.8 shows a device where three supports suspend the active region, formed by five alternating Au/Al_2O_3 layers that insulate the electrodes (Figure 4.8(a)) [121]. Compared with a thin-film phase plate, the Boersch concept is attractive because the phase shift is tunable and the amount of material in the beam path is reduced, which limits unwanted scattering, time-dependent contamination, and charging. Practical limits arise from the need for mechanical supports at the BFP and from the electrode outer diameter: both introduce obstruction, block low spatial frequencies, and create angular gaps.

A similar ring-electrode geometry has also been used in voltage-controlled phase-shifter arrays. In this approach, each aperture is surrounded by an independently biased electrode, enabling *in situ* and tunable phase shifts. By controlling the three ring electrodes individually, the device produces adjustable three-beam interference [122]. This demonstrates how the same Boersch-type structure can be adapted for dynamic and programmable electron-wavefront modulation rather than a fixed phase delay.

Using the same principle, the Zach phase plate replaces the central Einzel lens with a coaxial geometry: a biased inner rod and a grounded outer electrode [125]. As shown in Figure 4.9, it is placed in the BFP or a conjugate diffraction plane. The open tip creates a near field that

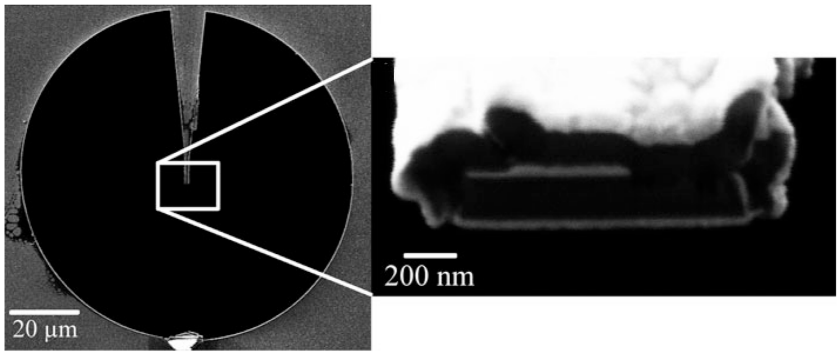


Figure 4.9. Zach phase plate [123, 124]

gives the unscattered beam a distance-dependent, voltage-controlled phase shift. Because only a single rod intersects the plane, the amount of material is small; this simplifies alignment and lowers the risk of charging [126]. However, the rod blocks one side of the spectrum and can create single-sideband contrast that depends on the rod–specimen orientation [124, 123].

Vortex Beams and Orbital Angular Momentum

Electromagnetic and matter waves can be structured such that their phase winds helically around the propagation axis. Such beams, whose transverse phase varies as $e^{i\ell\varphi}$ in cylindrical coordinates (r, φ, z), are referred to as *vortex beams*. The ℓ is the *topological charge*: it counts how many times the phase advances by 2π when circling once around the axis. In the paraxial regime, a convenient model for a vortex mode is

$$\psi_\ell(r, \varphi, z) = A(r, z)\, e^{i\ell\varphi}, \tag{4.22}$$

where $A(r, z)$ describes the radial envelope. Acting with the z-component of the orbital angular momentum (OAM) operator,

$$L_z = -i\hbar \frac{\partial}{\partial \varphi}, \tag{4.23}$$

yields $L_z\psi_\ell = \ell\hbar\,\psi_\ell$, showing that each electron (or photon) carries a well-defined OAM $\ell\hbar$ along the propagation direction [127, 128]. The helical phase implies a *phase singularity* on the axis, where the amplitude vanishes and the phase is undefined; destructive interference around this singularity produces the characteristic donut-like intensity profile with a dark central core.

The winding of the phase around the singularity makes the topological charge robust against moderate perturbations, which explains the characteristic stability of vortex beams compared with conventional Gaussian modes. A schematic illustration of the three-dimensional helical wavefront and the corresponding ring-shaped intensity distribution is shown in Figure 4.10.

The concept of OAM was first established in optics, where paraxial solutions of the scalar wave equation with azimuthal phase factors $e^{i\ell\varphi}$ were shown to carry quantized OAM and to

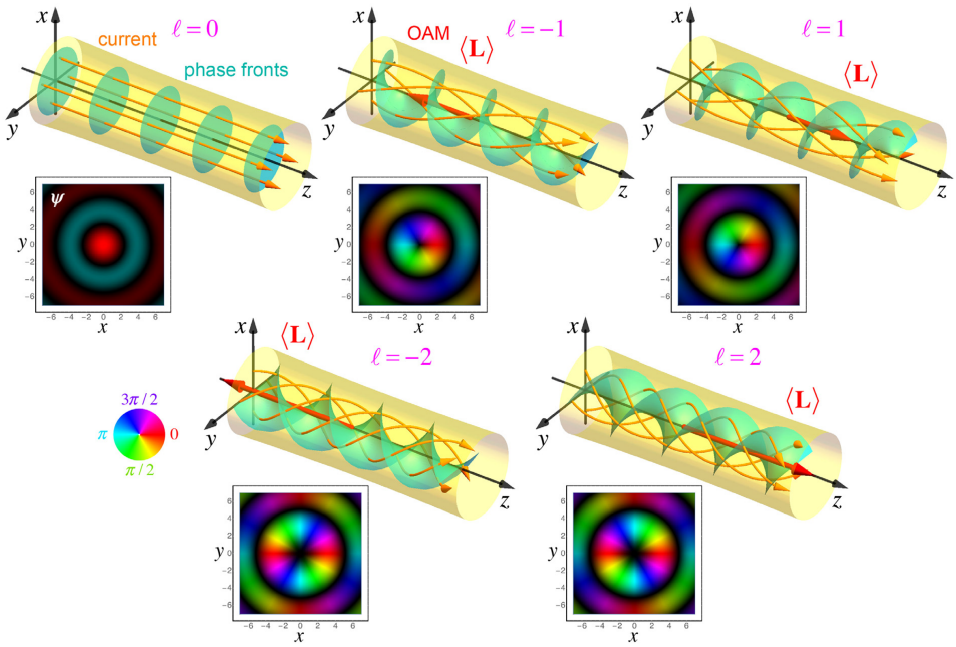


Figure 4.10. Vortex beams with different vortex charges ℓ, showing helical phase fronts (cyan), circulating probability current (orange), and the transverse amplitude–phase distribution at z=0 with the corresponding radial profiles [128].

exhibit ring-shaped intensity profiles [129, 130]. Analogous solutions exist for fast electrons in a transmission electron microscope (TEM). In the paraxial approximation, the electron wave obeys a scalar paraxial wave equation, which has vortex solutions of Bessel or Laguerre–Gaussian type,

$$\psi_\ell(r, \varphi, z) = J_\ell(k_r r)\, e^{i\ell\varphi}\, e^{ik_z z}, \tag{4.24}$$

with longitudinal wave number k_z and transverse wave number k_r [131].

In conventional TEM, the probe is typically approximated as a plane wave or a convergent beam whose phase is as uniform as possible. A vortex beam introduces an additional, controllable degree of freedom: the quantized OAM. This extra degree of freedom enables types of contrast and spectroscopy that are inaccessible, or much less efficient, with unstructured beams.

First, the OAM of an electron beam is associated with a magnetic moment proportional to ℓ, even in the absence of spin polarization. This magnetic moment can couple to internal magnetic fields in a specimen, making vortex beams promising probes for mapping localized magnetic moments [132]. Second, the helical phase structure allows selective coupling to chiral and multipolar excitations. The interaction of a vortex beam with atoms, molecules, or nanostructures can involve the exchange of OAM between the beam and the internal degrees of freedom of the specimen. Dynamical simulations have shown that an incident electron vortex beam can transfer OAM to specific atomic columns in a crystal, leading to atom-dependent scattering patterns that encode information about orbital character and local symmetry [133]. More generally, the distribution of scattered OAM channels can act as a *fingerprint* of the local electronic structure and symmetry, and, in inelastic scattering, of

the topology of excitations such as plasmons, phonons, or magnons.

The existence of OAM modes is only useful experimentally if both preparation and *detection* of OAM are possible. While dedicated beam-shaping optics can create an incoming vortex probe, the wave emerging from a specimen is typically a complicated superposition of many OAM components,

$$\psi_{\text{out}}(r,\varphi) = \sum_{\ell} c_\ell(r)\, e^{i\ell\varphi}. \tag{4.25}$$

To extract the coefficients c_ℓ, one needs an element that converts the azimuthal degree of freedom into a measurable spatial separation: this is the role of an *OAM sorter*.

An OAM sorter is an electron-optical system that implements an almost unitary mapping from helical modes $\{e^{i\ell\varphi}\}$ to spatially separated spots indexed by ℓ in the detector plane [132, 134]. Conceptually, a typical sorter consists of two successive phase transformations placed in conjugate planes. The first element (often called the "unwrapper") performs an approximate log–polar transformation, converting the azimuthal dependence of the wave into a linear phase ramp. The second element (the "corrector") compensates for residual distortions so that, after propagation, each OAM component focuses into a distinct lateral position. The resulting intensity distribution on the detector directly yields an "OAM spectrum" $I(\ell)$ or, in the spectroscopic case, a two-dimensional dataset $I(\ell, \Delta E)$ for OAM-resolved EELS [135].

Vortex beam generation with holographic fork holograms

A widely used method to generate electron vortex beams is the holographic "fork" (or pitch-fork) grating, implemented as a synthetic hologram on a thin Si_3N_4 membrane.[136, 137, 99] A fork hologram encodes the interference pattern between a plane wave and a helical wave $\exp(i\ell\varphi)$, such that the resulting grating contains a phase dislocation at its center. When this grating is illuminated by a quasi-parallel electron beam, each diffraction order n carries an orbital angular momentum (OAM) of $n\ell\hbar$. By selecting, for example, the +1 diffraction order with an aperture, one obtains an isolated vortex beam with topological charge ℓ at the specimen plane.

Fork holograms can be realized as amplitude, phase, or mixed amplitude–phase gratings, using focused ion beam (FIB) milling or electron-beam lithography.[99] Phase and blazed designs improve the diffraction efficiency and allow the generation of high-ℓ beams, but the method is still intrinsically diffractive: several diffraction orders are produced, and part of the current remains in the zero order. The pattern is also static, so changing ℓ requires exchanging the hologram.

Generation of electron vortex beams using electrostatic phase plates ("chopsticks")

Figure 4.11 shows a MEMS based vortex beam generator with "chopstick" developed in our group. Electrostatic "chopstick" phase plates generate vortex beams by imprinting a continuous azimuthal phase onto the electron wavefront. In practice, two micron-scale electrodes

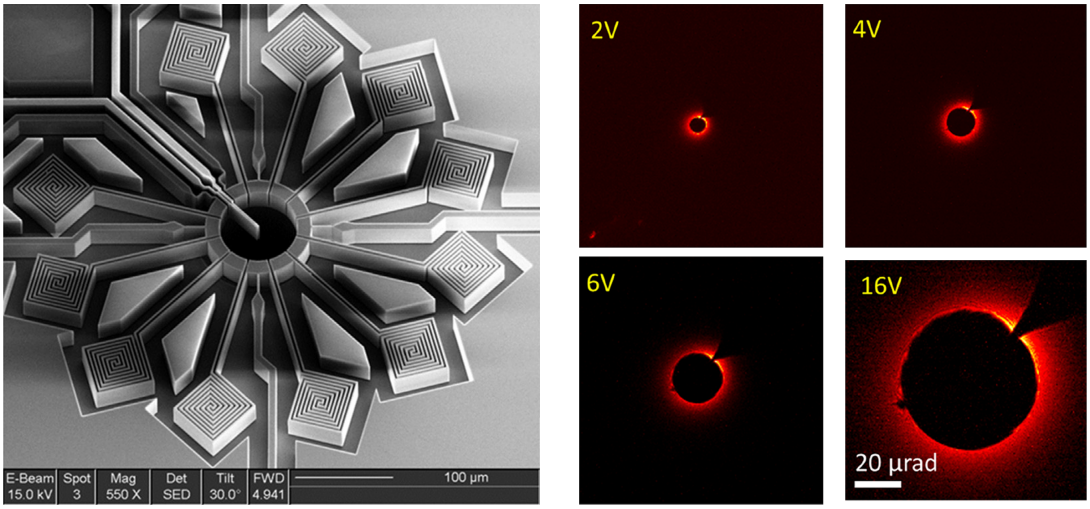


Figure 4.11. MEMS based Vortex beam generator with chopstick (left) and Vortex beams (recorded in the diffraction plane of the aperture) shown as a function of bias voltage applied to one of the chopsticks (right) [98].

extend into the beam and are biased symmetrically to $\pm V$ across a narrow gap [138, 139]. The resulting transverse electric field is approximately azimuthal around the gap, so that electrons picking up a dynamical phase along different azimuthal paths acquire a net helical phase factor $\exp(i\ell\varphi)$. The effective topological charge ℓ, not quantized in this case, grows with the applied voltage and depends on the detailed device geometry.[98]

Modern implementations place the chopstick electrodes on a MEMS chip mounted in a condenser or objective aperture plane. In this configuration, the device acts as a tunable OAM source with a single main output beam and no higher diffraction orders, in contrast to holographic gratings. The same chip can be used to scan ℓ continuously, to calibrate OAM-sensitive measurements, and to provide clean reference beams for alignment and for the OAM sorters.

OAM Spectroscopy of electron vortex beams using electrostatic phase plates ("OAM sorter")

Beyond serving as tunable OAM sources, similar MEMS-based electrostatic elements can also be configured as orbital angular momentum (OAM) sorters. The MEMS based OAM sorter phase plate developed in our group in collaboration with Forschungszentrum Jülich demonstrated a two-element electrostatic sorter that implements a log–polar conformal mapping of the electron wavefront, separating different ℓ components into laterally displaced spots at the detector with an experimental resolution of $\Delta\ell \approx 1.1$–2 and nearly lossless throughput.[135] In this design, the first element $S1$ (the "unwrapper") applies a phase distribution [140]

$$\varphi_1(u) = \frac{s}{\lambda}\,\mathrm{Re}\left\{u\ln\left(\frac{u}{L}\right) - u\right\}, \tag{4.26}$$

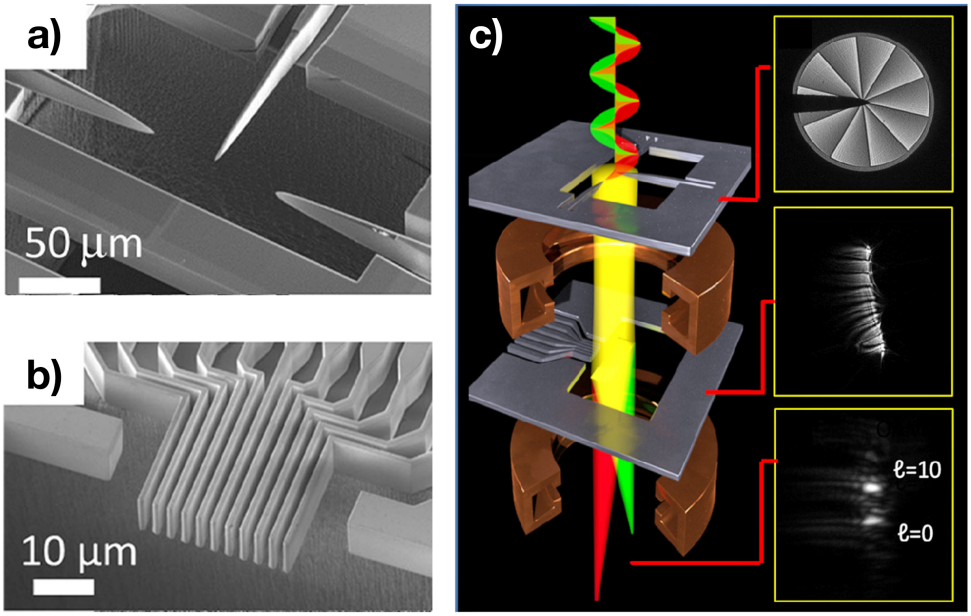


Figure 4.12. SEM images of electrostatic OAM sorter elements (a) S1 (unwrapper) and (b) S2 (corrector). (c) Schematic diagram of part of an electron microscope column that contains a sorter, with experimental images showing the evolution of the electron wave function [135].

while the second element $S2$ (the "corrector") introduces a compensating phase

$$\varphi_2(u) = \frac{Ls}{\lambda} \operatorname{Re}\left\{\exp\left(\frac{u}{sf}\right)\right\}, \tag{4.27}$$

where $u = x + iy$ is the complex transverse coordinate, λ is the electron wavelength, s is a scaling factor, L is a characteristic length, and f is an effective focal length. Together, these tailored phase profiles implement the log–polar mapping and reflatten the wavefront, so that different OAM modes emerge as well-separated intensity peaks suitable for OAM-resolved imaging and spectroscopy [135].

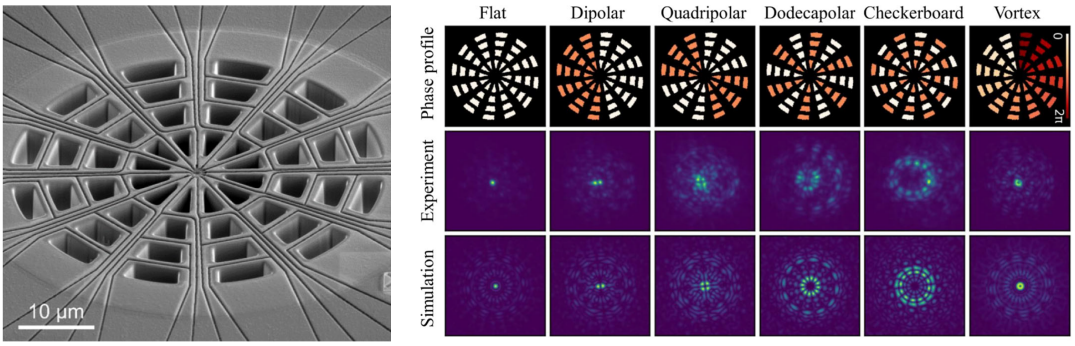


Figure 4.13. 48-element programmable phase plate (left) and the three rows of figures on the right. From top to bottom, the phase configurations set on the phase plate, the simulated probe shapes, and the resulting experimental probe images, respectively [141].

Vortex beams from programmable MEMS phase plates

A more general and programmable approach uses MEMS-based electrostatic phase plates with many individually addressable pixels, such as the 48–element device developed by Verbeeck and co-workers.[142, 141] In this design, each pixel is a through-hole in a Si_3N_4 membrane whose inner wall is coated with a metallic electrode (also see Figure 4.13). By applying different low voltages to each electrode, the local electrostatic potential inside each aperture modifies the phase of the transmitted electrons, with sub-milliradian phase resolution at typical TEM voltages.

Vortex beams are obtained by programming an approximate azimuthal phase ramp across the pixels, so that the transmitted wavefront reproduces $\exp(i\ell\varphi)$ to good accuracy. Because the device is reconfigurable, the same hardware can switch between different ℓ values, superpositions of OAM modes, or completely different phase profiles, such as defocus, aberration patterns, and Hadamard-like masks. In this sense, the 48–element phase plate acts as the electron analog of a spatial light modulator (SLM), enabling dynamic and versatile vortex-beam generation inside a conventional TEM.

4.2.4. Magnetic Phase Plates

Zernike phase contrast can also be produced by a magnetic vector potential via the Aharonov Bohm (A–B) effect. In the Aharonov–Bohm effect, electrons acquire a phase shift even in field-free (i.e. B=0) regions if their trajectories enclose magnetic flux. The accumulated phase is given by [143]

$$\Delta\varphi \;=\; -\frac{e}{\hbar}\oint \mathbf{A}\cdot d\mathbf{l} \;=\; -2\pi\,\frac{\Phi}{\Phi_0}$$

where $\Phi_0 = h/e = 2.07\times10^{-15}\,T\,\mathrm{m}^2$ is the flux quantum. For more details, see section 2.1.3.

In practice, the required magnetic vector potential field can be realized in two broad ways: (i) using *magnetic materials* that confine flux (e.g., toroidal or closed-flux elements, magnetized needles or rings), and (ii) using *Ampère-current elements* (micro-coils, current-carrying wires) that generate a controllable flux.

Depending on where the element is placed in the microscope, these implementations serve two complementary roles. Placed at the objective back focal plane (or a conjugate diffraction plane), they act as *phase-contrast plates*, imparting a controlled reference-wave shift (Zernike-like, typically near $\pi/2$) to enhance in-focus low-spatial-frequency transfer. Inserted in the illumination path (condenser aperture planes), they function as *beam shapers*, imposing designed phase profiles such as helical phases $e^{i\ell\varphi}$ for vortex/OAM beams or the harmonic phases used by OAM sorters [144].

Several groups have explored different strategies to realize phase modulation through confined magnetic fields. A representative implementation of a magnetic phase plate for Zernike-type contrast was reported by Edgcombe *et al.* [146], who investigated ferromagnetic thin-film rings operating through the Aharonov–Bohm effect. When magnetized into a vortex state, the confined magnetic flux within the ring produces a well-defined vector potential in the

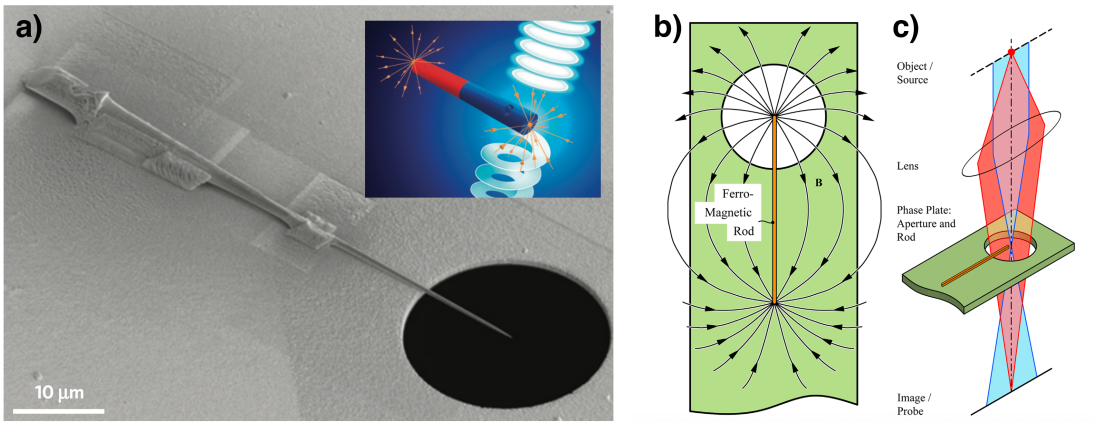


Figure 4.14. Magnetic-material phase plates. (a) Magnetized Ni needle generating an electron vortex. (b) CoFe rod in the condenser forming a vortex probe. (c) The samerod in the objective yielding spiral-type phase contrast (Zernike–Hilbert) [144, 145].

surrounding vacuum. Electrons passing through and outside the ring acquire a relative phase shift $\Delta\phi = (e/\hbar)\Phi$, which reaches $\pi/2$ for a flux $\Phi \approx h/4e$, equivalent to the phase condition of a Zernike phase plate. Béché *et al.* [145] demonstrated that a magnetized Ni needle can mimic a magnetic monopole at its apex, producing an Aharonov–Bohm phase that converts a plane electron wave into a vortex beam with quantized orbital angular momentum (also see Figure 4.14). Shortly after, Blackburn and Loudon [144] introduced a *magnetic spiral phase plate* based on a narrow CoFe rod mounted at an aperture plane. Its stray field imposes an azimuthal phase winding that (i) produces an electron *vortex probe* when the rod is used in the illumination path (condenser aperture), with the topological charge set by the enclosed flux ($\Phi \approx n\,h/e$), and (ii) yields a *spiral-type phase-contrast transfer* when placed at the objective aperture, enhancing low-frequency contrast with characteristics intermediate between Zernike and Hilbert plates when $\Phi \sim h/e$ (maximum phase shift $\approx 2\pi$).

An Ampère (current-driven) phase plate uses the magnetic vector potential around a short, current-carrying wire placed parallel to the electron beam to impose a tunable phase in the back focal plane. Positioned beside the unscattered beam (see Figure 4.15), it acts like a controllable Zernike phase shifter: the phase increases with current and reverses sign when the current is inverted. Because the device contains only a small wire segment and no ferromagnetic material, it introduces little matter into the plane, avoids dielectric charging, and remains stable in the strong objective field. A one-sided obstruction from the wire can create single-sideband contrast that depends on the wire orientation; this can be reduced by orientation choice and corrected a posteriori in Fourier space. Small residual lens effects (e.g., slight astigmatism) are handled by routine focus and stigmation adjustments. Overall, the Ampère phase plate provides simple, reversible, and substrate-free phase tuning suitable for near-focus imaging [147]. More recently, Pozzi *et al.* [134] extended this principle to design a magnetic implementation of the OAM sorter, in which current-carrying or magnetized elements reproduce the same phase transformations as electrostatic ones. Such devices exploit the formal analogy between electrostatic potentials and magnetic vector potentials, offering

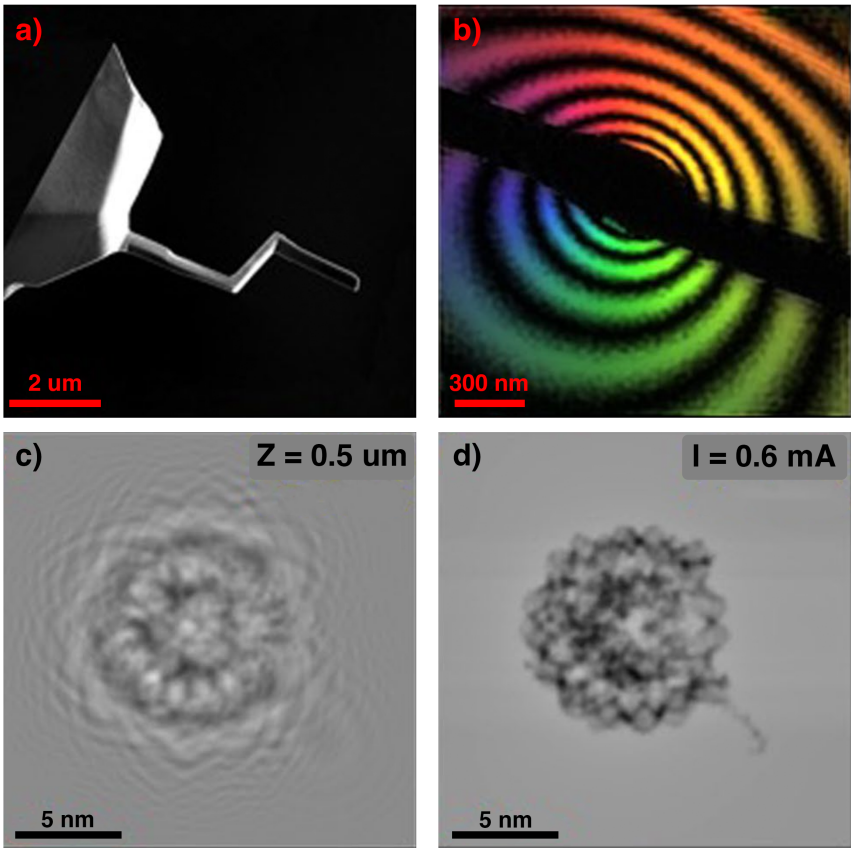


Figure 4.15. Ampère (current-driven) phase plate. (a) Electron micrograph of the tunable Ampère phase plate (TAPP). (b) Off-axis holography phase maps of the surrounding vacuum region for a 4 mA current. (c) Simulated out-of-focus Fresnel images of an NCP at 0.5 µm defocus. (d) Simulated TAPP images for a 0.6 mA current.

potential advantages of reduced charging and energy-independent phase response. Although these magnetic approaches are experimentally demanding and not yet widely adopted, they highlight the versatility of magnetic vector potentials for phase control in electron optics. Magnetic implementations can reduce the material in the beam path and avoid dielectric charging. Current-driven designs provide simple tunability through the supply current. However, care is needed to limit stray fields and hysteresis, stabilize magnetization or current, and minimize mechanical supports that may obstruct low spatial frequencies [148].

4.2.5. Photonics Phase Plates

Photonics (laser-based) phase plates use the interaction between free electrons and a coherent optical field to control the phase of the electron wavefront without inserting any material into the beam path [149]. Since, to the first order, the electron-photon interaction averages to zero over a period, scientists have to find different approaches to produce a phase effect on electrons. The current main ideas are to use a more intense laser beam and use second order effects or take advantage of near-field effects in the presence of material. The former process in a quantum description, can be treated as a *second-order stimulated Compton process*: the electron interacts with a single optical mode by absorbing and re-emitting a photon from the

same mode, so there is no net energy exchange, but the electron wave acquires a controllable phase delay [150, 151]. The accumulated phase shift increases with optical intensity (photon density) and with the interaction time, and it also depends on the laser wavelength. For more relativistic (higher-energy) electrons, the same optical field produces a smaller phase shift, so stronger fields or longer interaction lengths are needed [150].

In the classical limit, the same physics is often explained using the *ponderomotive* picture: an oscillating light field acts as an effective potential that pushes the electron away from high-intensity regions while also imprinting a phase modulation that follows the optical intensity pattern [152, 153]. This provides a direct way to implement Zernike-type phase contrast by shifting the phase of the unscattered beam relative to the scattered components, for example, by placing the unscattered beam at a chosen point of a standing-wave field [153, 154]. More generally, by shaping the optical field geometry, laser-based phase plates can synthesize tailored electron wavefronts with high tunability and minimal material-induced scattering [155].

Unlike electrostatic or thin-film phase plates, ponderomotive phase plates are *non-invasive*, *tunable*, and immune to *charging* or *contamination*, since the interaction occurs entirely in free space through the electromagnetic field. This consideration applies a bit less in the case of near field interaction where a material must be present within few micron from the point of interaction, this has not stopped the possibility of a few important project related to near field beam shaping [156]. Depending on how the optical field is generated and temporally shaped, three main classes of photonic phase plates can be [100]:

1. **Continuous-wave laser phase plates**: a resonant optical cavity maintains a steady, controllable phase retardation [157];

2. **Pulsed-laser phase plates**: short intense pulses produce transient ponderomotive potentials;

3. **PINEM-based phase plates**: near-field or evanescent optical fields exchange quantized energy with the electron beam, leading to coherent phase modulation through photon absorption and emission processes [156].

Continuous-wave laser phase plates

In a continuous-wave (CW) laser phase plate, a steady ponderomotive potential is created by a high-intensity standing light wave, so the electron phase can be shifted without inserting any material into the beam path. Schwartz *et al.* (2019) demonstrated this concept by integrating a near-concentric Fabry–Perot optical cavity into a TEM and resonantly enhancing a CW 1064 nm laser to reach local intensities on the order of tens of GW cm^{-2} at the optical antinodes [157]. By placing the unscattered beam at a single antinode in the diffraction plane, the unscattered component acquires a controlled retardation relative to the scattered electrons, analogous to Zernike phase contrast.

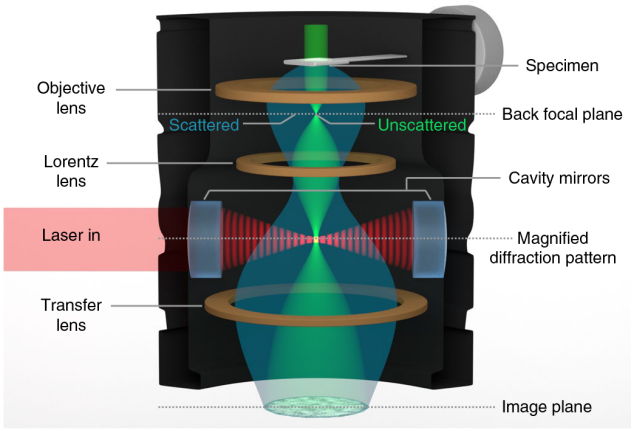


Figure 4.16. Laser-based control of the electron phase in a TEM: a resonantly enhanced standing laser wave in an optical cavity imparts a phase shift to the unscattered beam in the objective-lens back focal congugate plane, which is then recombined with the scattered beams to form phase-contrast images [157].

Because the interaction occurs in free space, CW laser phase plates avoid charging, contamination, and unwanted scattering that can affect solid or electrostatic phase plates. The phase shift is continuously tunable through laser power and cavity alignment, and experiments showed improved low-spatial-frequency contrast near focus, consistent with a modified contrast transfer [157]. In the reported implementation, the achieved retardation was below the optimal $\pi/2$, but the results indicate a clear path toward larger phase shifts by increasing the cavity enhancement and available optical power [157].

Pulsed-laser phase plates

Pulsed-laser phase plates replace the continuous optical field with ultrashort laser pulses that transiently modify the phase of a passing electron wave. The mechanism is the same as for continuous-wave laser phase plates: the phase shift is set by the optical intensity experienced by the electrons, but the interaction exists only during the femtosecond or picosecond pulse. Therefore, the electron beam must be synchronized to the laser so that they overlap in space and time at the back focal plane, for example using a laser-driven photocathode or RF phase locking. When the unscattered beam is aligned to the intensity maximum, it acquires a controlled retardation relative to the scattered components, analogous to a Zernike phase plate.

Because the interaction is purely electromagnetic, pulsed-laser phase plates are material-free and avoid charging, contamination, and scattering from solid supports. The phase shift is tunable via pulse energy, duration, and wavelength, and the timing enables time-gated phase contrast in ultrafast TEM. Recent work has demonstrated reproducible, spatially structured phase modulation without net energy exchange, supporting pulsed-laser fields as a flexible route toward optically programmable phase manipulation in pump–probe electron microscopy [158].

Part II

Concepts and Methods

5 Design Philosophy and Methodology

Contents

5.1 Phase plates: Why They Have Reached Their Practical Limit? 87

5.2 Motivation for MEMS-based Electrostatic Phase Plates 88

5.3 Current controlled electrostatic phase plate 91

5.4 Sandwich chip design 95

In the previous chapter, we introduced the general principles of electron beam shaping and reviewed the main classes of phase plates used in transmission electron microscopy. We discussed how these devices can be used either for contrast enhancement or for more general wavefront engineering, and we highlighted the physical mechanisms that each design relies on. Building on this foundation, the present chapter focuses on electrostatic phase plates, with particular attention to MEMS-based implementations. These devices offer several advantages over other phase plates but still face practical limitations that have prevented their full integration into routine TEM workflows.

The goal of this chapter is to clearly establish these advantages and limitations and to motivate the developments presented in the rest of the thesis. A central part of this work addresses the main barriers of MEMS phase plates by introducing and investigating a current-controlled electrostatic concept designed to improve stability and tunability. We also explore a sandwich-type geometry to localize the electric potential more effectively and achieve more precise beam shaping.

5.1. Phase plates: Why They Have Reached Their Practical Limit?

Despite their potential to expand the capabilities of TEM through controlled wavefront shaping, phase plates remain limited by practical constraints that are difficult to overcome. Material-based phase plates, such as the Zernike and Volta designs, suffer from fundamental physical limitations that arise whenever a thin film is inserted into the electron beam. In Zernike plates, even a few seconds of irradiation are enough to build up charges around the central hole, driven by contamination layers that act as electrical insulators. This charging produces parasitic electrostatic fields that distort the intended phase shift and create fringing halos, cut-on frequency drift, and low-frequency information loss. The Volta phase plate avoids the geometric discontinuity of a hole, but it still relies on a beam-induced surface potential that evolves with dosage. As a result, the phase shift is not constant: it starts small, approaches the desired $\pi/2$, and continues to grow, leading to overcharging, pseudo-astigmatism, and contrast inversion. In both cases, the interaction between the beam and the film surface creates an unstable phase profile that limits reproducibility and lifetime.

Beyond charging, the presence of material in the beam path introduces additional drawbacks that are intrinsic to the physics of electron–matter interaction. Any film, even a few nanometers thick, scatters electrons elastically and inelastically, damping the image signal across all spatial frequencies and lowering the signal-to-noise ratio. Moreover, thin film phase plates are energy-dependent, meaning that a Zernike plate designed for 200 kV will not provide the

same phase shift at 300 kV. In addition, thin films age rapidly under electron irradiation, accumulate contamination, and require periodic replacement or repositioning. Together, these mechanisms establish a clear limitation: as long as a phase plate requires a physical film to imprint phase contrast, it cannot provide stable, tunable, and high-resolution performance. This is the central motivation for replacing material phase plates with electrostatic MEMS devices, which generate the phase shift while minimizing the amount of material (in many cases eliminating it entirely) in the electron beam and thus avoid these inherent sources of instability and information loss.

While material phase plates suffer from direct scattering and charging, laser/ponderomotive and magnetic phase plates avoid these issues but introduce a different set of limitations. Ponderomotive phase plates require extremely high optical intensities sustained inside a high-finesse cavity, making their performance strongly sensitive to thermal drift, mirror deformation, and Johnson noise from surrounding conductive parts. These effects limit phase stability and can degrade the resolution of the microscope. In addition, a standing laser wave acts as an electron phase grating, leading to Kapitza–Dirac diffraction into unwanted higher orders that are difficult to suppress and can complicate image interpretation [154]. Magnetic phase plates typically require microfabricated magnetic structures in (or near) the diffraction plane, which can partially obstruct the Fourier space and introduce a finite cut-on that attenuates low-spatial-frequency transfer. In addition, their magnetic state can be perturbed by the strong objective-lens field, giving rise to hysteresis, domain instabilities, and stray-field artifacts such as apparent astigmatism or beam deflection.

5.2. Motivation for MEMS-based Electrostatic Phase Plates

Electrostatic phase plates, particularly MEMS-based implementations, offer capabilities that overcome many of the limitations of material- and field-based phase plates. Electrostatic MEMS devices provide a route to stable, tunable, and programmable phase manipulation inside a TEM. For these reasons, this thesis focuses on electrostatic, MEMS-fabricated phase plates as the most promising platform for advanced beam shaping and adaptive electron optics. In this section, we summarize their main advantages and the remaining limitations that motivate the work presented in the following chapters. Electrostatic MEMS phase plates exhibit advantages such as:

1. **Precision fabrication, cost, and scalability.** A key advantage of MEMS-based phase plates is the use of semiconductor nanofabrication, which provides a level of geometric precision that is difficult to achieve with conventional machining. Silicon nanofabrication provides nanometer-scale control over electrode geometry, which is essential for setting the electrostatic boundary conditions. Because these devices are fabricated in batches on silicon wafers, a single process run can produce tens to hundreds of nominally identical chips. This wafer-level fabrication reduces the cost per device and greatly improves reproducibility compared with individually machined devices, enabling consistent performance.

2. **Electro-mechanical efficiency and low-voltage operation.** One important advantage of MEMS-based phase plates is their micrometer-scale active area, where electrodes can be positioned very close to the electron-beam path. This geometry enables large electric-field gradients with comparatively low applied voltages; for example, in spiral phase-plate designs, a full 2π phase modulation can be achieved with voltage steps on the order of tens of millivolts, depending on the interaction length and electrode geometry. Although MEMS-based devices suffer from a finite thickness along the z direction, typically in the range of 1–50 μm, which limits the interaction length and thus the achievable phase shift for a given potential, this drawback is largely compensated by the strong fields produced in the small active region. The same scaling is especially beneficial for multipole-type designs, where the voltage required for a given phase profile decreases rapidly with the device radius. In addition, commercial silicon-based MEMS processes provide mechanically stiff and electrically well-defined structures, helping to maintain stable electrostatic fields under beam exposure and ensuring reliable long-term operation.

3. **Instrument-compatible integration and tunability** MEMS-based phase plates are realised as thin chips that can fit inside the narrow gap of high-resolution objective lens pole pieces without major modifications to the microscope column. They can be mounted on commercial electrostatic or biasing holders that provide multiple electrical feedthroughs, so that a standard TEM aperture or holder position effectively becomes a programmable phase module. Unlike a Zernike film with a fixed thickness (and thus a fixed phase shift), an electrostatic MEMS device is fully tunable: the phase shift can be adjusted in real time by changing the applied voltages without breaking vacuum.

Despite these promising advantages, MEMS-based electrostatic phase plates are not free from limitations. Their practical implementation still involves a number of engineering and operational challenges that must be addressed before they can become a robust, routinely used component in high-resolution TEM. In the following, we briefly summarize the main limitations relevant to this thesis:

1. **Pull-in instability.** Although electrodes fixed on a substrate have mechanical support strong enough to resist electrostatic attraction, MEMS phase plates that include suspended electrodes (such as "chopstick" structures in Figure 4.11) are fundamentally limited by pull-in instability. As the voltage increases, the electrostatic force grows faster than the mechanical restoring force, and beyond a critical displacement, the suspended electrode collapses onto neighboring electrodes, causing stiction, electrical failure, or short circuits. This sets a strict limit on the achievable tuning range and requires conservative safety margins in both geometry and applied voltage [159].

2. **Voltage ceiling and vacuum breakdown.** Even when mechanical stability is ensured, the maximum usable voltage on a MEMS phase plate is limited by vacuum breakdown and field emission. In the high-vacuum environment of a TEM, breakdown

is driven by vacuum arcs initiated by field emission from microscopic asperities and sharp edges on the electrodes, so devices can fail at relatively low bias voltages when gaps are on the micrometer scale.

3. **Dielectric charging and hysteresis.** MEMS phase plates are composite structures that combine conducting silicon with insulating layers, typically silicon nitride or silicon oxide. Under exposure to the electron beam and scattered electrons, these insulating regions can trap charge and develop local "patch" potentials that add unwanted electrostatic effects on top of the programmed field [160]. The resulting charging can cause slow, time-dependent drift of the phase shift during acquisition, and hysteresis effects where the optical state of the device depends on its previous voltage and dose history. Because charge trapping is often non-uniform, it can also introduce asymmetric fields and astigmatism. Partial mitigation is possible by adding conductive coatings or shielding electrodes around the dielectric regions, but this increases fabrication complexity and introduces a risk of short circuit.

4. **Limited number of controllable electrodes.** Although microfabrication allows us to define electrode structures on the micrometer scale, the number of independently addressable electrodes is still limited by the available electrical feedthroughs in commercial TEM biasing holders. In practice, only a few to a few tens of channels are available, which constrains the complexity of the boundary potential that can be biased. This limitation makes it difficult to implement highly resolved phase plates or finely tuned phase profiles, which are crucial for high-precision beam shaping and fully programmable wavefront control.

This thesis provides principal contributions toward overcoming the fundamental limitations of conventional electrostatic phase plates and introduces new device architectures for advanced electron-beam shaping. These contributions can be summarized as follows:

- **Current-controlled electrostatic phase plates.** I introduce and develop the concept of current-controlled electrostatic phase plates, which allows finer control of the boundary potential with a limited number of electrical contacts. This strategy improves tunability and reduces the constraints imposed by the limited feedthroughs in commercial TEM holders.

- **Sandwich fabrication strategy.** I propose and implement a new *sandwich* fabrication approach in which electrodes and dielectric layers are stacked to better confine the electric potential, improve breakdown performance, and increase the mechanical and electrical robustness of MEMS phase plates.

- **Redesign of phase plate architectures.** Building on these concepts, I redesign and design several families of electrostatic phase plates, including vortex beam generators, OAM sorters, and electrostatic multipoles, adapting them to the current-controlled and sandwich geometries to enhance their functionality and tunability.

- **Advanced numerical design and optimisation tools.** I develop gradient-based optimisation methods and high-precision numerical workflows in Python and COMSOL to compute, optimise, and validate the phase profiles produced by complex electrode layouts before fabrication, providing a systematic route from target phase map to device geometry.

- **Machine-learning-based control and aberration-orientation diagnosis for multipole phase plates.** I implement deep neural networks for image-based analysis and on-the-fly control of phase plates with millisecond-scale response

- **Aberration correctors** For the main part of this thesis, I design both electrostatics and current-controlled electrostatic multipoles, in particular hexapoles, with the specific goal of enabling aberration correction using MEMS-based electrostatic phase plates.

5.3. Current controlled electrostatic phase plate

Electrostatic phase plates are commonly fabricated on silicon-on-insulator (SOI) wafers. The purpose of these devices is to convey a well-defined electric potential from the external holder biases to the active region, where the electron beam acquires a phase shift. For this reason, it is essential that all conductive paths—from the bonding pads to the electrode structures—have sufficiently low resistance. Although highly doped silicon is typically used as the device layer, it is standard practice to deposit a thin gold film across the entire surface. This metal layer ensures uniform conductivity, suppresses charging under electron irradiation, and stabilizes the potential distribution.

The electrodes in these phase plates are usually patterned as individual elements, each driven independently and kept in an open circuit (also see Figure 5.1(a)) . This configuration is the reason they are referred to as electrostatic phase plates: the phase modulation is generated by static electric fields without requiring current flow. However, in some designs—such as certain vortex generators—closed-circuit geometries are used. In these cases, a single electrode may connect two bonding pads, allowing a current to flow through the structure. Figure 5.1(c) illustrates a simple closed-circuit layout.

Closed circuits introduce two important limitations. First, any current passing through the device produces Joule heating, which increases the temperature of the membrane and limits the maximum voltage that can be safely applied. Second, the path length from the pads to the active area can be several millimeters. In such cases, most of the voltage drop occurs along the long, narrow leads rather than across the active region, making precise phase control difficult or even impossible.

A more fundamental limitation concerns the number of electrodes available at the boundary of the active area. In practice, this number is small, which restricts the complexity of the potential landscape that can be generated. For example, if we want to create a smooth radial or azimuthal potential gradient, what options do we have? Figure 5.1 illustrates the two conventional strategies: (i) increasing the number of electrodes, which is technologically

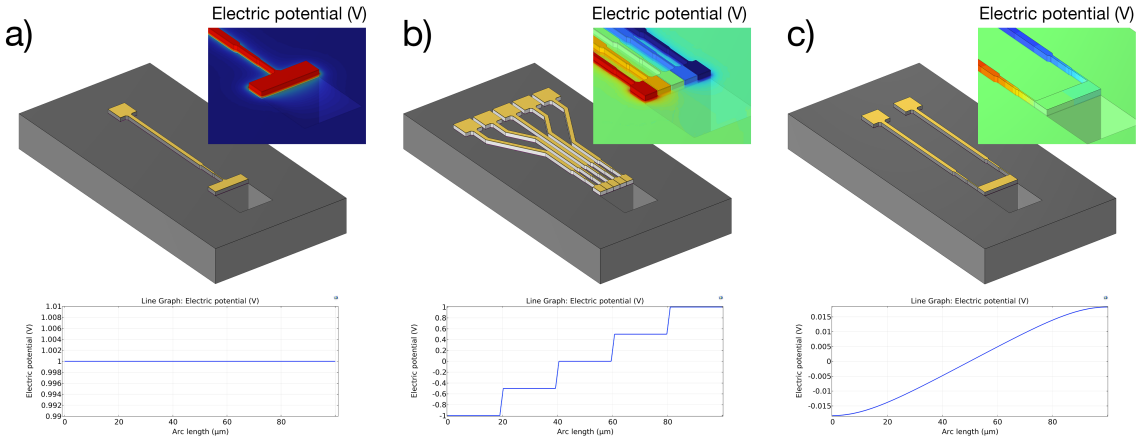


Figure 5.1. Three commonly used electrode layouts in conventional electrostatic phase-plate designs. (a) An open-circuit electrode that provides only a uniform potential. (b) A multi-electrode arrangement used to approximate a desired potential gradient. (c) A closed-circuit electrode in which the applied current creates a distributed voltage drop.

constrained by fabrication limits and wiring density, and still produces a step-like potential profile (Figure 5.1(b)); or (ii) using a closed-circuit design with large applied voltages, which is constrained by heating and voltage-drop issues (Figure 5.1(c)). In both cases, the approaches are impractical for generating arbitrary or high-resolution phase profiles.

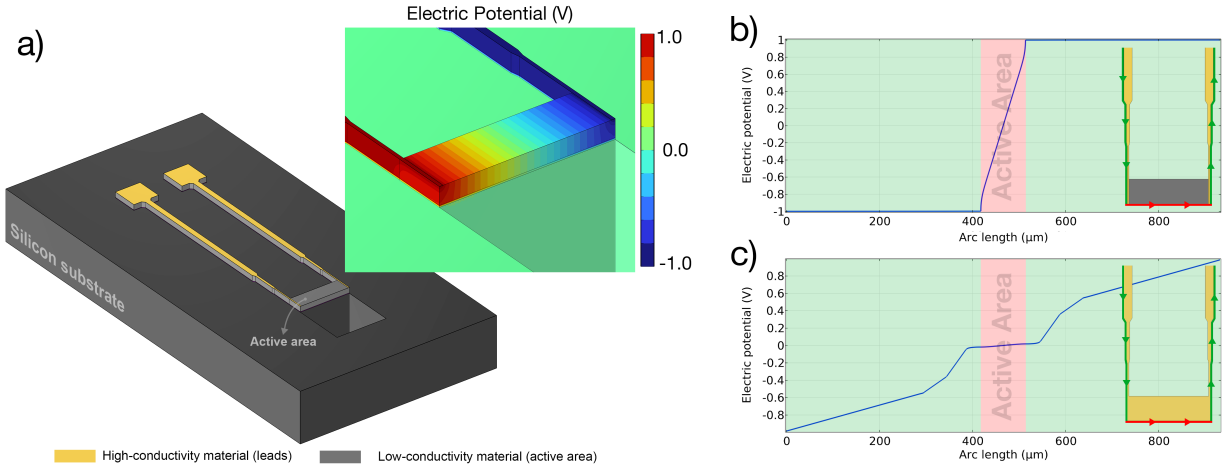


Figure 5.2. Current-controlled electrostatic phase plate. (a) Geometry and simulated electric potential distribution, showing a localized drop at the active area enabled by combining two materials with different electrical conductivity (high-conductivity leads and a lower-conductivity active area). (b) Line profile of the electric potential along the circuit, demonstrating that the voltage drop occurs almost exclusively at the active area. (c) Conventional circuit with uniform conductivity, in which the voltage drop is distributed along the entire path according to the local resistance.

The current-controlled electrostatic phase plate introduced in this thesis is designed to overcome the limitations described above. The central idea is to use two materials with different conductivities to control where the potential drop occurs. The first material is gold (or another commonly used, non-oxidizing metal in nanofabrication), which provides very high electrical conductivity and is standard in micro-fabricated circuits. The second material is

a moderately doped silicon layer, whose conductivity is significantly lower and can be tuned during wafer selection. As shown in Figure 5.2, when two voltages are applied to the closed circuit, the potential drop occurs almost entirely across the low-conductivity silicon region. This creates a well-defined potential gradient directly at the active area of the phase plate.

Mathematically speaking, under steady-state DC conditions, charge does not accumulate in the conductor, so the current density satisfies the continuity equation

$$\nabla \cdot \mathbf{J} = 0. \tag{5.1}$$

Using Ohm's law in local form, $\mathbf{J} = \sigma \mathbf{E}$, together with $\mathbf{E} = -\nabla V$, gives

$$\mathbf{J} = -\sigma \nabla V \qquad \Rightarrow \qquad \nabla \cdot (\sigma \nabla V) = 0. \tag{5.2}$$

If the conductivity is uniform ($\sigma = \text{const.}$), Eq. (5.2) reduces to the Laplace equation $\nabla^2 V = 0$. In regions where σ is constant, the potential therefore varies linearly along the transport direction.

Consider a one-dimensional stack of n regions in series, each with length L_i and conductivity σ_i, and a constant cross-sectional area A. In 1D, Eq. (5.2) becomes

$$\frac{d}{dx}\left(\sigma \frac{dV}{dx}\right) = 0 \quad \Rightarrow \quad J \equiv \frac{I}{A} = -\sigma_i \frac{dV}{dx} = \text{const.} \tag{5.3}$$

within each region. The potential drop across region i is then

$$\Delta V_i = \int_0^{L_i} \left(-\frac{dV}{dx}\right) dx = \frac{J\, L_i}{\sigma_i}. \tag{5.4}$$

Summing the drops gives the applied voltage,

$$V_0 = \sum_{i=1}^{n} \Delta V_i = J \sum_{i=1}^{n} \frac{L_i}{\sigma_i}, \tag{5.5}$$

so the current density is

$$J = \frac{V_0}{\displaystyle\sum_{i=1}^{n} \frac{L_i}{\sigma_i}}. \tag{5.6}$$

Equivalently, defining the series resistance of each region as

$$R_i = \frac{L_i}{\sigma_i A}, \tag{5.7}$$

where R_i is the resistance of region i. one obtains the standard Ohmic form $V_0 = I \sum_i R_i$,

and

$$\Delta V_i = IR_i = V_0 \frac{R_i}{\sum_{j=1}^{n} R_j} = \frac{V_0}{\sum_{j=1}^{n} \frac{L_j}{\sigma_j}} \frac{L_i}{\sigma_i}. \tag{5.8}$$

This expression shows that regions with larger *local resistance* L_i/σ_i take a larger fraction of the total voltage, so the low-conductivity segment dominates the potential drop and creates the desired gradient at the active area.

A practical consideration in this design is the metal–semiconductor interface. Gold deposited on doped silicon forms a Schottky contact; however, for mid-range doping levels the silicon behaves sufficiently conductive for the voltages used here. Under these conditions, thermionic-field emission or tunnelling currents remain small. The total current flowing in the circuit is therefore limited by the silicon conductivity, which allows us to avoid excessive Joule heating while still achieving a controlled potential distribution.

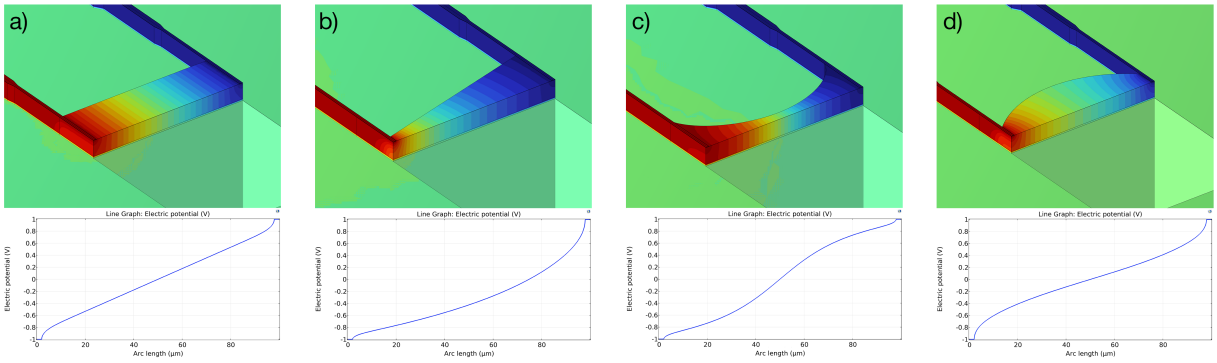


Figure 5.3. Control of a monotonic potential drop by optimizing the shape of the low-conductivity region. (a) Rectangular shape producing an approximately linear voltage drop. (b) Triangular shape. (c,d) Tapered geometries enabling tailored non-linear potential profiles.

An additional advantage of this approach is the possibility of shaping the monotonic potential drop by geometrically tailoring the electrode. By adjusting the width, length, or tapering of the low-conductivity region, the voltage drop can be distributed in a linear or non-linear way. Figure 5.3 illustrates four examples where the electrode geometry results in a linear potential gradient or different curved profiles. These geometries can be optimized using gradient-based methods in finite-element simulations (for example, COMSOL Multiphysics), enabling precise control of the electric field and improving the design accuracy of the phase plate. In the next chapter, we provide a detailed analysis of this design strategy, including its modeling and optimization.

5.4. Sandwich chip design

In MEMS-based electrostatic phase plates, the phase modulation is induced by electrodes biased at fixed voltages. Although the electrode layout is typically designed in a two-dimensional plane, the resulting electrostatic potential extends into three dimensions. Therefore, the axial (z) distribution of the field is significant because it determines the domain where the electrons interact with the potential.

For harmonic phase plates (e.g., multipoles), the axial (z) decay of the potential can often be determined so that the interaction region remains well confined along z. However, electrical connections on the chip—especially the leads from the contact pads to the active area—can distort the intended potential distribution and the resulting phase shift, making it harder to achieve the target phase profile. For non-harmonic phase plates, the situation is more challenging. In addition to on-chip connections, electrodes placed closer to the electron path can further disturb the electrostatic potential, making it harder to produce the target phase shift. This also reduces z-localization and can introduce unwanted coupling to nearby grounded boundaries or neighboring electrodes. Therefore, it is beneficial to localize the electrostatic potential along the optical axis so that most of the phase modulation is accumulated within a well-defined interaction volume.

A clear example is the Boersch phase plate, where the biased electrode is surrounded by grounded electrodes that act as an electrostatic shield (often described as a Faraday-cage-like enclosure). This shielding suppresses field leakage outside the intended region and reduces crosstalk with other electrodes. As a result, electrons experience the designed potential primarily within a small, localized interaction area, which improves the stability and control of the phase modulation.

To address these parasitic phase contributions, I introduced a fabrication and chip design strategy that we call the *sandwich design* . In this approach, a second, grounded chip containing a small aperture (typically on the order of tens of micrometers) is placed above the phase-plate chip, separated by a gap of a few micrometers. The two chips are mechanically aligned using complementary step-like aligners (similar to *LEGO* bricks). After alignment, the chips can be permanently joined using standard silicon wafer/chip bonding methods to form a stable two-chip assembly. The active electrodes are therefore confined and *sandwiched* between two grounded boundaries: the top grounded aperture chip and the grounded silicon substrate of the phase-plate chip. This geometry provides electrostatic shielding and localizes the interaction volume along the optical axis.

Although this design method can be used for any kind of phase plate, particularly non-harmonic phase plates, one example is provided here. Figure 5.4 shows the design of a thin-film phase plate that we fabricated in collaboration with DTU Nanolab. This device was designed to compensate for spherical aberration by imposing a quartic phase profile across the electron beam. The structure consists of several biased Au rings (30 nm thick) patterned on a 30 nm Si_3N_4 membrane. By applying an optimized voltage to each ring, the device generates a three-dimensional electrostatic potential that yields a phase shift approximating

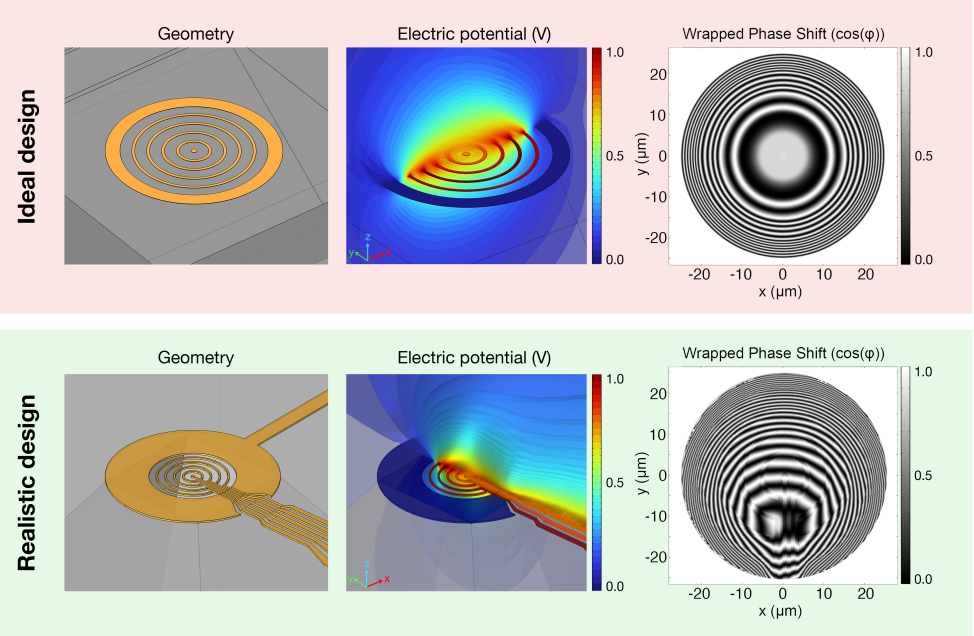


Figure 5.4. Ideal

the target quartic profile. In the realistic design, each ring must be connected to external bias pads through conducting leads. As shown in Figure 5.4, these leads introduce parasitic fields along the y-axis and additional phase contributions that break cylindrical symmetry and cause unwanted phase distortions, thereby reducing the accuracy of the intended phase profile. Figure 5.5 compares the conventional model with the sandwich design. For this phase plate, localizing the electric potential between two grounded electrodes significantly reduces parasitic distortions from the lead electrodes, resulting in a more symmetric phase shift.

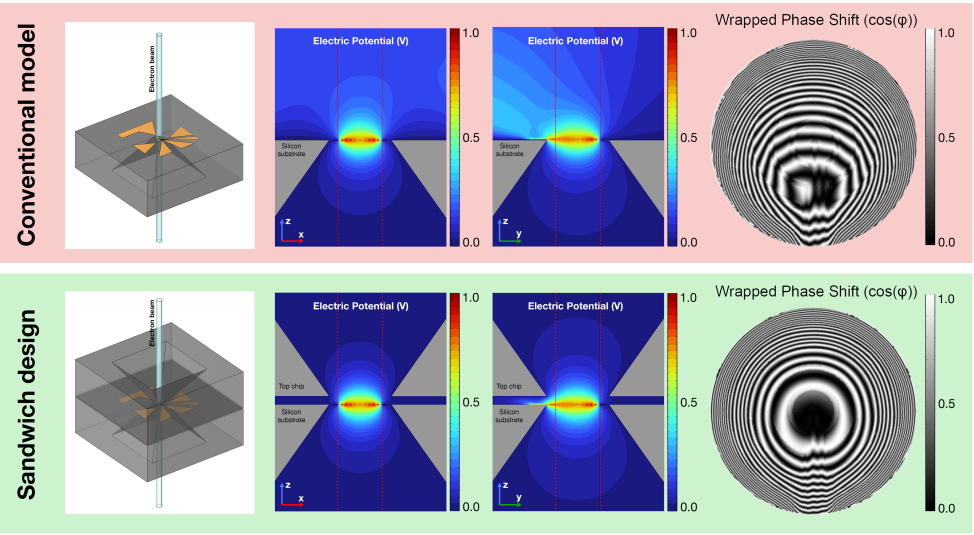


Figure 5.5. Comparison between a conventional single-chip design and the sandwich design. Field localization in the sandwich configuration reduces parasitic distortions from external connections and results in a more symmetric phase profile.

Figure 5.6 shows the fabrication process flow for this design. Although this process is pre-

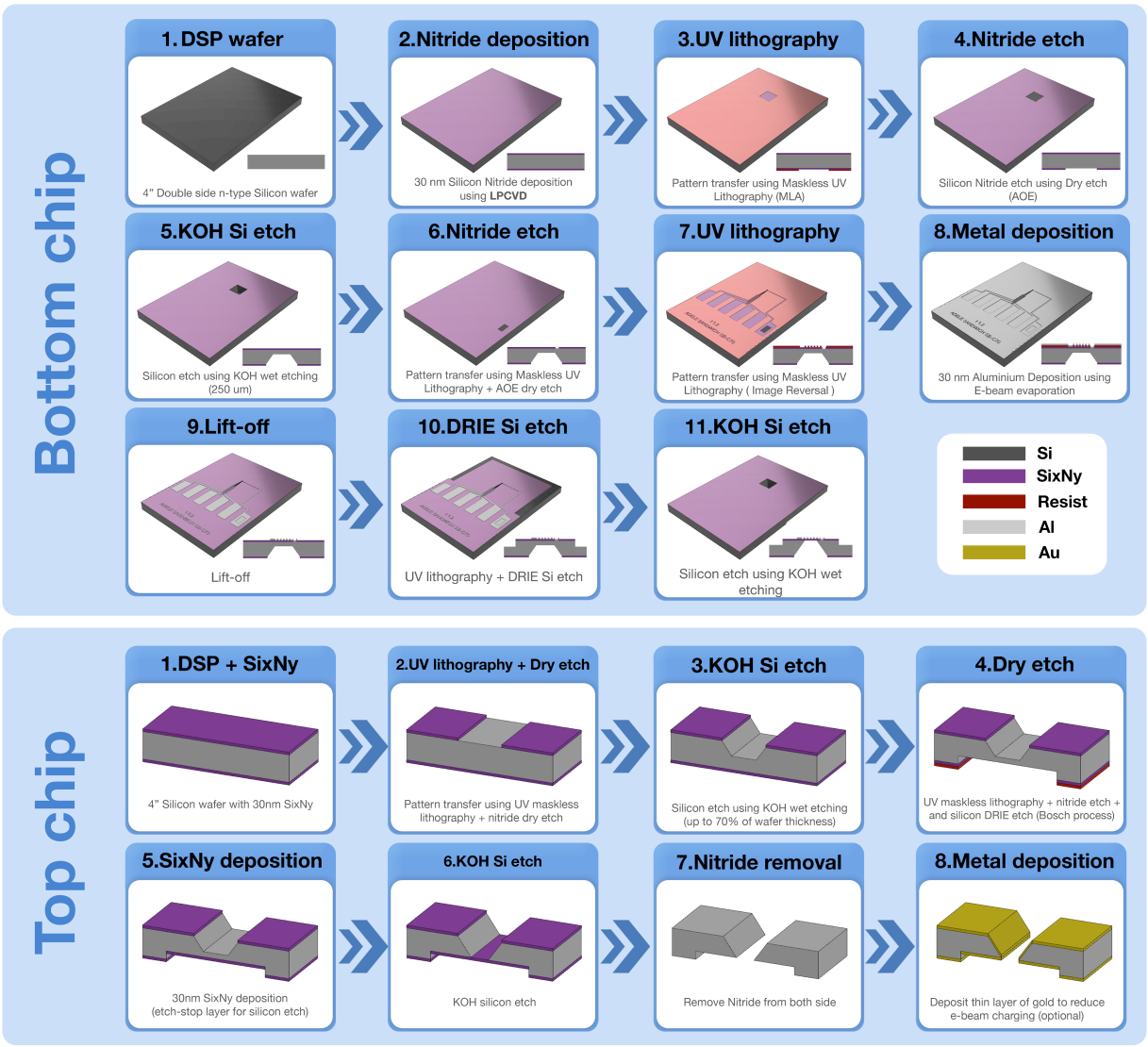


Figure 5.6. Sandwich-design fabrication process flow for the thin-film phase plate. The device is fabricated as two separate chips (top and bottom), which are then aligned and bonded to form the final assembled chip.

sented for a thin-film phase plate chip, the key fabrication concept—splitting the device into a "top" and a "bottom" chip—can also be applied to other phase plate architectures. One important design constraint is the overall MEMS thickness, which must be compatible with both the TEM pole-piece geometry and the electrical-biasing sample holder. Figure 5.7 shows the full-chip layout and the placement of the device in a dedicated TEM electrical-biasing holder (Protochips Fusion AX). Although this holder provides only six independent bias lines, its open-top configuration allows integration of the MEMS device. The main practical requirement is that the top chip must be designed with smaller dimensions so that the holder contacts can reliably reach the on-chip pads. Figure 5.7(b) shows an example of a fabricated device produced using the process flow in Figure 5.6. Optical micrographs of the active area acquired from the top and bottom sides of the chip confirm precise alignment. One advantage of this fabrication strategy, beyond the sandwich concept, is that it can be extended

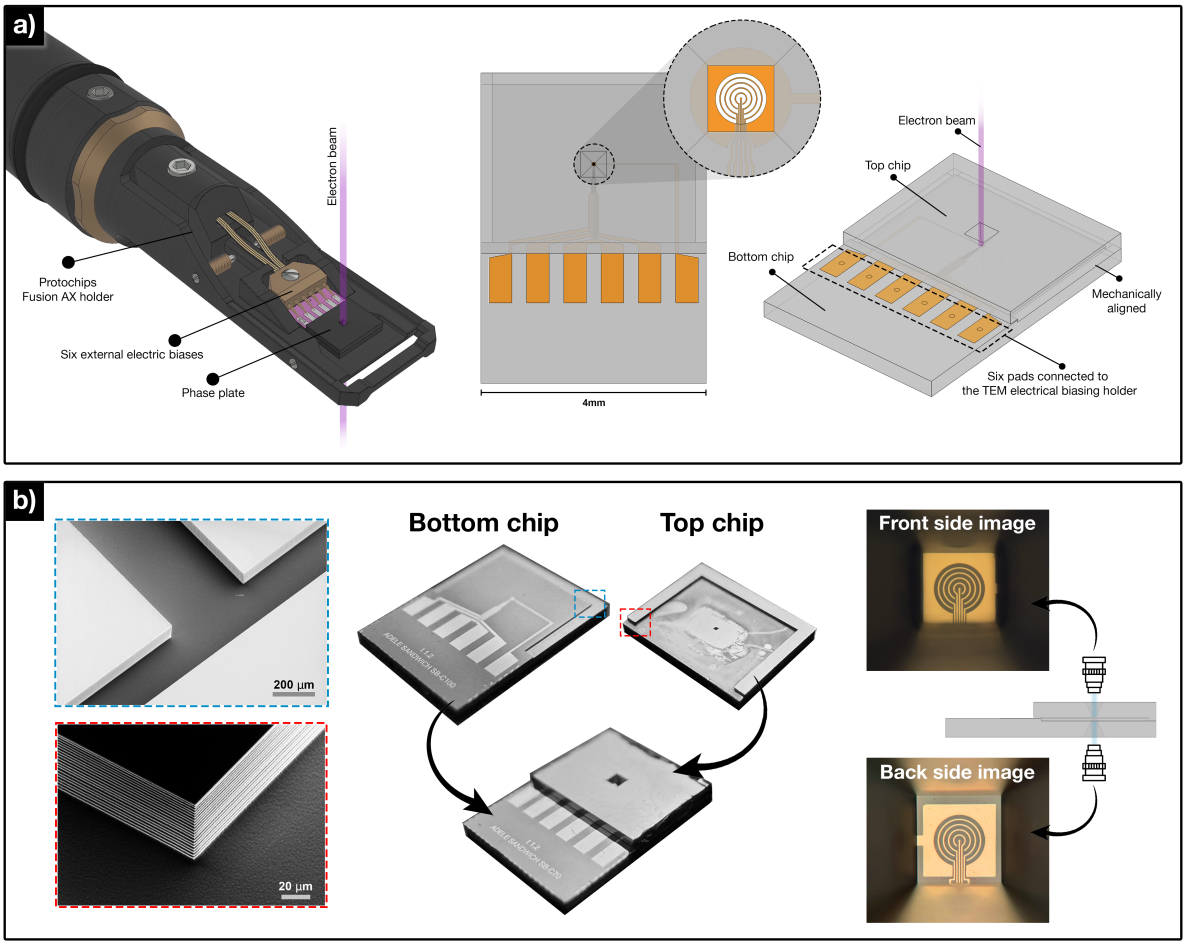


Figure 5.7. Full sandwich-chip layout. (a) Protochips electrical-biasing holder with an open-top design (left) and the sandwich-chip design with mechanical alignment (right). (b) Fabricated devices: (left) SEM image of the mechanical alignment steps (dashed square in the middle image), (middle) final chip formed by stacking the two chips, and (right) optical micrographs of the top- and bottom-side apertures showing accurate alignment achieved by the mechanical aligners.

to *stack* two MEMS devices in series along the optical axis. Such vertical integration would enable more complex phase-plate architectures (e.g., multi-stage or multi-electrode designs) while keeping each MEMS layer manufacturable and independently optimized. This stacking approach is therefore a promising route for future device designs.

At the time of writing, electrical biasing experiments could not be performed because the specific holder configuration required for these measurements was not available; therefore, this section focused only on the fabrication strategy.

6 Numerical Simulation and Modeling

Contents

6.1 COMSOL Simulations 100

6.1.1 Choosing the Physics Modules and Study Type 100

6.1.2 Geometry Import and Model Setup 101

6.1.3 Setting Boundary Conditions 102

6.1.4 Meshing and Computing 103

6.1.5 Optimization and Inverse Design 104

6.2 Phase-Shift Calculation: Projected Potential and Multislice Propagation 106

Paraxial Wave Equation and Projected Potential 106

Multislice Recurrence for a Thick Phase Plate 107

Thin-Phase (Single Projection) Versus Multislice method 108

6.3 Python-Based Simulation, Control, and Data-Analysis Tools . . 110

6.3.1 Multislice Simulation GUI 110

6.3.2 Electron Microscopy Toolkit GUI 113

6.4 Machine Learning and Neural Network 118

6.1. COMSOL Simulations

COMSOL Multiphysics® is a finite-element simulation environment used to solve coupled partial differential equations across many areas of physics. It provides a unified platform for modeling systems that involve electromagnetics, structural mechanics, heat transfer, fluid dynamics, acoustics, and chemical transport, either alone or in fully coupled multiphysics setups. The software includes flexible tools for geometry creation, meshing, parametric sweeps, and numerical optimization. These features allow accurate analysis of physical behavior and make it possible to introduce custom physics when required. Because of its modular structure and the ability to combine several physics interfaces in a single model, COMSOL is widely used in engineering, materials science, and device-level research.
In this thesis, COMSOL Multiphysics was used to model and optimize the electromagnetic, thermal, and mechanical response of micro- and nanoscale devices relevant to electrostatic phase plates. The AC/DC module was used for electrostatic, electric-current, and magnetostatic simulations. The MEMS and structural-mechanics interfaces supported the design of microfabricated components and allowed us to evaluate stress, deformation, and mechanical stability. Thermal effects, including Joule heating, were studied with the heat-transfer module. In addition, the electromagnetic-wave, wave-optics, and ray-optics interfaces were used to analyze wave propagation in some simulations. Finally, COMSOL's gradient-based optimization tools were employed for parameter tuning, inverse design, and shape optimization, providing a systematic approach to improving device performance across multiple coupled physics domains.
The next section outlines the full simulation workflow, including model construction, material assignment, study selection, validation steps, and the refinement process used throughout this thesis.

6.1.1. Choosing the Physics Modules and Study Type

Once the model dimensionality has been chosen (3D, 2D, or 2D axisymmetric), the next step is to select the appropriate physics interfaces for the problem. In this thesis, the following COMSOL modules were used:

1. **AC/DC – Electrostatics (es)**: Solves for the electric potential and electric field under static conditions. It was mainly used to model electrostatic phase plates, where the potential distribution defines the phase shift.

2. **AC/DC – Electric Currents (ec)**: Computes current density, resistive behaviour, and voltage distribution in conductive materials. In this work, it was used to model current-driven electrostatic phase plates.

3. **Joule Heating (Electric Currents + Heat Transfer in Solids)**: Solves the coupled problem in which electrical currents generate heat. It was used to calculate the temperature rise and steady-state thermal profiles in MEMS structures under applied voltage.

4. **Charged Particle Tracing (CPT)**: Computes electron trajectories in prescribed electric and magnetic fields. In this work, it was used to analyse electron motion through phase plates by coupling the trajectories to the electric and magnetic field distributions obtained from the AC/DC modules.

5. **Electromechanics**: Couples electrostatics with structural mechanics to model deformation of MEMS elements under electrostatic forces. It was used to assess shape stability, pull-in behaviour, and endurance at high voltage biases.

6. **Optics – Wave Optics and Ray Optics**: Simulates electromagnetic wave propagation in dielectric structures. In this work, it was used to draw analogies between light optics and electron optics.

After selecting the physics interfaces, an appropriate *study type* is chosen:

- **Stationary (static) study** – Used for electrostatic, electric-current, and steady-state thermal simulations, where only the equilibrium field distribution is required and time dependence can be neglected.

- **Time-dependent study** – Used for charged-particle trajectory simulations and for problems with time-varying fields (e.g., AC excitation or pulsed operation). This study type allows analysis of transient behavior and the dynamic response of the system.

All of the above interfaces were employed in the investigation of phase plates. However, the Electrostatics and Electric Currents modules played a central role in this work. Therefore, the following sections focus on these two physics models.

6.1.2. Geometry Import and Model Setup

The simulation geometry in COMSOL can be created in two ways: either by building it directly with the built-in CAD tools or by importing an external 2D or 3D CAD model. For simple structures, the internal tools are usually sufficient. For more complex or highly detailed devices, it is more efficient to import CAD files generated in software such as Rhino 3D or SolidWorks. In all cases, the geometry should be kept as simple as possible. Very large or overly detailed structures—such as mixing millimeter-scale parts with micro- or nanoscale features in the same model—make the meshing process difficult and increase the computational cost, mostly without improving physical accuracy. For MEMS-based phase plates, the geometry should include only the regions that influence the physics of interest, especially the features that determine the boundary conditions and the resulting electrostatic

or thermal fields.

Once the device geometry is built or imported, it must be placed inside a surrounding domain representing the microscope environment, which, in this work, is vacuum. Because computational resources are limited, it is not practical to model a very large vacuum region around the device. Instead, a finite vacuum box is used, and the outer boundaries are treated to mimic an open space extending far from the structure. In COMSOL, this is achieved with *infinite element domain* subdomains or equivalent far-field boundary conditions, which approximate an effectively unbounded vacuum. These boundary treatments allow the fields to decay smoothly and prevent artificial reflections at the outer boundaries of the domain. This approach provides an accurate approximation of the real microscope environment while keeping the simulation size and computation time manageable.

In simulations of the electric potential or magnetic field around the active area of the MEMS phase plate—that is, the region traversed by the electron beam—it is often useful to define an explicit vacuum subdomain along the electron path through the device. In practice, this is modeled as a cylindrical region aligned with the electron beam. This design has two main advantages. First, it allows the use of a finer mesh in the beam region, which increases the accuracy of the calculated potential or field seen by the electrons. Second, it provides a well-defined domain for computing the projected potential by using COMSOL's built-in *general projection* operator to integrate the electric potential along the beam direction. This projected potential is later used to calculate the corresponding electron phase shift (See also Section 6.2).

The next step is to define and assign materials to the different subdomains of the model. COMSOL provides an extensive built-in materials library (metals, semiconductors, dielectrics, structural materials, etc.), which can be used as a starting point. However, whenever possible it is preferable to use measured or device-specific material properties instead of generic values, especially for parameters that strongly influence the solution.

For electrostatic and electric-current simulations, the most critical parameters are the electrical conductivity and the relative permittivity of the materials. These quantities determine the potential distribution, charge accumulation, and current paths in the device. In this thesis, material parameters were taken, when available, from data sheets of the specific silicon wafers and metals used in fabrication, or from experimentally measured values, rather than relying only on default library entries. This reduces systematic errors and makes the simulations better reflect the actual behaviour of the fabricated MEMS phase plates and supporting structures.

6.1.3. Setting Boundary Conditions

The next critical step in phase-plate simulations is defining the boundary conditions associated with each physics interface. For the most important physics—electrostatics and electric currents—the essential boundary conditions are as follows:

1. **Electric Potential Boundary Condition**

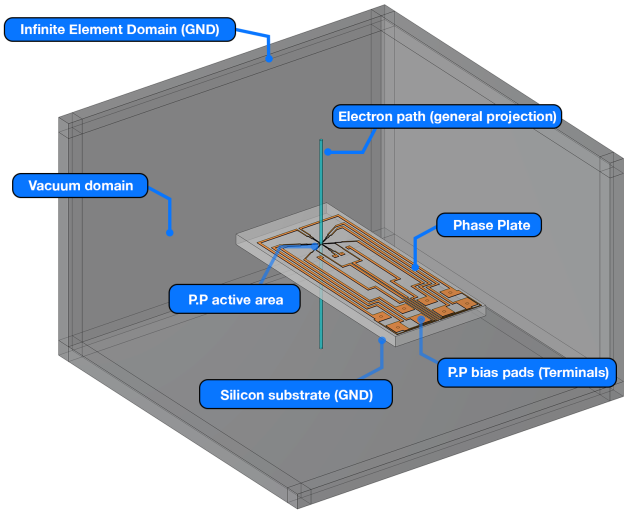


Figure 6.1. COMSOL simulation domain for the electrostatic phase plate: the device is placed in a vacuum region enclosed by a grounded infinite-element domain; electric potentials are applied via Terminal boundary conditions, the silicon substrate is grounded, and an electron-path subdomain is used to compute the phase shift using the General Projection function

This boundary condition is applied to the parts of the device that are connected to an external voltage source, such as the biased pads in the TEM holder. In COMSOL, this can be implemented using either the *Electric Potential* or the *Terminal* boundary condition. Both impose a fixed potential, but the *Terminal* condition also allows COMSOL to compute quantities such as the total current, resistance, and power dissipation between the terminal and ground, which is often useful for analysing MEMS phase plates.

2. **Ground Boundary Condition**
   This condition defines the parts of the device that are electrically grounded. In MEMS-based phase plates, two regions are typically grounded: (i) the contact surface between the MEMS chip and the TEM holder, which is electrically connected to the microscope ground; and (ii) the outer surfaces of the vacuum domain (or the outer boundary of an infinite-element domain), which represent the grounded microscope chamber. Applying these boundaries correctly ensures that the potential distribution around the active region of the phase plate matches the real microscope environment.

Figure 6.1 shows how these two boundary conditions are applied to a representative phase-plate geometry.

6.1.4. Meshing and Computing

Once all parameters and boundary conditions are defined, the next step is to generate a suitable mesh for the geometry. COMSOL can create a *physics-controlled* mesh that is automatically adapted to the selected physics interfaces. The user can then choose the global

element size, from extremely coarse to extremely fine. For many geometries, this automatic meshing is sufficient; however, for accurate phase-plate simulations, it is often preferable to use a *user-controlled* mesh.

In a user-defined mesh, the element size can be refined specifically in the regions near the electron path, such as narrow electrodes, membrane edges, and the vacuum cylinder that represents the electron beam path. A finer mesh in these subdomains improves the accuracy of the calculated electric potential and field, and therefore, of the projected potential and phase shift. At the same time, less critical regions of the vacuum or support structure can be meshed more coarsely to minimize the total number of mesh elements and computation time.

After meshing, the model is solved using the chosen study type (stationary or time-dependent). The results can then be visualized in several ways: as 3D volume plots, slices through the geometry, or fields mapped on selected boundaries. In addition, COMSOL allows exporting the solutions to 2D or 3D datasets, which can be saved as numerical matrices and further processed in external tools such as Python for post-processing and data analysis.

6.1.5. Optimization and Inverse Design

COMSOL Multiphysics provides a comprehensive framework for optimization-driven design, allowing users to refine device performance by systematically adjusting model parameters. The *parameter optimization* tools focus on tuning scalar quantities, such as voltages, geometric parameters, or material properties, to minimize or maximize a chosen objective function. COMSOL integrates sensitivity analysis directly into the solver, ensuring that the derivatives of the objective with respect to parameters are computed consistently within the multiphysics environment. Beyond parameter tuning, COMSOL supports *shape optimization* and *topology optimization* for problems in which the geometry itself must evolve. *Shape optimization* modifies boundary positions using control functions defined on contours or surfaces, making it suitable for refining electrode edges or aperture shapes. *Topology optimization*, in contrast, operates on a material distribution inside a domain, allowing the solver to "discover" new structural layouts without predefined shapes. This method is powerful for exploring non-intuitive solutions in electrostatics, mechanics, or photonics, where the optimal configuration may involve complex internal voids or material rearrangements.

Throughout this thesis, COMSOL optimization tools are used to design and refine several electrostatic phase plates. The specific optimization setup for each device is described in the corresponding chapters. In all cases, the central idea is to match the simulated phase shift to a desired analytical target, which is our *objective function*.

To evaluate the phase response, we define a set of vertical sampling lines parallel to the optical axis (z) at selected transverse positions (x_0, y_0) in the region where the electron beam propagates (see Figure 6.2). Along each line, the projected electrostatic potential is computed using a one-dimensional integration,

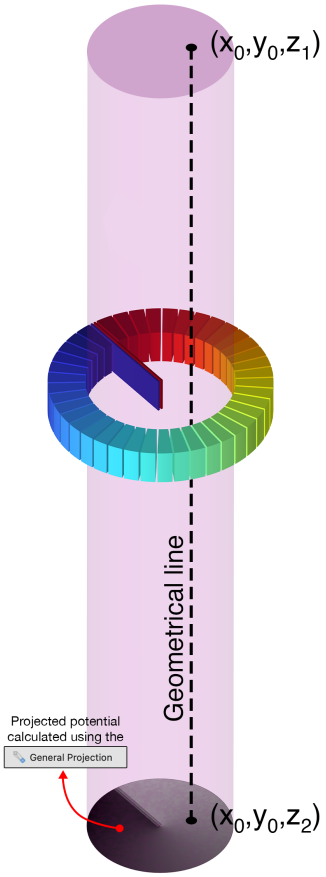


Figure 6.2. Vertical evaluation geometry used in COMSOL to extract the electron phase shift. The purple cylinder defines the projection volume used with COMSOL's General Projection operator to compute the phase map $\varphi(x, y)$ on a plane by integrating along z. The dashed line indicates the integration path at a fixed lateral position (x_0, y_0), yielding $\varphi(x_0, y_0)$. The same approach can be applied to tilted illumination by orienting the projection direction.

$$\varphi(x_0, y_0) = C_E \int_{z_1}^{z_2} V(x_0, y_0, z)\, \mathrm{d}z, \tag{6.1}$$

which is proportional to the phase shift at that transverse coordinate. The objective function is then constructed as a least-squares error between the simulated phase profile and the analytical target phase $\varphi_{\text{target}}(x_0, y_0)$ evaluated at the same sampling positions:

$$\mathcal{O} = \sum_{(x_0,y_0)\in\Omega} \left|\varphi_{\text{target}}(x_0, y_0) - \varphi(x_0, y_0)\right|^2, \tag{6.2}$$

where Ω denotes the discrete set of sampling lines. The optimization minimizes $\mathcal{O}$ by adjusting the device control parameters (e.g., electrode voltages) or the device geometry, thereby driving the simulated phase response towards the desired analytical phase profile. The number and position of these lines depend on the complexity of the optimization and are adjusted for each specific design. In the next chapter, we discuss the optimization of different phase plates.

6.2. Phase-Shift Calculation: Projected Potential and Multislice Propagation

In this work, the phase shift introduced by a three-dimensional electrostatic (or magnetostatic) potential is calculated using two related approaches:

1. **Single projected potential**, obtained with COMSOL's *General Projection*
2. **The Multislice Solution** for thick devices where free-space propagation inside the structure cannot be neglected.

Although the standard derivation is usually presented for specimens (see, e.g., Kirkland, *Advanced Computing in Electron Microscopy* [50]), the same mathematics applies directly to electrostatic phase plates, where the "specimen potential" is replaced by the designed device potential.

Paraxial Wave Equation and Projected Potential

For fast electrons, the elastic interaction with a slowly varying potential $V(x, y, z)$ can be written in paraxial form as a first-order equation along the optical axis z [50]:

$$\frac{\partial \psi(x, y, z)}{\partial z} = \left[i \, \frac{\lambda}{4\pi} \nabla_{xy}^2 + i \, C_E \, V(x, y, z) \right] \psi(x, y, z), \tag{6.3}$$

where $\psi(x, y, z)$ is the electron wavefunction, λ is the relativistic wavelength, ∇_{xy}^2 is the transverse Laplacian, and C_E is the interaction constant. Equation 6.3 contains two non-commuting operators A and B. A free-space propagation operator,

$$A = i \, \frac{\lambda}{4\pi} \nabla_{xy}^2, \tag{6.4}$$

and an interaction operator,

$$B = i \, C_E \, V(x, y, z). \tag{6.5}$$

Because A and B do not commute, the exponential $\exp[(A + B)\Delta z]$ cannot be factorized exactly. To first order in Δz, one can approximate

$$\exp[(A + B)\Delta z] \approx \exp(A\Delta z) \, \exp(B\Delta z) + \mathcal{O}(\Delta z^2), \tag{6.6}$$

A detailed discussion of the numerical treatment of this equation and its implementation is given in Ref. [161]. For a thin slice of thickness Δz, the formal solution of Eq. (6.3) can be written as

$$\psi(x, y, z + \Delta z) = \exp\left[i \, \frac{\lambda}{4\pi} \nabla_{xy}^2 \Delta z + i \, C_E \, v_{\Delta z}(x, y, z) \right] \psi(x, y, z) \tag{6.7}$$

It is convenient to introduce the projected potential over the slice,

$$v_{\Delta z}(x,y,z) = \int_{z}^{z+\Delta z} V(x,y,z')\,\mathrm{d}z', \tag{6.8}$$

which will determine the local phase shift within that layer.
Equation 6.7 is then separated into a *transmission function* for the slice.

$$t_{\Delta z}(x,y) = \exp[i\,C_E\,v_{\Delta z}(x,y)]\,, \tag{6.9}$$

which describes the phase shift due to the projected potential of that layer; and A *propagation operator* describing free-space propagation over Δz.
In Fourier space ($\mathcal{F}$ denoting the 2D Fourier transform over x and y), the propagation operator becomes

$$\mathcal{F}\left\{\exp\left(i\,\frac{\lambda}{4\pi}\Delta z\,\nabla^2_{xy}\right)t_{\Delta z}(x,y)\right\} = \exp\left(-i\,\pi\lambda\,k^2\Delta z\right)\,\mathcal{F}\{t_{\Delta z}(x,y)\}, \tag{6.10}$$

with $k=(k_x,k_y)$ and $k^2=k_x^2+k_y^2$. The corresponding propagator in reciprocal space is

$$P(k,\Delta z) = \exp\left(-i\,\pi\lambda\,k^2\Delta z\right). \tag{6.11}$$

Transforming back to real space, one obtains the Fresnel propagator

$$p(x,y,\Delta z) = \mathcal{F}^{-1}\{P(k,\Delta z)\} = \frac{1}{i\lambda\Delta z}\exp\left[i\,\frac{\pi}{\lambda\Delta z}(x^2+y^2)\right], \tag{6.12}$$

so that the wave after one slice is

$$\psi(x,y,z+\Delta z) \approx p(x,y,\Delta z) * [t_{\Delta z}(x,y)\,\psi(x,y,z)]\,, \tag{6.13}$$

where $*$ denotes 2D convolution.

Multislice Recurrence for a Thick Phase Plate

For a thick potential region (specimen or phase plate), the device volume is divided into N thin slices with thickness Δz_n. For each slice n:

$$v_n(x,y) = \int_{z_n}^{z_{n+1}} V(x,y,z)\,\mathrm{d}z, \tag{6.14}$$

$$t_n(x,y) = \exp[i\,C_E\,v_n(x,y)]\,, \tag{6.15}$$

$$P_n(k) = \exp\left[-i\,\pi\lambda\,k^2\,\Delta z_n\right]. \tag{6.16}$$

Denoting the wave function at the entrance of slice n as $\psi_n(x,y)$, the multislice recurrence relation reads

$$\psi_{n+1}(x,y) = \mathcal{F}^{-1}\left\{P_n(k)\,\mathcal{F}\left[t_n(x,y)\,\psi_n(x,y)\right]\right\}, \tag{6.17}$$

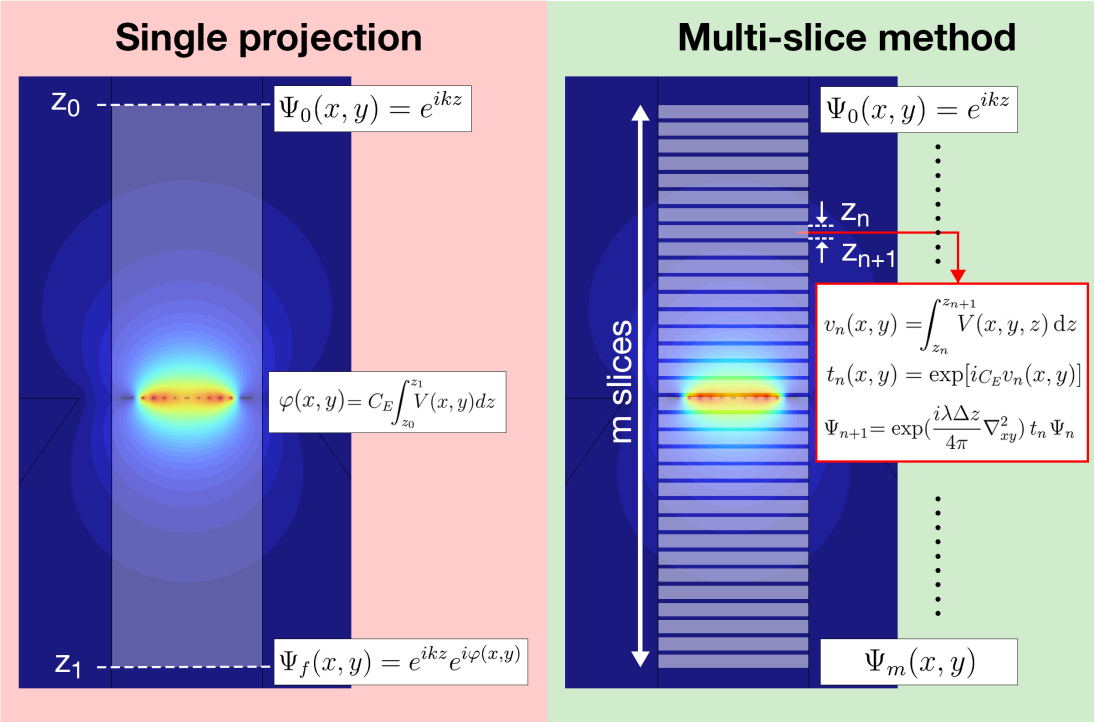


Figure 6.3. Single projection vs Multislice.

which is the standard "transmit–propagate" sequence. Algorithmically, for a phase plate:

1. compute the 3D potential $V(x, y, z)$ from COMSOL;
2. slice the device along z into thin layers and form $v_n(x, y)$ for each slice;
3. construct the transmission functions $t_n(x, y)$;
4. start from the incident wave $\psi_0(x, y)$ and apply Eq. (6.17) successively until the exit surface is reached.

The final wavefunction $\psi_m(x, y)$ represents the electron wavefunction after passing through the entire phase plate, including both local phase modulation and free-space propagation within the structure.

Thin-Phase (Single Projection) Versus Multislice method

If the phase plate is sufficiently thin and the potential varies slowly along z, the free-space propagation inside the device can be neglected to a good approximation. In this *phase-grating* limit, the full 3D potential is replaced by a single projected potential

$$V_{\mathrm{proj}}(x, y) = \int V(x, y, z)\,\mathrm{d}z, \tag{6.18}$$

and the exit wave is approximated by

$$\psi_{\mathrm{exit}}(x, y) \approx \exp[i\,\sigma\,V_{\mathrm{proj}}(x, y)]\,\psi_{\mathrm{in}}(x, y), \tag{6.19}$$

optionally followed by a single free-space propagation step to the next plane in the microscope. To calculate the phase shift using single projection, there are two options: (i) export the simulated potential and use a Python script to compute the integral, or (ii) more conveniently, use COMSOL's *General Projection* operator to evaluate the integral directly (see Figure 6.2). For thicker devices, or for phase plates with strong longitudinal variations of the potential, the single-projection approximation is not sufficient. In that case, the multislice scheme in Eq. (6.17) must be used, where the potential is sliced, the phase shift in each slice is applied, and the wave is propagated between slices. In this thesis, both approaches are used: thin electrostatic phase plates are treated by a single projected phase, while thick or strongly structured devices are simulated using a full multislice treatment based on COMSOL-generated three-dimensional potentials.

6.3. Python-Based Simulation, Control, and Data-Analysis Tools

Python was used extensively throughout this thesis as a core computational framework for simulation, analysis, and visualization. Its flexibility and rich ecosystem of scientific libraries made it possible to develop a unified workflow that integrates multiple aspects of electron-optical modelling and data processing. In this work, Python supported five main tasks:

1. Numerical simulations, including wave-optics propagation, ray tracing, and aberration modelling.
2. Data analysis and interpretation of experimental and simulated signals.
3. Image processing, particularly for TEM and STEM datasets.
4. Phase-shift calculations, using both projected-potential and multislice implementations.
5. Machine learning and deep neural networks (DNN), used for inverse design, parameter prediction, and automated analysis.

To streamline these tools and make them accessible to non-experts, I developed an interactive Python-based graphical user interface (GUI). This interface links the simulation modules, analysis routines, and visualization tools into a single environment that allows users to adjust parameters, inspect results, and run computational workflows in real time.
The following sections introduce the main Python modules, functions, and design choices used in this thesis, along with their roles in supporting phase-plate modeling, optimization, and experimental data analysis.

6.3.1. Multislice Simulation GUI

The multislice simulations used in this thesis were implemented in Python and integrated into a graphical user interface (GUI). The full application comprises approximately 2400 lines of Python code. Only the key numerical components are summarised here to illustrate the computational workflow; the remaining parts of the implementation consist of utility functions for data loading, plotting, user interaction, and GUI control, which are not directly relevant to the underlying physical model.
A screenshot of the graphical user interface is shown in Figure 6.4. The complete implementation, including all simulation modules and GUI tools, is available in the public GitHub repository listed in the appendix A.

Fresnel Propagator

Free-space propagation between adjacent slices is carried out using the Fresnel propagator. In Python, this is implemented by constructing the transfer function in reciprocal space,

$$P(k_x, k_y; \Delta z) = \exp\left[-i\pi\lambda(k_x^2 + k_y^2)\Delta z\right], \tag{6.20}$$

and applying it to the Fourier transform of the wave function. The corresponding pseudocode is:

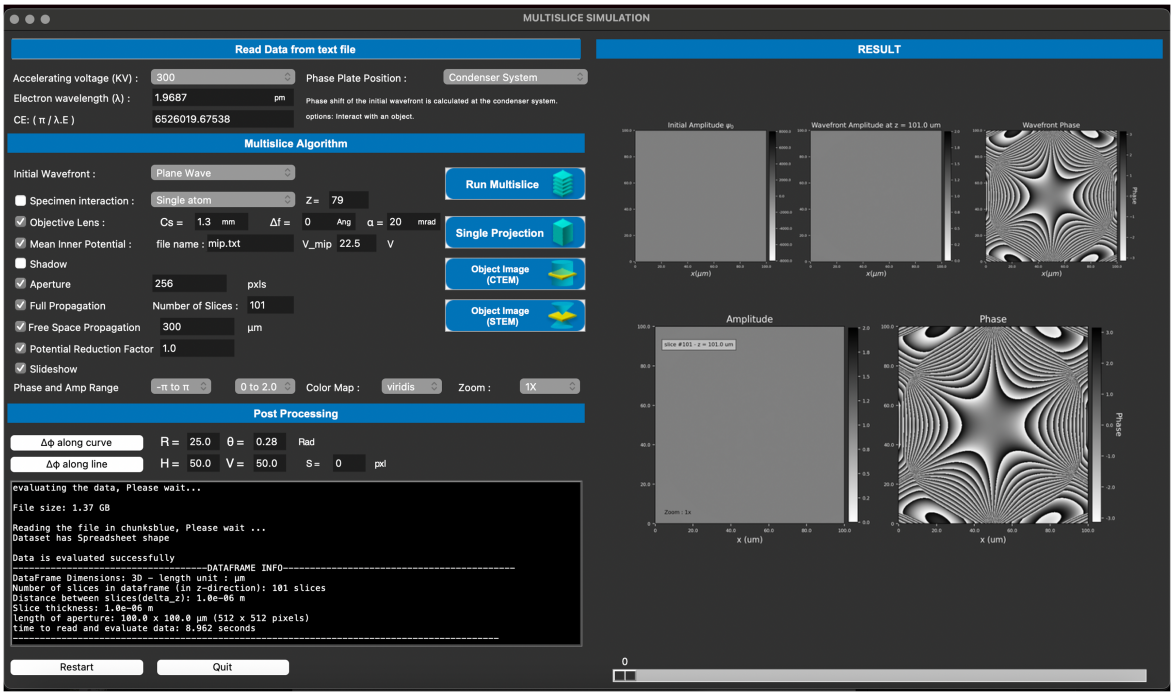


Figure 6.4. Graphical user interface (GUI) developed in Python for interactive control of the multislice simulation, wave–optics propagation, and phase-plate analysis. The interface provides real-time parameter adjustment and visualisation of wavefronts, potentials, and simulation outputs.

```
def fresnel_propagator(wave, kx, ky, dz, lam):
    P = np.exp(-1j * np.pi * lam * (kx**2 + ky**2) * dz)
    return np.fft.ifft2(P * np.fft.fft2(wave))
```

This function performs one free-space propagation step of thickness Δz.

Transmission Function

The phase shift generated by a single slice is computed from the projected potential,

$$v_n(x, y) = V(x, y, z_n)\,\Delta z,$$

resulting in the transmission function

$$t_n(x, y) = \exp[i\sigma v_n(x, y)].$$

The corresponding Python implementation extracts the potential for the current slice and constructs t_n:

```
def transmission(potential_slice, dz, sigma):
    v = potential_slice * dz
    return exp(1j * sigma * v)
```

Multislice Recurrence

The multislice update applies the transmission function followed by the Fresnel propagator:

$$\psi_{n+1}(x,y) = \mathcal{F}^{-1}\left\{P_n(k_x,k_y)\,\mathcal{F}[t_n(x,y)\,\psi_n(x,y)]\right\}.$$

The Python pseudocode for the full multislice loop is:

```
def multislice(potential, dz_list, psi0, sigma, kx, ky, lam):
    psi = psi0
    for n, dz in enumerate(dz_list):
        # Extract 2D potential of slice n
        Vn = potential[n]

        # Transmission
        t = exp(1j * sigma * Vn * dz)
        psi = t * psi

        # Fresnel propagation
        psi = fresnel_propagator(psi, kx, ky, dz, lam)

    return psi
```

Here, `psi0` is the incident electron wavefront, and `potential[n]` contains the (x, y) potential map for slice n.

Implementation Notes

The Python implementation used in this work builds on the core multislice routines and adds a number of practical tools for data handling and visualisation. These include loading slice-by-slice potentials exported from COMSOL, reshaping the data into (pxl × pxl) arrays, applying zero-padding or circular apertures, and providing a graphical interface for real-time parameter control. In addition, the GUI offers several features that assist both simulation and analysis:

- Setting the acceleration voltage to study the effect of the electron wavelength on the phase shift,
- Defining the incident beam as a plane wave, a Gaussian wave, or an imported wavefront,
- Adding the mean inner potential along the electron path,
- Loading a phase object or crystal lattice to model beam–specimen interaction,
- Inserting lens functions to visualise the effect of an objective lens,
- Propagating the wavefunction in free space to examine the beam evolution at different distances,

- Placing apertures at user-controlled positions,
- Loading large COMSOL datasets in batches to reduce memory usage,
- Analysing the wavefunction through one-dimensional line profiles or radial and angular intensity curves.

These additions improve usability and allow rapid testing of different beam parameters, but they do not alter the underlying physics of the multislice method. For this reason, the technical implementation details are not included in the main text.

6.3.2. Electron Microscopy Toolkit GUI

As part of this thesis, I developed a Python-based GUI that integrates several simulation, analysis, and control tools for electron microscopy. The application consists of approximately 9000 lines of Python code and combines wave-optics simulation, ray-optics modelling, image processing, Deep Neural Network and hardware control in a single environment. This toolkit represents one of the main technical developments of the thesis and was essential for both the design of electrostatic phase plates and the interpretation of experimental TEM data. A screenshot of the GUI is shown in Figure 6.5.
The interface is organised into multiple tabs, each dedicated to a specific task: (i) probe and diffraction-pattern simulation under aberration, (ii) Ronchigram simulation, (iii) image processing and neural-network-based aberration analysis, (iv) microscope ray-optics simulation, and (v) controller tools for the external voltage-control board. Only the key numerical concepts are summarised here only for the tabs that are related to the thesis; implementation details and full source code are provided in the public GitHub repository (see appendix A).

Probe and Diffraction-Pattern Simulation

One of the central components of the GUI is the probe and diffraction-pattern simulation tab, which allows interactive control of lens aberrations and real-time visualisation of their effect on the electron probe. Aberration coefficients such as defocus, two-fold astigmatism (A_2), three-fold astigmatism (A_3), spherical aberration (C_3), and coma can be adjusted continuously using slider widgets. Thanks to the low computational overhead of the PyQt5 interface and the use of fast Fourier transforms (FFT), the probe is updated within a few milliseconds after each slider movement. This real-time response was one of the key motivations for developing the tool and proved valuable for exploring aberration effects during multipoles phase-plate design.
The numerical model is based on the standard Fourier-optical description of the TEM illumination system: an aperture function is defined in real space, multiplied by the phase of an ideal lens, and then Fourier-transformed into the back focal plane. The user-defined aberrations are implemented through an aberration phase function (see Equation 2.26). The aberration phase is evaluated on the same reciprocal grid used for the FFT, ensuring that changes in the sliders correspond directly to physical modifications of the wavefront.

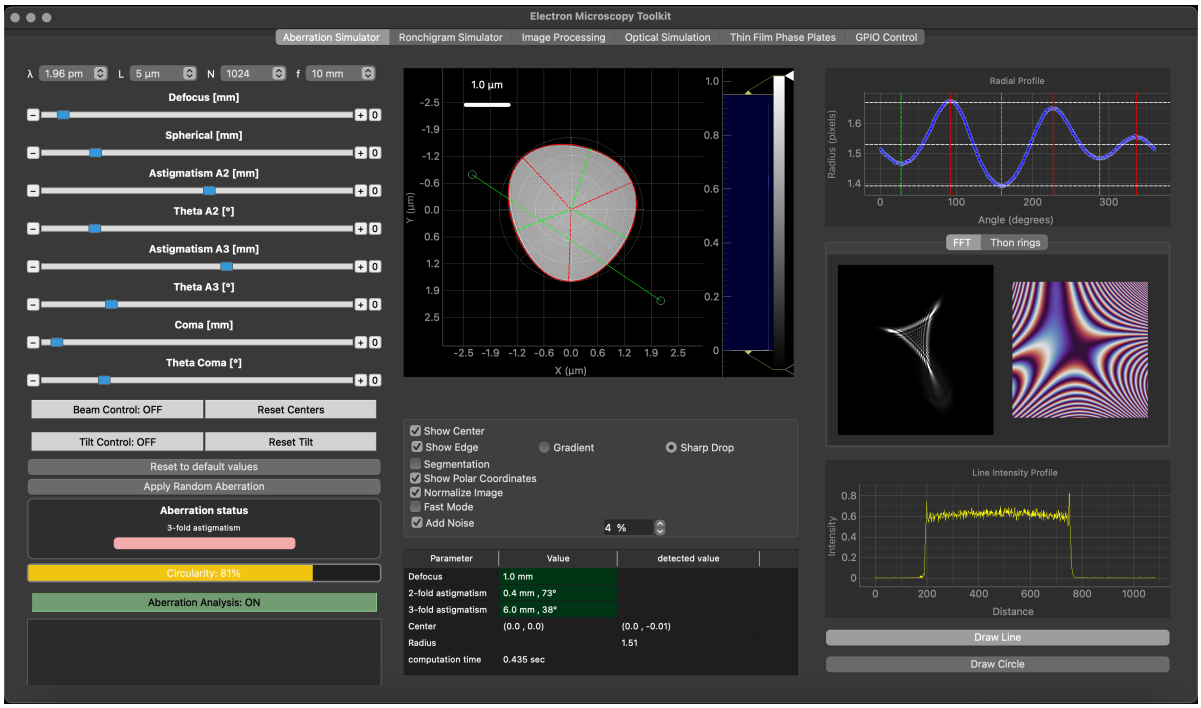


Figure 6.5. Graphical user interface (GUI) developed in Python for interactive Simulation

In Python, the aberration phase function is implemented in a compact numerical routine such as:

```
def aberration_phase(defocus, C3, A2, th_A2, A3, th_A3, coma, th_coma,
                     lam, rho, phi):
    chi_def = -np.pi * lam * defocus * rho**2
    chi_C3  = 0.5 * np.pi * C3 * lam**3 * rho**4
    chi_A2  = np.pi * A2 * lam * rho**2 * np.sin(2*(phi - th_A2))
    chi_A3  = (np.pi/3) * A3 * lam**2 * rho**3 * np.cos(3*(phi - th_A3))
    chi_B   = (2*np.pi/3) * coma * lam**2 * rho**3 * np.sin(phi - th_coma)
    return chi_def + chi_C3 + chi_A2 + chi_A3 + chi_B
```

Once the aberration phase $\chi(k_x, k_y)$ is calculated on the reciprocal-space grid, it is applied directly to the Fourier transform of the initial wavefront,

$$\Psi_{\mathrm{ab}}(k_x, k_y) = \Psi_0(k_x, k_y)\, \exp[i\chi(k_x, k_y)]\,.$$

The inverse Fourier transform of Ψ_{ab} then yields the aberrated probe in real space, from which both the probe intensity $|\psi(x, y)|^2$ and phase $\arg\{\psi(x, y)\}$ are obtained. In addition, the diffraction pattern can be computed by Fourier transforming the pupil function $\exp[i\chi]$, allowing the user to visualise how individual aberration terms modify the contrast and symmetry of the diffraction plane.

The numerical routine `simulate_aberrated_probe` implements this full sequence: construction of the aperture function, application of lens and tilt phases, generation of the aberration phase in reciprocal space, and computation of the real-space probe and the diffraction pattern

through forward and inverse FFTs. This tab became an essential exploratory tool throughout the thesis. It was used to compare simulated aberration effects with experimental beam behaviour, to validate the wavefront modifications introduced by electrostatic phase plates, and to develop an intuitive understanding of how individual aberrations shape the probe in both TEM and STEM illumination modes.

```python
import numpy as np

def simulate_aberrated_probe(N, L, lam,
                             defocus=0.0, C3=0.0,
                             A2=0.0, th_A2=0.0,
                             A3=0.0, th_A3=0.0,
                             coma=0.0, th_coma=0.0,
                             tilt_x=0.0, tilt_y=0.0,
                             f_lens=np.inf):
    """
    Compute the real-space probe intensity, probe phase,
    and diffraction pattern for a given set of aberration coefficients.
    """

    # --- Real-space grid ---
    dx = L / N
    x = np.linspace(-L/2, L/2, N)
    X, Y = np.meshgrid(x, x)
    R = np.sqrt(X**2 + Y**2)

    # --- Reciprocal-space grid ---
    fx = np.fft.fftfreq(N, d=dx)
    FX, FY = np.meshgrid(fx, fx)
    rho = np.sqrt(FX**2 + FY**2)
    phi = np.arctan2(FY, FX)

    # --- Soft aperture ---
    radius = L / 3
    softness = 1e-9
    t = np.clip((R - radius) / softness, -100, 100)
    aperture = 1 / (1 + np.exp(t))

    # --- Lens phase (optional) ---
    if np.isinf(f_lens):
        lens_phase = 1.0
    else:
        lens_phase = np.exp(-1j * np.pi * R**2 / (lam * f_lens))

    # --- Beam tilt ---
    k = 2 * np.pi / lam
    tilt = np.exp(1j * k * (tilt_x * X + tilt_y * Y))
```

```
42
43      # --- Initial wave function ---
44      psi0 = aperture * lens_phase * tilt
45      psi0_ft = np.fft.fft2(np.fft.ifftshift(psi0))
46
47      # --- Aberration phase chi(k) ---
48      chi = aberration_phase(defocus, C3, A2, th_A2,
49                             A3, th_A3, coma, th_coma,
50                             lam, rho, phi)
51
52      # --- Apply aberrations in k-space ---
53      psi_ft_ab = psi0_ft * np.exp(1j * chi)
54
55      # --- Aberrated probe in real space ---
56      psi_ab = np.fft.fftshift(np.fft.ifft2(psi_ft_ab))
57      intensity = np.abs(psi_ab)**2
58      phase = np.angle(psi_ab)
59
60      # --- Diffraction pattern (from pupil function) ---
61      pupil = np.exp(1j * chi)
62      diff = np.fft.fftshift(np.fft.fft2(np.fft.ifftshift(pupil)))
63      diff_int = np.abs(diff)**2
64
65      return intensity, phase, chi, diff_int
```

Once the real-space intensity distribution is computed, a set of image-analysis routines is applied to quantitatively characterize the probe. These procedures operate directly on the simulated beam profile and include:

- Detection of the beam center using centroid analysis of a thresholded intensity mask,
- Estimation of the beam radius from the outer contour of the bright region,
- Extraction of the angular radial profile $r(\theta)$ by sampling the intensity along rays originating from the center,
- Identification of local maxima and minima in $r(\theta)$ to reveal characteristic aberration symmetries,
- Computation of global metrics such as the beam circularity, defined as the deviation of $r(\theta)$ from its mean value.

These methods allow rapid identification of twofold and threefold astigmatism, quantification of probe distortions, and visualisation of aberration-induced modulations in the beam shape. The algorithms were implemented in Python using custom analysis classes and are documented for reproducibility. The complete scripts, together with additional implementation notes, are provided in Appendix B.

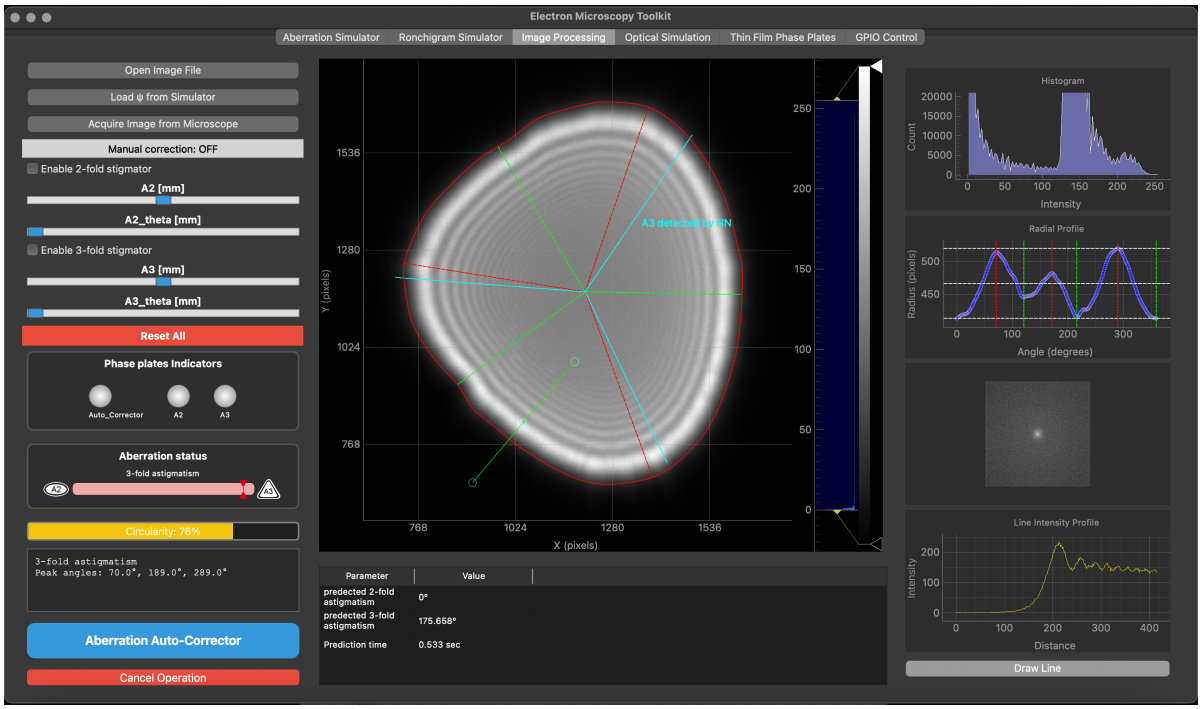


Figure 6.6. Graphical user interface (GUI) developed in Python for interactive Simulation

Image Processing and Neural-Network Analysis

A dedicated image-processing tab allows the application to connect to the microscope and acquire images for on-the-fly processing, as well as to analyse simulated or previously recorded data. The intensity distribution can be imported in three different ways:

- From the simulation tab, where only the intensity is transferred and the image-processing module has no direct information about the phase applied to the beam,
- From saved experimental data, by loading real microscope images from disk,
- Via real-time acquisition, by connecting to the microscope control software and streaming either the sample image or the diffraction pattern directly into the GUI.

Once an intensity image is available, a series of processing routines is applied to extract quantitative descriptors of the probe. The implemented tools include cropping, filtering, background subtraction, Fourier transforms, centre finding, edge detection, segmentation, and computation of radial profiles. From these quantities, the circularity of the probe is evaluated and characteristic features of the beam shape are extracted.
The key functionality of this tab is the automatic detection of aberrations using a deep neural network. The network operates on the processed intensity patterns and can identify twofold and threefold astigmatism within a few hundred milliseconds, providing the user with a fast estimate of the dominant aberration terms. Based on this information, an *Auto Correction* mode can be activated. In this mode, the software communicates with the hardware that controls the voltages applied to the phase plate, or with the microscope stigmators, and iteratively adjusts these controls to compensate both twofold and threefold astigmatism.

The architecture of the neural network and the design of the controller GUI are described in later sections of this chapter.

Controller and Automatic Tuning

The controller tab provides a direct interface to the external voltage-control board used to bias the MEMS phase plates. The GUI supports both manual control, where the user sets voltages channel by channel, and automatic control based on machine-learning models. In automatic mode, the code communicates with a trained neural network that predicts optimal voltage configurations from measured or simulated observables (for example, probe metrics or image features). The GUI sends the suggested voltages to the controller board, reads back status signals, and updates the display in real time. This tight integration between simulation, optimisation, and hardware control enabled rapid testing of new phase-plate designs and voltage configurations directly on the microscope.

6.4. Machine Learning and Neural Network

As discussed in Section 3.8, detecting and measuring residual aberrations is essential for achieving stable aberration correction. Because a significant part of this thesis focuses on MEMS-based multipoles, which will be introduced in later chapters, I developed a dedicated neural–network–based method for aberration detection. The goal of this approach is to determine the orientation of non-symmetric aberrations—specifically twofold and threefold astigmatism—directly from the shape of the electron probe.

As explained in Section 6.3.2, the Python-based GUI developed in this work can generate both the real-space probe intensity and the corresponding diffraction pattern. It also includes image-processing routines for edge detection, segmentation, line-profile extraction, and radial-profile computation. These features allow us to obtain a robust and compact description of the probe shape that is suitable for machine-learning analysis.

Using this GUI, I generated a large synthetic dataset consisting of 100,000 probe shapes with randomized aberration coefficients. For each probe, the full set of ground-truth parameters—including the magnitude and orientation of A_2 and A_3—was stored together with the computed radial profile and other geometric descriptors. All data were exported into a structured `.csv` file, which serves as the input for training the neural network. This dataset forms the basis for the supervised learning framework developed in this thesis and enables accurate prediction of aberration orientation directly from probe-shape information. An example of a simulated probe and its corresponding radial profile is shown in Figure 6.7.

To recover the orientation of non-symmetric aberrations directly from the probe shape, a supervised neural-network model was developed. The input to the model is the 360-point radial intensity profile extracted from each simulated probe. This representation captures the rotational structure introduced by aberrations such as twofold and threefold astigmatism while remaining compact and robust to noise. Before training, each profile was mean-centred

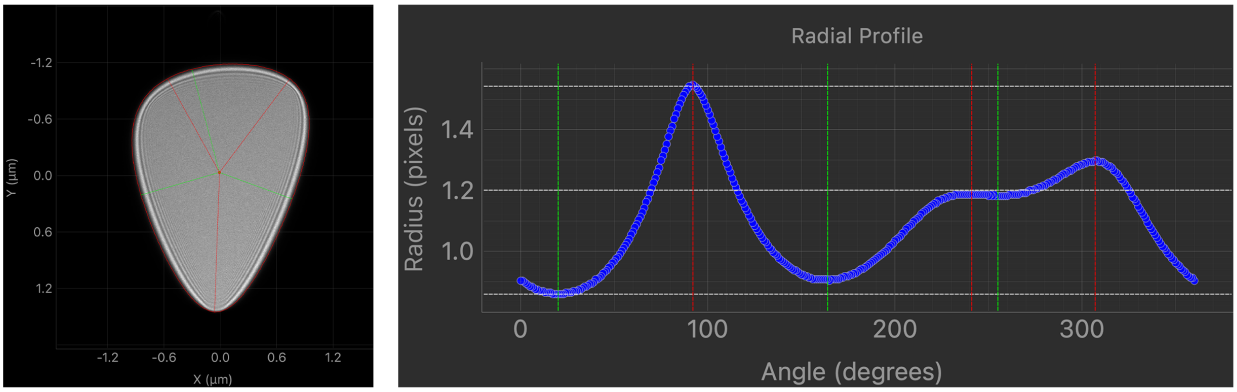


Figure 6.7. Simulation of an abberated probe shape (left) and corresponding radial profile used as ML training dataset (right)

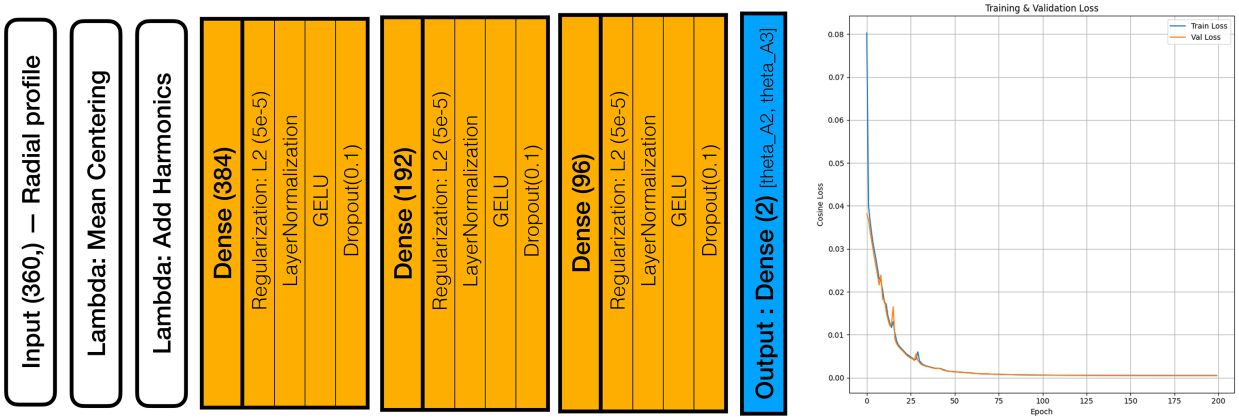


Figure 6.8. (left) Neural-network architecture with preprocessing, harmonic augmentation, fully connected layers, and a two-neuron output encoding the aberration orientation; (right) training and validation loss history for the same model.

to remove global intensity variations, and weak-aberration cases were filtered to ensure that the training set contained sufficiently strong orientation information.

For each aberration type, the angular parameter θ was encoded using its natural harmonic symmetry. In the case of A_3, the angle was transformed into the threefold invariant representation $[\cos(3\theta), \sin(3\theta)]$, which avoids discontinuities and ensures smooth learning. A similar encoding was used for twofold astigmatism. To help the network identify the characteristic periodic patterns associated with different aberration modes, additional harmonic projections of the radial profile were computed inside the network and concatenated with the input features.

The neural-network architecture consists of several fully connected layers with GELU activation, layer normalisation, and dropout regularisation. This structure provides a good compromise between expressive capacity and numerical stability. A diagram of the complete architecture, including the preprocessing, harmonic augmentation, and dense layers, is shown in Figure 6.8. The output layer contains two neurons corresponding to the encoded angular components of the aberration orientation.

Training was performed using the Adam optimiser and a cosine-similarity loss function,

which directly penalises angular discrepancies between predicted and ground-truth harmonic vectors. A learning-rate scheduler and an early-stopping criterion were employed to avoid overfitting and ensure convergence. To increase sensitivity to aberrations with larger amplitudes, a smooth weighting scheme was applied during training so that samples with stronger A_2 or A_3 contributions were emphasised.

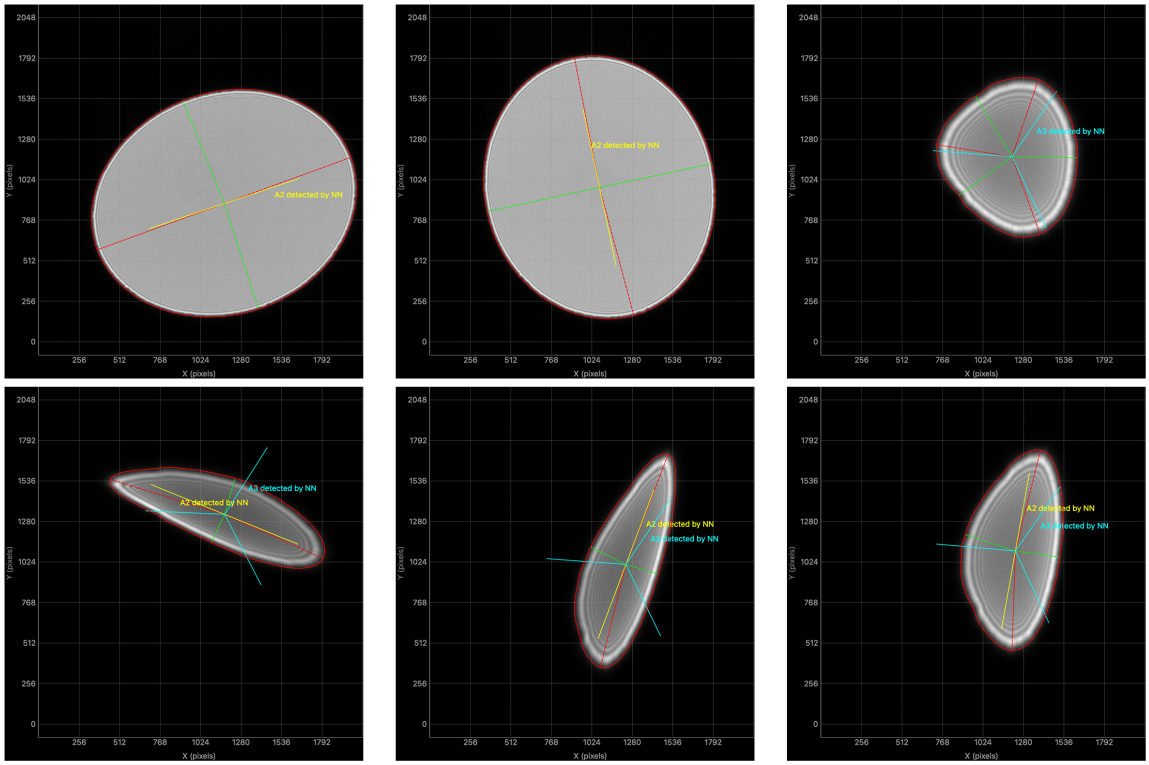


Figure 6.9. Experimental probe images with deliberately applied twofold astigmatism at different magnifications. Overlaid markers indicate the neural-network–predicted orientations of the twofold and threefold astigmatism.

The model was trained on a synthetic dataset of 20,000 simulated samples, using an 80/20 train–test split, and showed stable convergence across a wide range of aberration strengths. On the test set, the network predicted the aberration angle with sub-degree accuracy (median error 0.08°; mean error 0.12°), with a maximum error of 3.89° and 100% of predictions within 5° angular error. The final trained model was integrated into the Python GUI, enabling real-time estimation of aberration orientation from simulated or measured probe profiles. To verify the performance of the trained model on real experimental data, the neural network was also applied to probe images acquired directly from the microscope. These measurements provide an independent validation of the model and confirm its ability to extract aberration orientation from realistic probe shapes affected by noise, detector response, and residual misalignments. Representative examples of the detection results, together with the corresponding predicted orientations, are shown in Figure 6.9.

7 Designs

Contents

7.1 Overview 122

7.2 Multipoles 123

7.2.1 Electrostatic multipole : Hyperbola phase plate 124

7.2.2 Current-controlled electrostatic multipole : "Helios" phase plate . 129

7.3 Spiral phase plate 133

7.4 OAM sorter phase plate 137

7.1. Overview

This chapter presents the design strategies used for developing electrostatic phase plates. The design workflow, summarized in Figure 7.1, starts with an analytical target phase profile that defines the desired electron-wave modulation. This analytical model is first used to evaluate the feasibility of the phase plate. The design is then simulated using finite-element tools combined with wave-optics calculations, which allow us to predict the phase plate's performance from different perspectives and verify its agreement with the target phase profile. Chapters 5 and 6 introduced the general design methods for MEMS-based phase plates and the numerical tools required for their implementation. Building on these foundations, this chapter focuses on four specific phase plate designs: (i) an electrostatic dodecapole phase plate for compensating threefold astigmatism, (ii) a redesigned hexapole phase plate based on a current-controlled approach, analogous to the dodecapole, (iii) a current-controlled vortex-beam generator, and (iv) a current-controlled orbital angular momentum (OAM) sorter. For each design, we begin with the analytical phase model, proceed to a complete simulation workflow to evaluate functionality, and conclude with the final device layout prepared for silicon-chip fabrication.

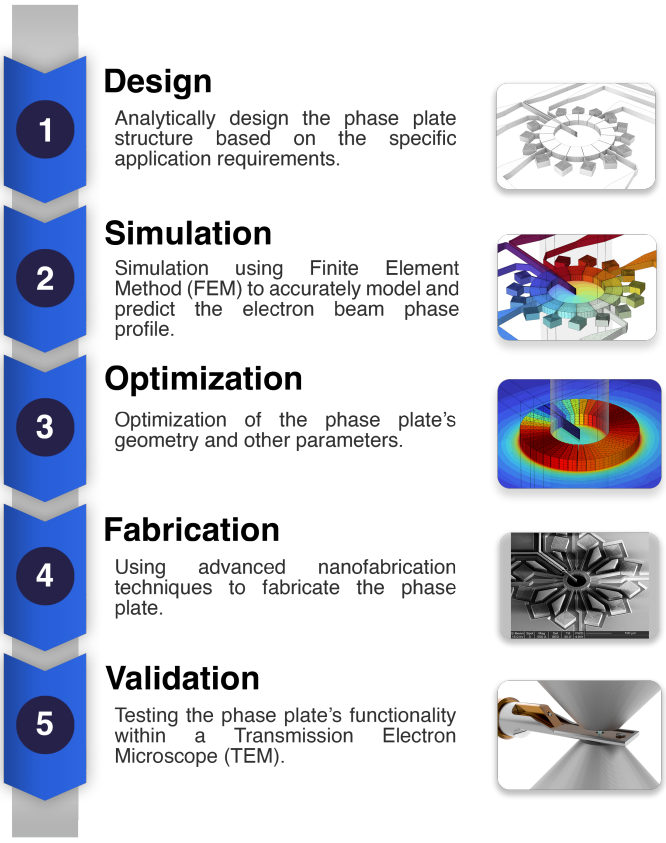


Figure 7.1. Phase plate design process flow

7.2. Multipoles

In the previous chapters, we discussed that modern aberration correctors rely heavily on multipole elements to compensate higher-order aberrations. Among these elements, the hexapole plays a central role in spherical aberration (C_s) correctors (also see section 3.7). Although these correctors provide excellent performance and represent a mature technology, they remain bulky and expensive, require complex manufacturing and precision machining, and demand tight integration into the microscope column. This motivates the use of compact MEMS-based multipole phase plates. For a hexapole to function as a threefold astigmatism generator, two main criteria must be satisfied. First, the induced phase shift must reproduce the Zernike polynomial corresponding to the threefold astigmatism term:

$$\chi_{\text{threefold}}(k) = \frac{2\pi}{3} f_{a3} \lambda^2 k^3 \sin[3(\phi - \phi_{a3})] . \tag{7.1}$$

Second, the design must allow for controlled rotation of this phase term, since tuning the orientation ϕ_{a3} is essential for precise aberration correction. In principle, a single electrostatic hexapole with six electrodes can generate a fixed threefold astigmatism (see Figure 7.2(a)); however, such a configuration cannot provide continuous and independent phase rotation.

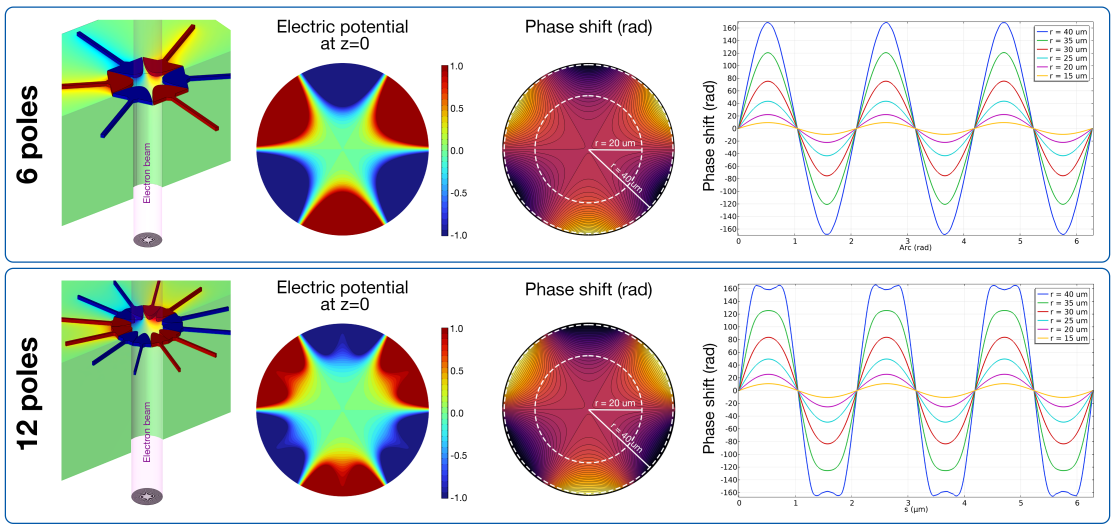


Figure 7.2. Comparison between two hexapole configurations used for generating three-fold astigmatism: a conventional 6-pole design (Hexapole: fixed orientation) and an extended 12-pole (dodecapole: tunable orientation).

In this work, I investigated two complementary approaches for implementing multipole-type phase control. The first approach is a purely electrostatic design based on a dodecapole geometry with twelve independently biased electrodes. This architecture enables an effective hexapole field whose orientation can be tuned by optimizing the applied potentials, without mechanical alignment. The second approach is a current-controlled electrostatic phase plate that generates a comparable hexapole-like phase profile within a circular aperture and provides tunability, including smooth rotation of the phase orientation. The following sections discuss both approaches in detail.

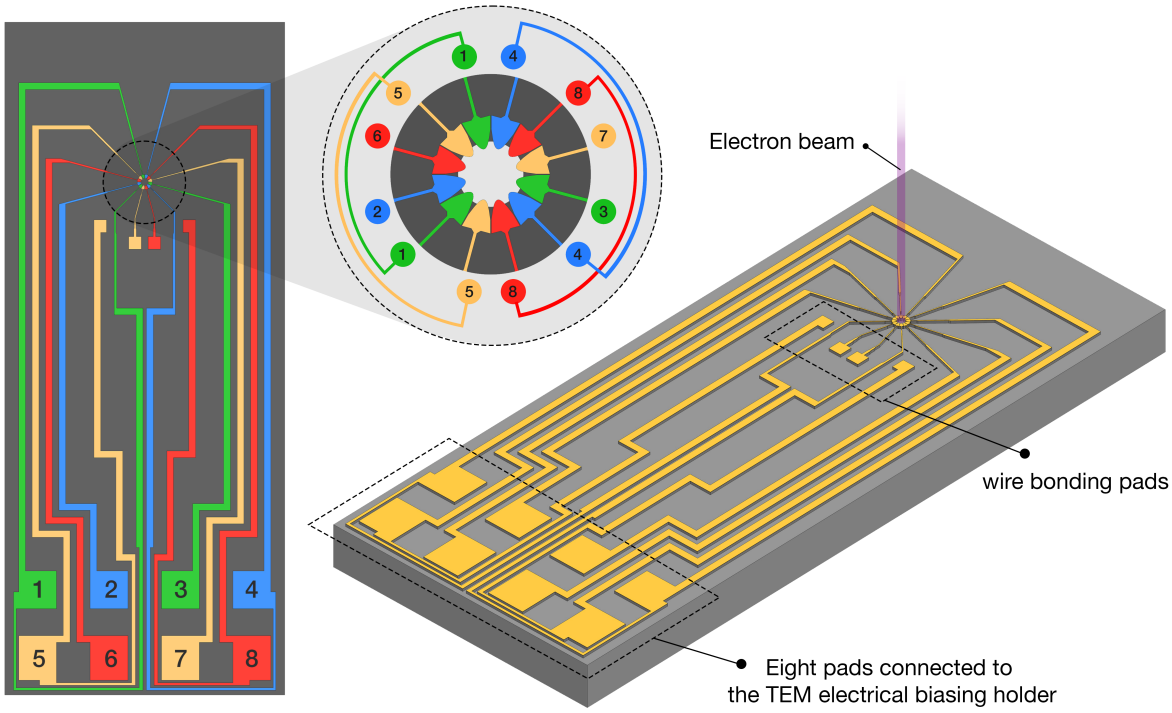


Figure 7.3. hyperbola-layout

7.2.1. Electrostatic multipole : Hyperbola phase plate

Figure 7.3 shows the chip layout for the dodecapole design. The chip is fabricated on a silicon-on-insulator (SOI) wafer composed of three main layers: the device layer containing the patterned electrodes, the box layer that electrically insulates the circuit from the substrate, and the silicon handle layer serving as the substrate. A detailed discussion of the wafer structure is provided in the next chapter.

To make the device fabrication-feasible and to reach the targeted phase shift, several practical aspects must be considered. First, because the electrical biasing holder provides only eight independent biases, the chip must generate twelve potentials in the active area using these eight biases. In silicon chip designs, due to the 2D pattern, routing limitations arise from the inability to cross electrodes on a single layer. This constraint can be addressed using two common strategies:

1. **Wire bonding**, where a thin metallic wire connects two circuit nodes externally.
2. **Through-Silicon Vias (TSVs)**, which are high-aspect-ratio metal-filled vertical interconnects etched through the silicon substrate.

For our device, wire bonding is sufficient. As shown in Figure 7.3, by optimizing the 2D connections layout and adding four additional pads for wire-bond connections, we are able to apply all twelve required potentials. Although this configuration means that not all electrodes are fully independent, the design still provides the necessary degrees of freedom for generating the double hexapole field, which requires four independent applied potentials (shown as

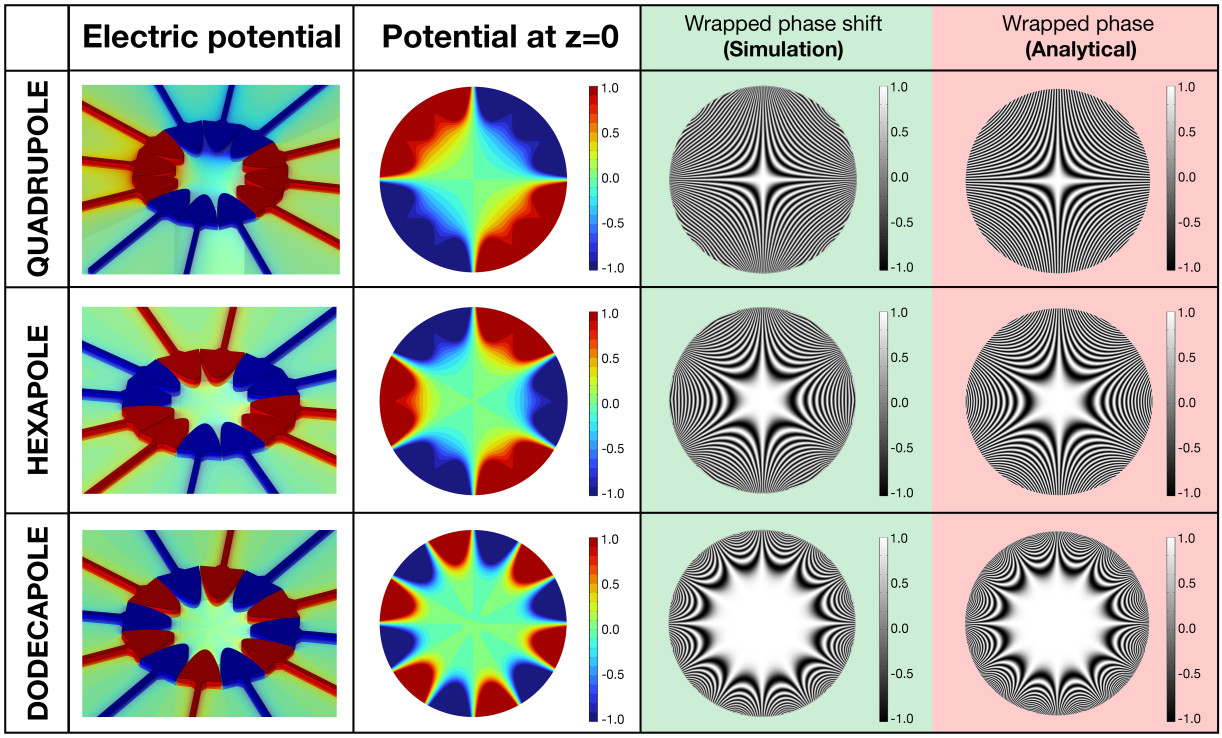


Figure 7.4. Simulated phase profiles generated with the dodecapole phase plate, including quadrupole, hexapole, and dodecapole modes, shown together with the corresponding analytical Zernike polynomials.

four different colors in Figure 7.3). However, a future design with twelve truly independent electrodes would offer finer control and enable fully tunable phase shifting.

Secondly, the electrode geometry must be optimized. In all simulations, we consider the realistic chip layout, including the full 3D structure, to ensure accurate modeling of the electric potential. As a result, the electrode shapes cannot perfectly follow the mathematical ideal since the grounded silicon substrate breaks the symmetry of the three-dimensional field distribution. Nevertheless, hyperbola-shaped electrodes provide a good initial guess for the optimization. For this purpose, I employed the shape optimization tools in COMSOL. Using Zernike polynomials as the objective function for the phase shift, I optimized the electrode shapes while keeping the sidewalls straight, which is a fabrication constraint inherent to silicon chip processing. Details of the optimization process are given in Section 6.1.5.

Using this dodecapole layout provides sufficient freedom to generate a wide range of phase profiles. Although the main goal is to realize a rotatable and tunable hexapole, the geometry can, in principle, produce orientable dipoles and quadrupoles, as well as a dodecapole with a fixed orientation. The limited degrees of freedom given by the 8 contacts further limit the available terms. Figure 7.4 shows several of these configurations, together with their simulated phase shifts and a comparison with the corresponding analytical Zernike polynomials.

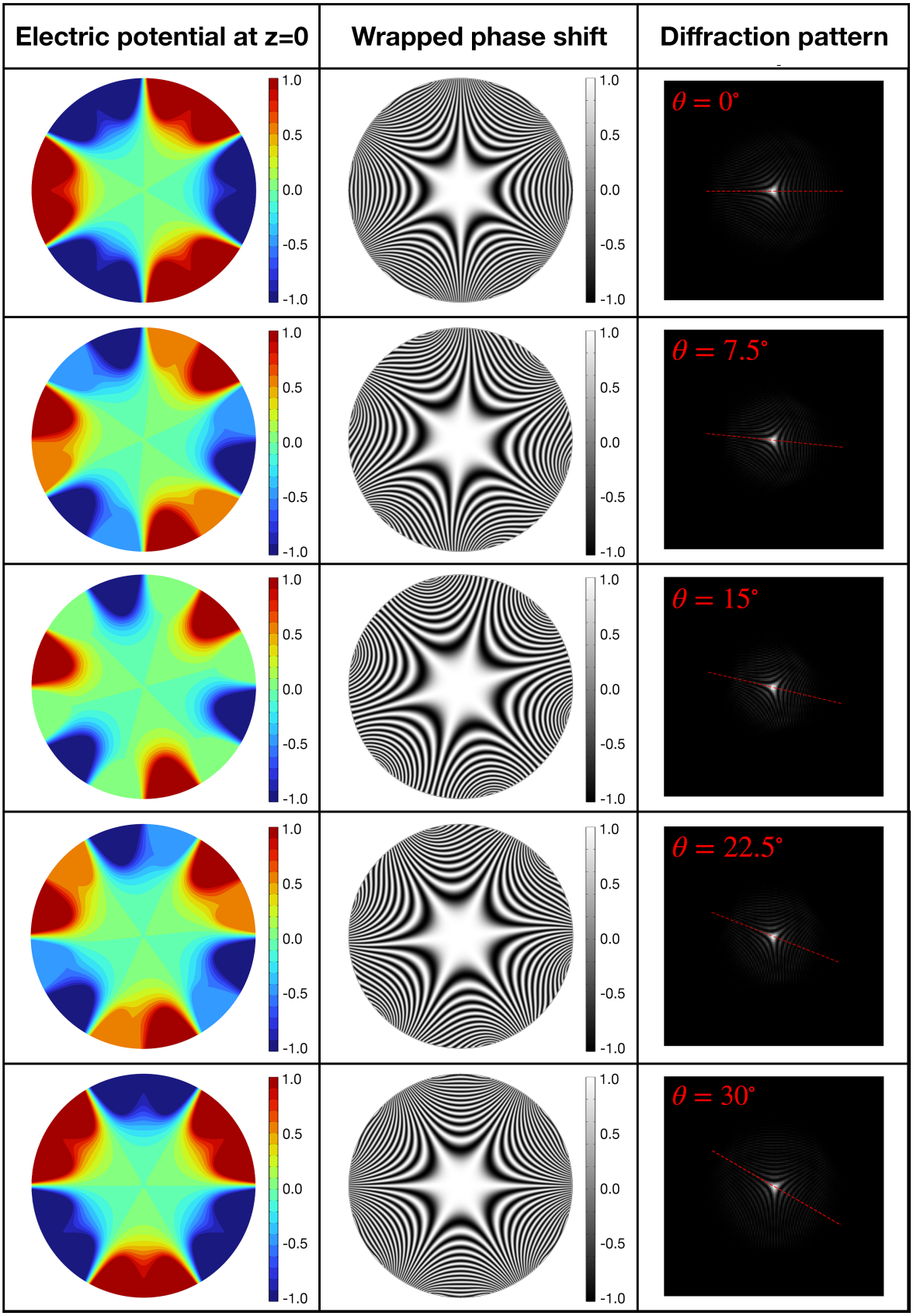


Figure 7.5. Hyperbola phase plate: optimized electrode potentials for rotating the three-fold phase from 0° to 30°, together with the corresponding simulated phase and Fourier-transform (in-focus diffraction patterns).

The next step toward designing a tunable phase plate is to determine the optimized potentials required for any orientation of the three-fold phase. By using COMSOL's gradient-based parameter optimization and employing the corresponding Zernike polynomial for different values of ϕ as the objective function, I obtained the set of potentials required for each rotation angle. Because three-fold astigmatism has a periodicity of $\pi/3$, full rotational control can be achieved by optimizing only within a 60° range. Figure 7.5 illustrates five optimized sets of applied potentials from 0° to 30°, together with the simulated phase and the Fourier transform of the wavefunction (corresponding to the in-focus diffraction pattern in the microscope). These simulations were performed for a device with an 80 μm diameter and a 20 μm thickness. To extend the rotation to the full 60°, the negative sign of the reported potentials can be used.

Although the simulations were performed using a +1 V excitation, it is worth noting that the total phase shift can be increased by applying higher potentials. Figure 7.6(a) and (b) show the radial phase shift for different applied voltages and two device diameters. The results clearly demonstrate that reducing the device diameter decreases the voltage required to achieve the same phase shift. This is an important consideration and an advantage of using MEMS-based phase plates compared with magnetic multipoles. Since the required phase shift depends on the beam diameter at the position of the hexapole within the optical column, different device sizes may be needed for different microscope configurations.

Another parameter that can be used to increase the phase shift—particularly relevant for applications such as spherical-aberration correctors—is the thickness of the device. Figure 7.6(c) shows how increasing the device thickness leads to a larger phase shift, especially at larger beam radii.

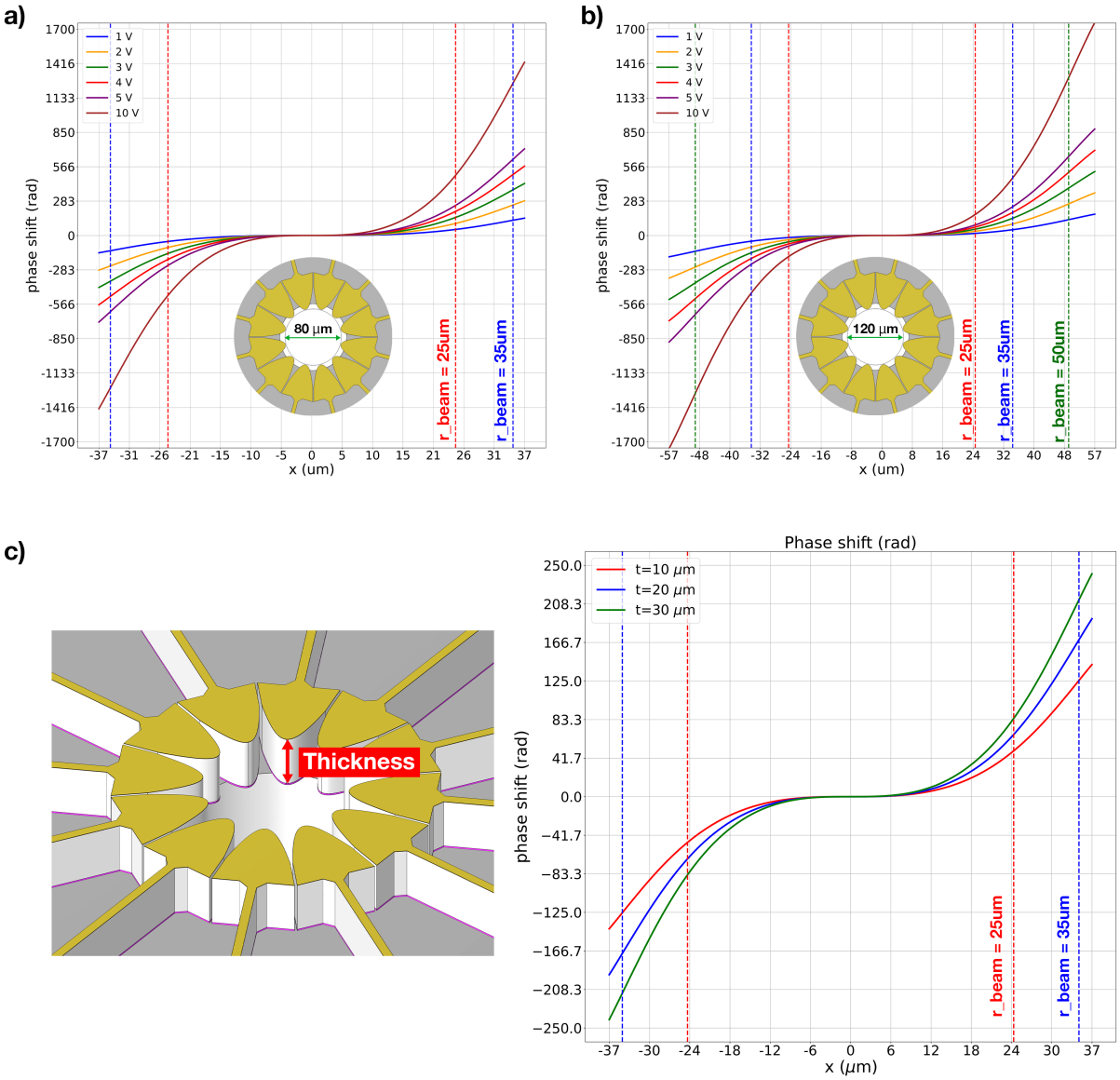


Figure 7.6. Effect of device radius and thickness on the total phase shift. The phase is evaluated along a line profile through the center of the phase profile. (a) Phase shift for an 80 μm device diameter at different applied voltages. (b) Phase shift for a 120 μm device diameter under the same voltage conditions. (c) Phase shift for different device thicknesses.

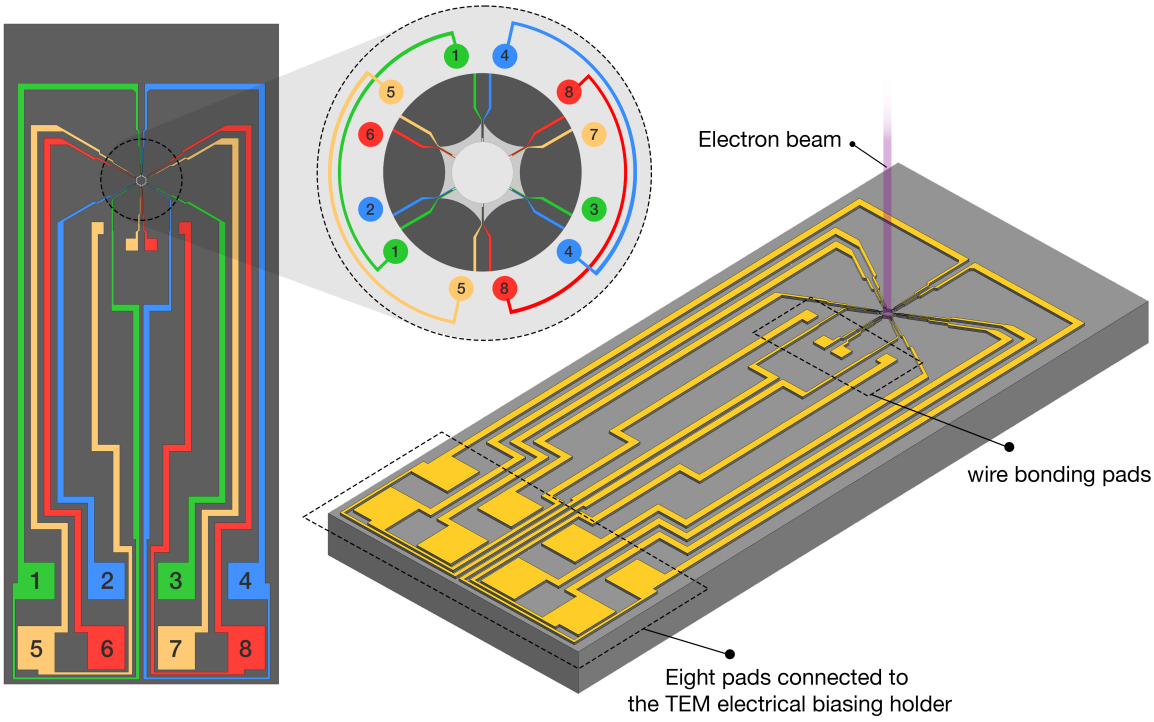


Figure 7.7. Helios layout and design configuration

7.2.2. Current-controlled electrostatic multipole : "Helios" phase plate

Although the hyperbola design exhibits near-perfect phase shifting and full rotatability, at larger beam radii—where the electrons pass close to the electrodes—the beam begins to experience the discrete nature of the electrode arrangement. This introduces an unwanted dodecapole component in the phase profile. To suppress this effect, one would need to increase the device diameter and add an aperture to limit the beam. A more effective approach, however, is to design a device that provides the same phase profile while also offering a circular aperture. For this purpose, I introduce the current-controlled design.

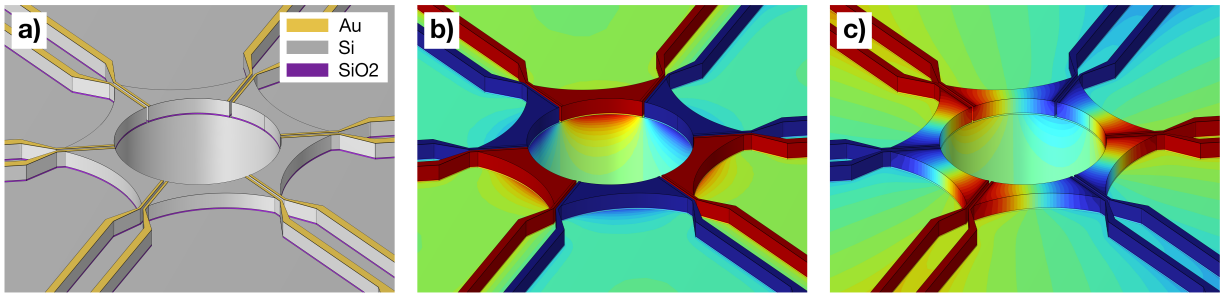


Figure 7.8. Helios current-controlled electrostatic phase plate: (a) geometry layout showing the materials and overall design; (b) electrostatic mode, where both ends of the electrode are connected to the same potential; (c) current-controlled mode, where different potentials are applied to the two ends of the electrode.

Figure 7.7 shows the schematic layout of the chip pattern, which is identical to the layout used for the hyperbola design, except in the active area. I added four additional pads for

wire bonding, enabling the use of twelve electrodes. As discussed in Chapter 5, in current-controlled designs, we employ a mid-range conductive material—such as doped silicon—for the electrodes near the active region, while all long connections are made from a highly conductive material. This is achieved by coating the full chip with a thin layer of metal (e.g., gold), except on the electrodes themselves (see also Figure 7.8(a)). The fabrication process will be described in detail in the following chapters. In this design, by applying optimized potentials to the head and tail of each electrode, we can generate a controlled potential gradient along the electrode length. Alternatively, applying the same potential to both ends produces a fixed potential equivalent to an electrostatic phase plate. Figures 7.8(b) and (c) illustrate two operating modes of the current-controlled design.

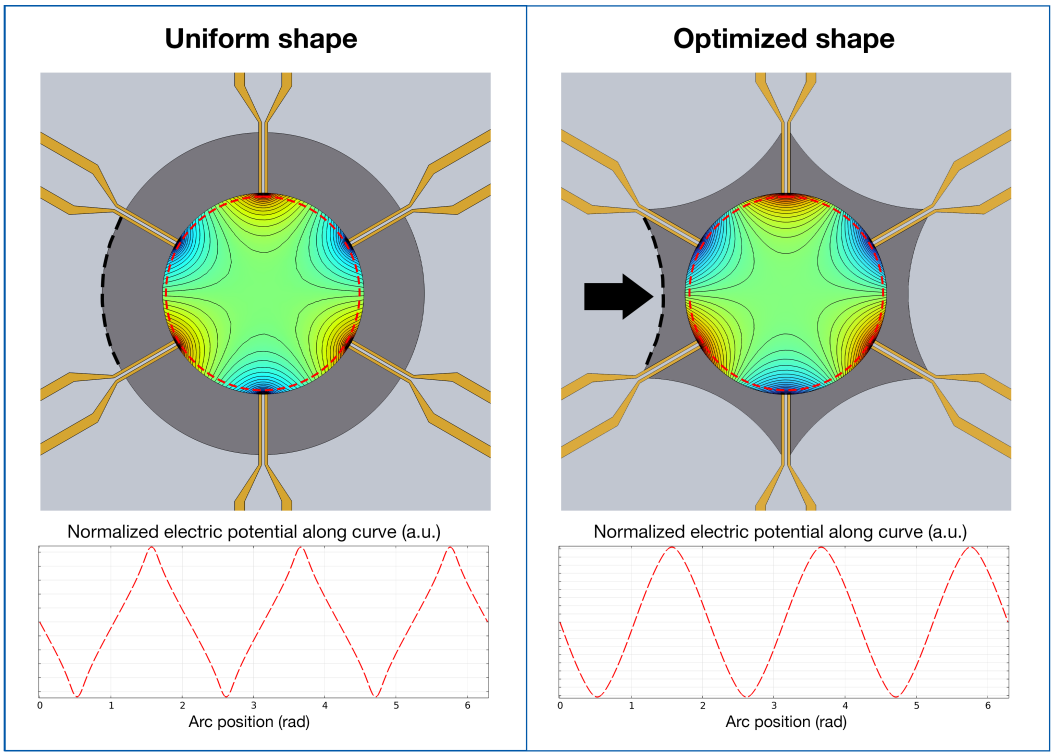


Figure 7.9. Optimized electrode shape (black dashed line) and electric potential at $z = 0$ along the circular contour near the boundary electrodes (red dashed line).

As shown in Figure 7.9, the electrode shapes were optimized to achieve the desired phase shift for all rotation angles ϕ. To accomplish this, we followed the same COMSOL shape-optimization procedure used for the hyperbola design; however, this time, the objective was to control the potential drop along the electrodes in the active area. The optimization results showed that using a curved electrode contour instead of a fixed, uniform shape satisfies the objective function for all rotational modes. Figure 7.10 shows five optimized sets of applied potentials that allow continuous rotation of the phase performed for a diameter of 80 μm and a device thickness of 20 μm. It is worth noting that parameter optimization must be performed for each device geometry with different diameters and device thicknesses. However, by using the same set of potentials as an initial guess, the optimization converges quickly and typically requires only a few minutes to find the optimal solution.

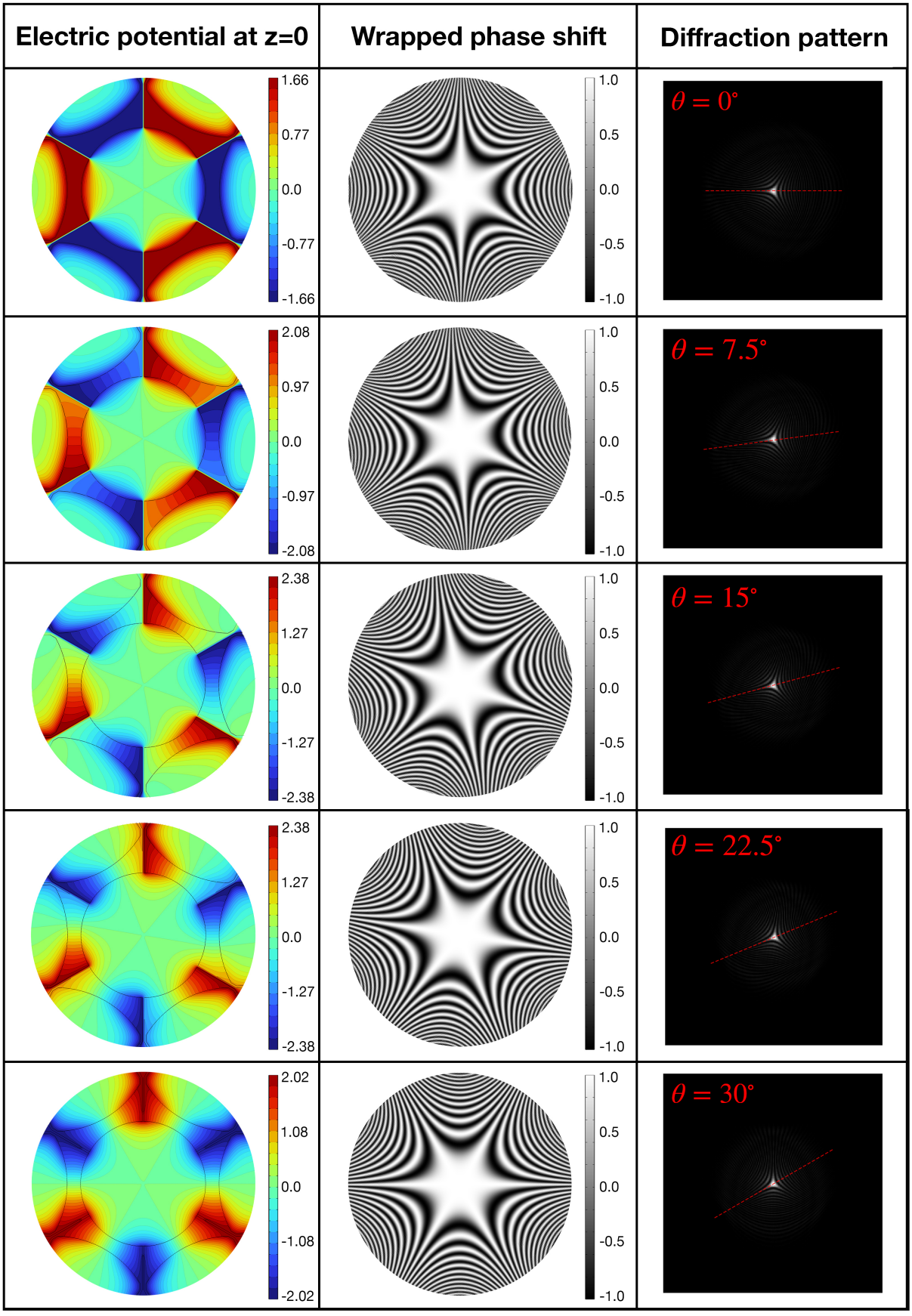


Figure 7.10. Helios phase plate: optimized electrode potentials for rotating the three-fold phase from 0° to 30°, together with the corresponding simulated phase and Fourier-transform (in-focus diffraction patterns).

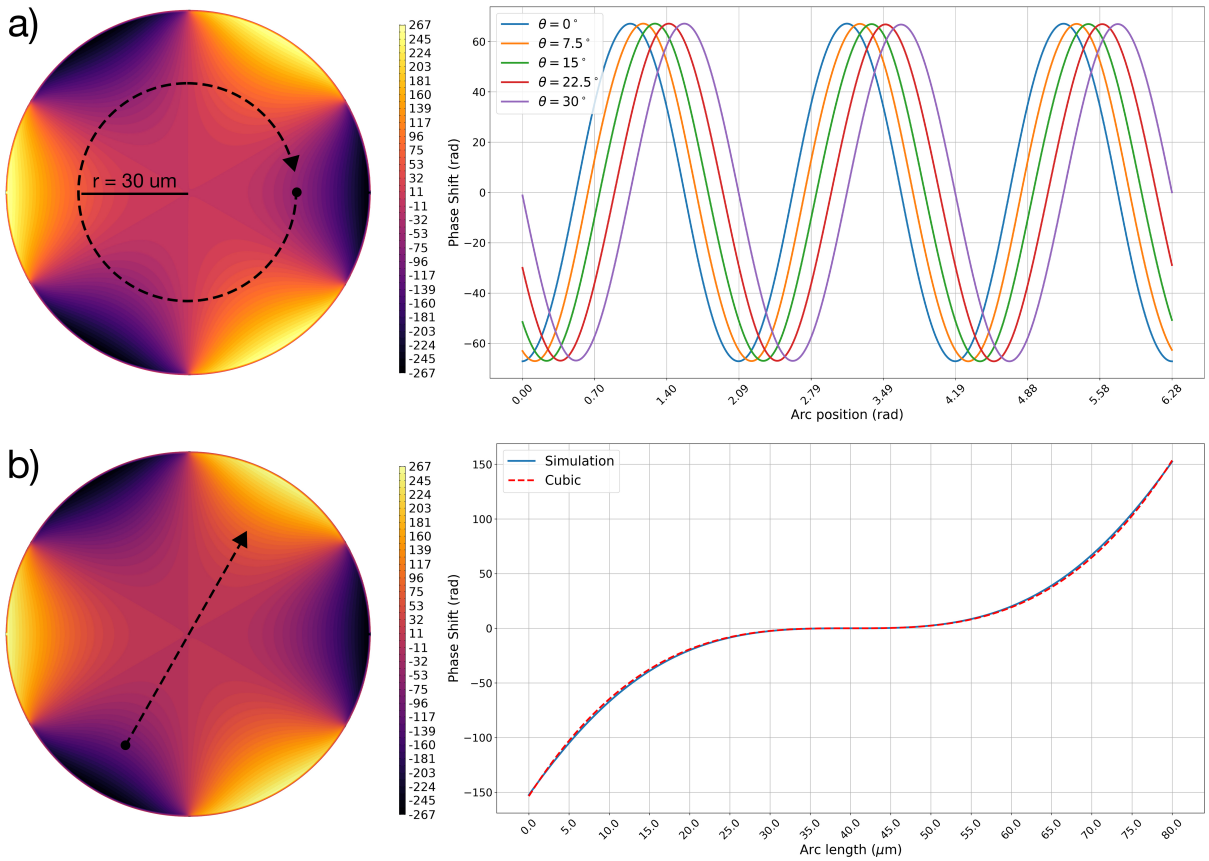


Figure 7.11. Helios phase plate: simulated phase shift showing (a) the angular phase profile along a circular contour for five orientations and (b) the phase line profile along a central line through the center of the phase profile.

Figure 7.11 shows the simulated phase shift along a circular contour and the corresponding radial profiles for all five orientations. The results demonstrate that the optimized potentials produce phase profiles that match the expected $\cos(3\phi)$ angular dependence (Figure 7.11(a)), while the radial component follows the cubic behavior expected from the analytical model (Figure 7.11(b)).

7.3. Spiral phase plate

As mentioned in Section 4.2.3, generating a vortex beam with a MEMS phase plate requires an azimuthally varying phase with an approximately linear dependence on the azimuthal angle θ, while maintaining (to first order) a constant phase with radius over the usable aperture. Therefore, the analytical vortex phase $\phi(\theta) = \ell\theta$ provides a natural reference for the device design and for defining the target phase response. Figure 7.12 presents the analytical vortex phase in wrapped and unwrapped for for $l = 10$, together with the phase profile evaluated along the dashed circular contours indicated in Figure 7.12(b).

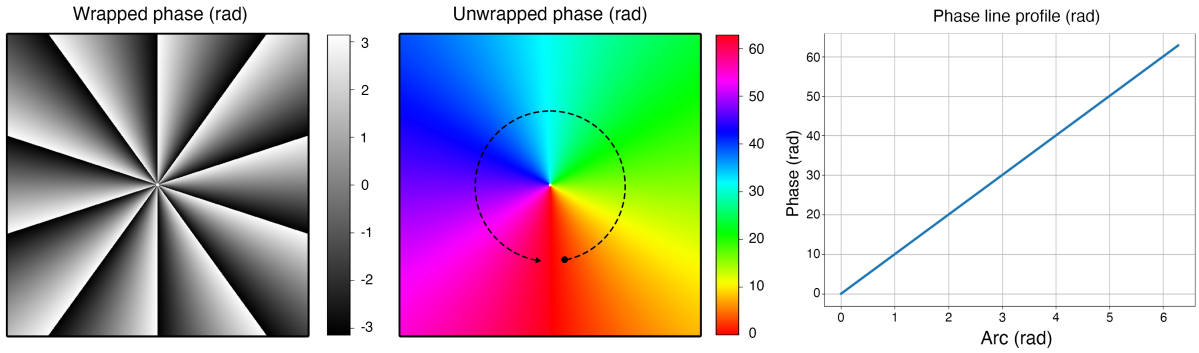


Figure 7.12. Analytical vortex phase in wrapped and unwrapped form, and the phase profile extracted along the dashed circular contour, showing the linear azimuthal dependence.

In our implementation, the core of the vortex generator is a pair of needle electrodes (the "chopsticks") biased with opposite polarities similar to a previous design developed in our group [98]. This geometry produces the desired phase variation close to the needle tips [138, 98]. However, for wider beams, additional boundary elements are required to preserve the target phase over a larger region. We therefore introduce boundary electrodes distributed along the circular aperture to shape the three-dimensional electrostatic potential and extend the usable region in which the phase follows the vortex reference. In this section, parameter optimization is used to determine an optimal set of boundary-electrode voltages. These results are then used as inputs for a second optimization stage based on *shape optimization*. In this step, we design a current-controlled electrostatic vortex phase plate that reduces the number of independent control channels to eight biased electrodes, while still providing fine phase control.

The optimization was performed in COMSOL using a gradient-based parameter-optimization Nelder-mead method. Figure 7.13 shows the simulated geometry, which includes the chopstick electrodes, 41 boundary-electrode pairs arranged around the aperture, and a substrate to mimic a realistic environment. Although 41 independently addressable voltages are not practical with current electrical

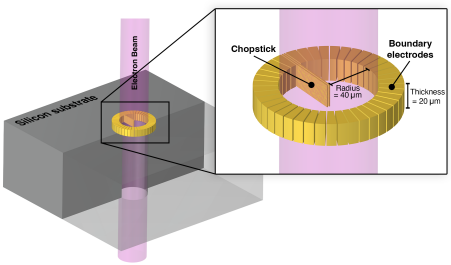


Figure 7.13. Simulation geometry for parameter optimization, including 41 discretized boundary electrodes.

biasing holders, this discretized configuration provides a high-resolution estimate of the boundary voltage distribution. The device diameter is 80 μm and the thickness is 20 μm, which is typical for MEMS fabrication in our process flow. It is important to note that the optimized voltage distribution depends on the device thickness; here we present one representative thickness.

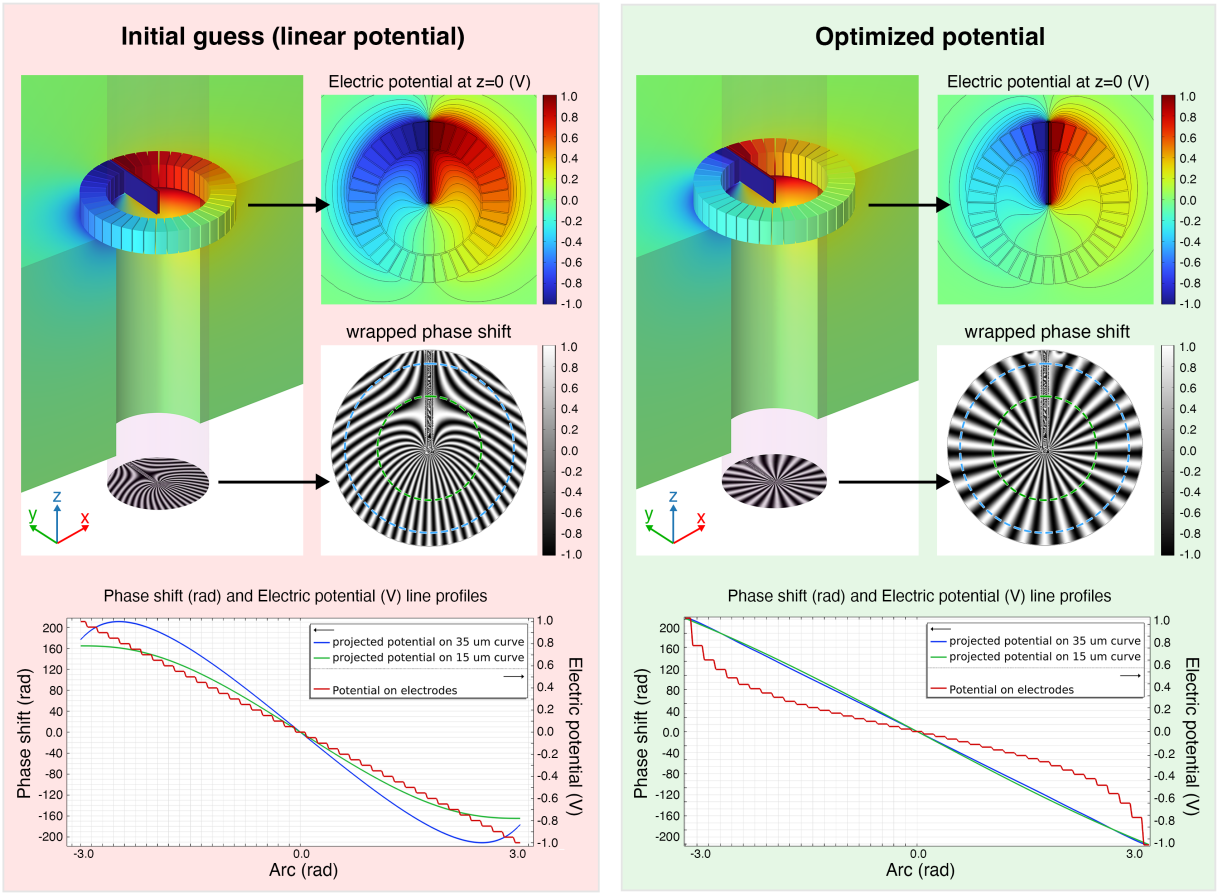


Figure 7.14. Boundary-electrode parameter optimization. Left: initial linear voltage ramp used as the initial guess. Right: optimized solution. The red profile indicates the required boundary-electrode voltage trend.

As an initial guess, the boundary-electrode voltages were assigned a linear variation from $+1$ (matching the positive needle) to -1 (matching the negative needle) along the azimuthal coordinate. The objective function enforced a linear phase ramp on two circular contours at $r = 15\,\mu$m (inner) and $r = 35\,\mu$m (near the boundary). Figure 7.14 shows the simulated electrostatic potential and the total phase shift, calculated using the COMSOL "general projection" operator, both before and after optimization. The parameter-optimization results indicate that the boundary-electrode voltage profile exhibits a steep drop over the first few electrodes, followed by an approximately semi-linear trend over the remaining electrodes. This observation motivated the move to a current-controlled electrostatic phase-plate concept, where the slope and the overall monotonic trend can be enforced through geometry (via controlled current flow) and refined using shape optimization.

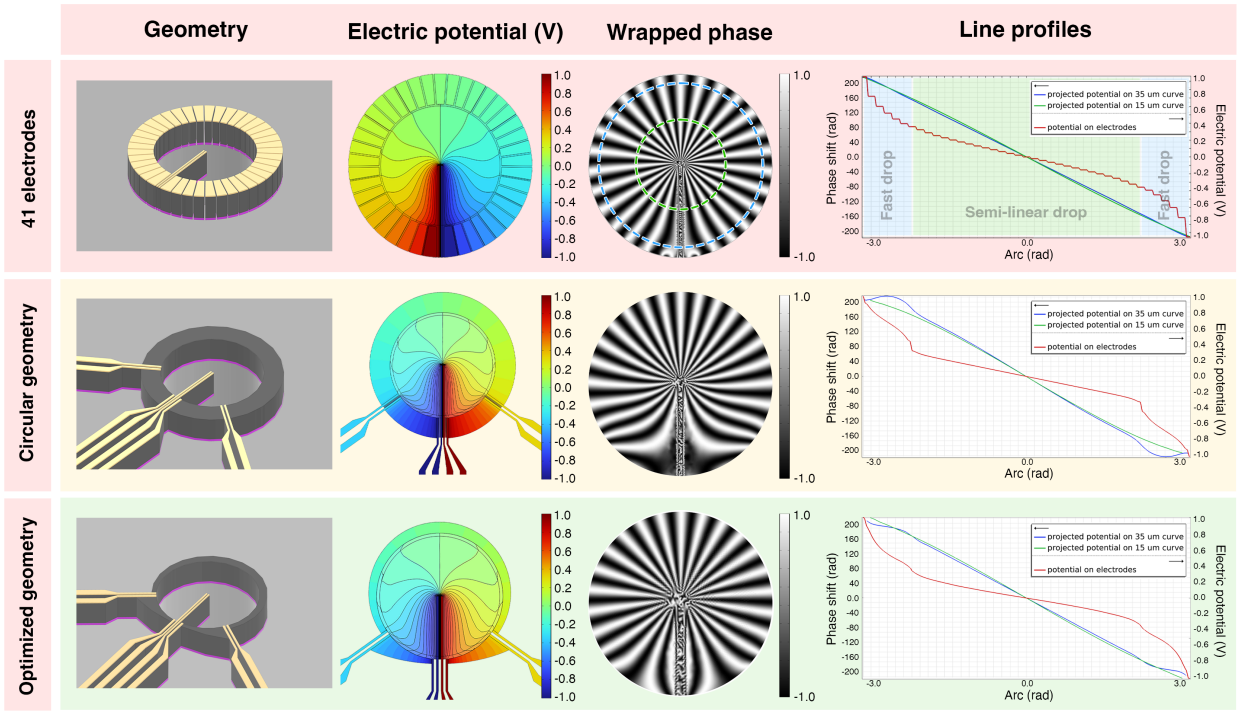


Figure 7.15. Shape-optimization workflow for designing a current-controlled electrostatic vortex-beam generator.

Figure 7.15 summarizes the three-stage shape optimization strategy. In Stage 1, we obtain a high-resolution boundary-voltage distribution using the discretized design with 41 independently biased electrode pairs. In Stage 2, we construct an initial current-controlled design by dividing the boundary-electrode region into three segments that represent the "fast-drop" and "semi-linear drop" parts of the voltage trend, and we implement these segments using a circular, uniform baseline geometry. In Stage 3, we apply shape optimization to deform the boundaries between the segments and adjust their geometry until the simulated phase response matches the target vortex phase over the usable aperture. the total number of vortex l can be tuned by adjusting the potential. Figure 7.16 shows the full-chip design layout.

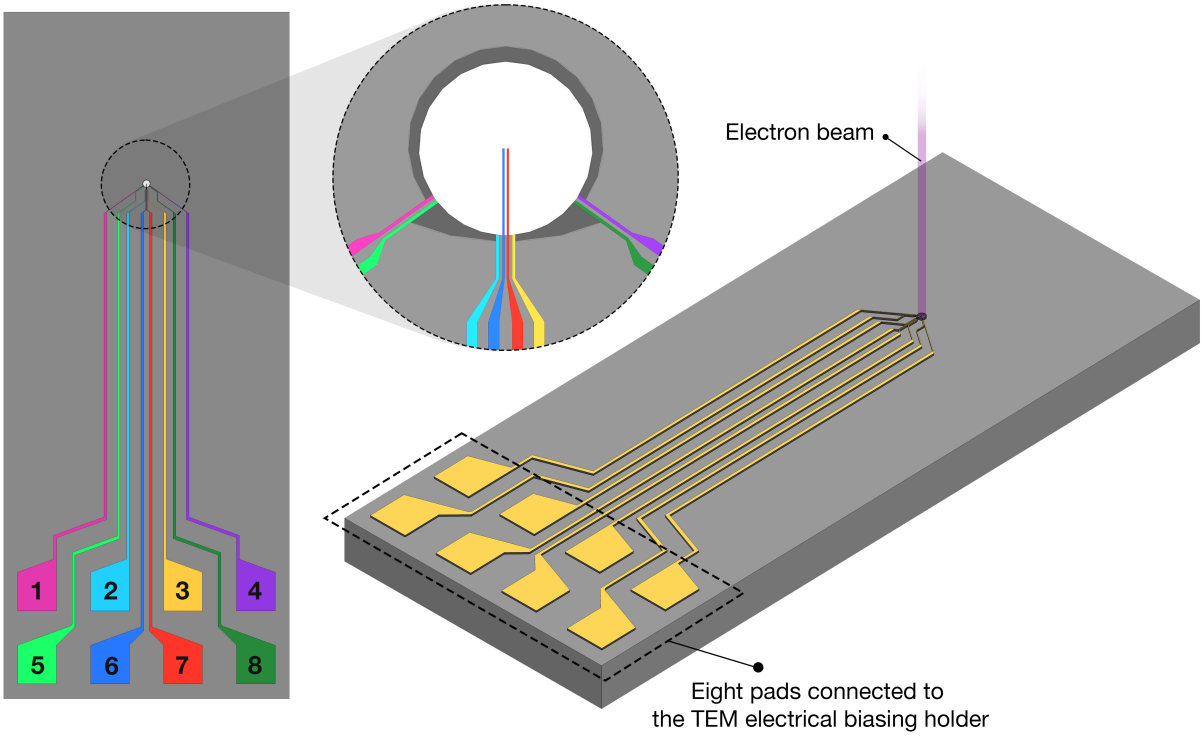


Figure 7.16. Current-controlled electrostatic vortex beam generator phase plate design and chip layout.

7.4. OAM sorter phase plate

In designing our OAM sorter, we adopted the electrostatic "unwrapper" (sorter 1) proposed by Pozzi *et al.* [162], which in phase profile is given by [163]

$$\phi(x,y) = \frac{d}{\lambda f}\left[y \arctan\left(\frac{y}{x}\right) - x \ln\left(\frac{\sqrt{x^2+y^2}}{b}\right) + x\right]. \tag{7.2}$$

In the ideal case, this phase implements a log–polar mapping that converts the azimuthal dependence of the incoming beam into a linear phase ramp. Pozzi *et al.* showed that a long biased needle placed in front of a conducting plane can reproduce this phase distribution with good accuracy, so that the projected electrostatic potential of the needle provides a practical realization of sorter 1. Its feasibility has already been demonstrated experimentally in the TEM using electrostatic needles and a corresponding corrector stage [135]. The analytical phase map is shown in Figure 7.17 together with the phase line profile, which is used as the target phase.

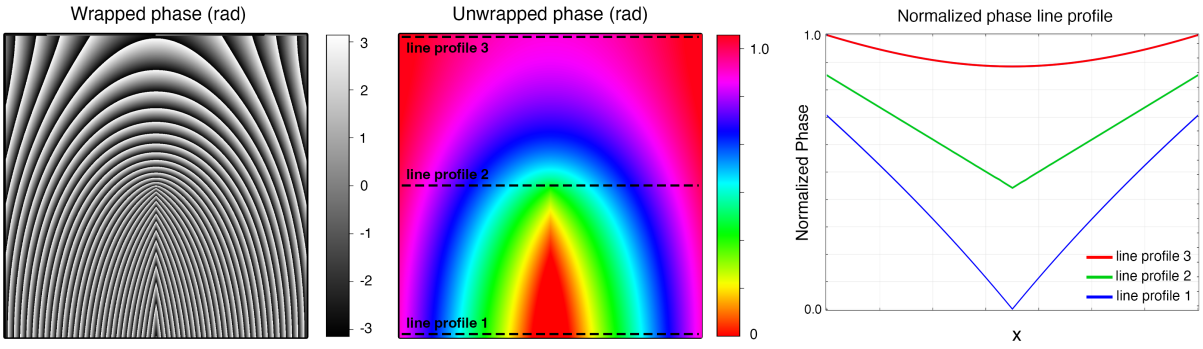


Figure 7.17. OAM sorter analytical target.

Building on this idea, I redesigned and optimised the geometry, position, and bias of the needles within a current-controlled electrostatic design. In this approach, the high degree of freedom in setting the boundary-electrode potentials allows finer control of the phase shift by tuning the biases. The goal is to make the simulated electrostatic phase shift closely follow Eq. 7.2, while minimising astigmatism and other distortions introduced by finite-size effects and boundary conditions.

The optimization process follows the same strategy as for the other phase-plate designs: (1) parameter optimization to determine the bias potentials on the needles and boundary electrodes; (2) shape optimization to refine the boundary-electrode geometry, which is particularly important in the current-controlled design.

Figure 7.18 shows the final optimized geometry and potential distribution. The device consists of five boundary electrodes (E1–E5) and three needles (N1–N3). In the current-controlled implementation, each boundary electrode is connected to two external contacts to control the potential via resistive paths. Thanks to the symmetry of the target phase, mirror-symmetric electrodes can be biased with corresponding mirror potentials, so that the

required potential landscape is realized with only eight external biases. The wire-bonding pads are arranged accordingly in a symmetric layout.

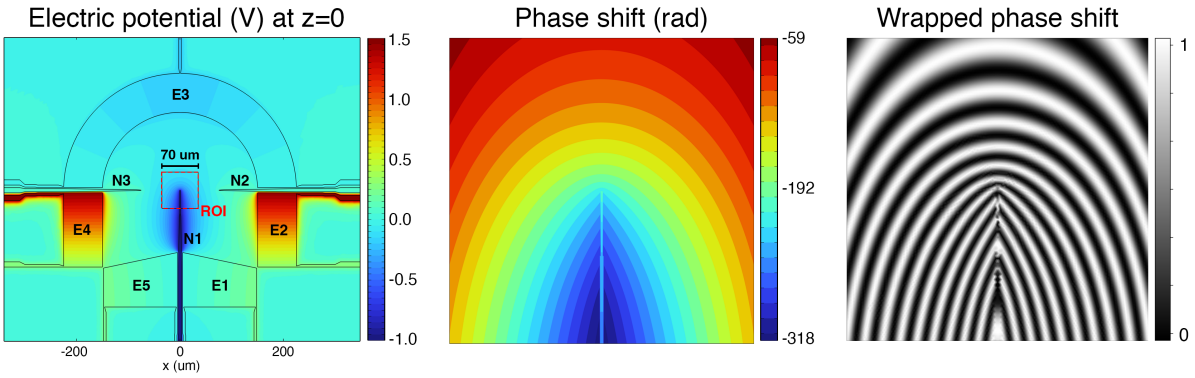


Figure 7.18. Current-controlled OAM sorter: (a) optimized device geometry (300 μm size) with the 70 μm region of interest (ROI) indicated by a red dashed square; (b) unwrapped phase shift in the ROI; (c) wrapped phase shift in the ROI.

Figure 7.19 compares the analytical phase map with the simulated phase of the optimised device. As shown in Figure 7.19(c), which reports the relative phase error in percent, the phase shift is almost identical over most of the 70 μm square domain, while a maximum deviation of about 14% appears near the tip of the main needle. This discrepancy arises from the finite tip thickness, which is unavoidable in fabrication; it could be further reduced by optimising the tip shape. Figure 7.20 shows the final device layout, including the full chip geometry and electrical connections ready for fabrication.

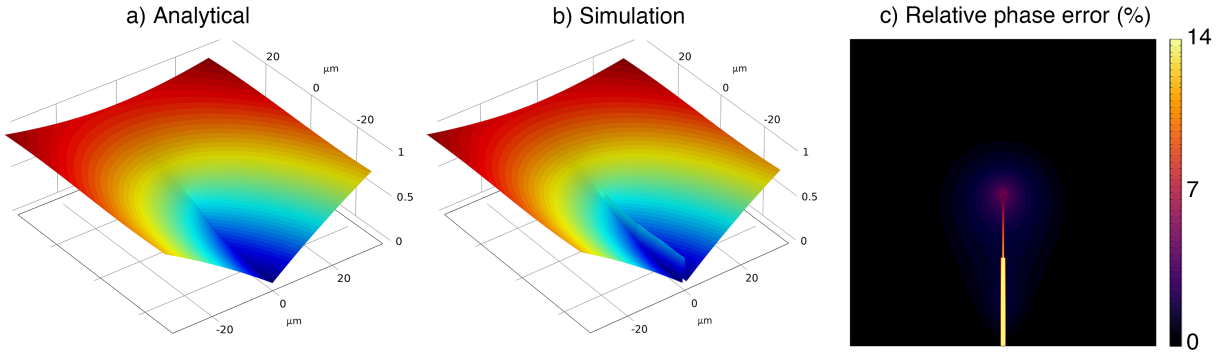


Figure 7.19. Comparison between (a) the analytical target and (b) the simulated phase of the OAM sorter phase plate, together with (c) the relative phase error between simulation and target (percentage).

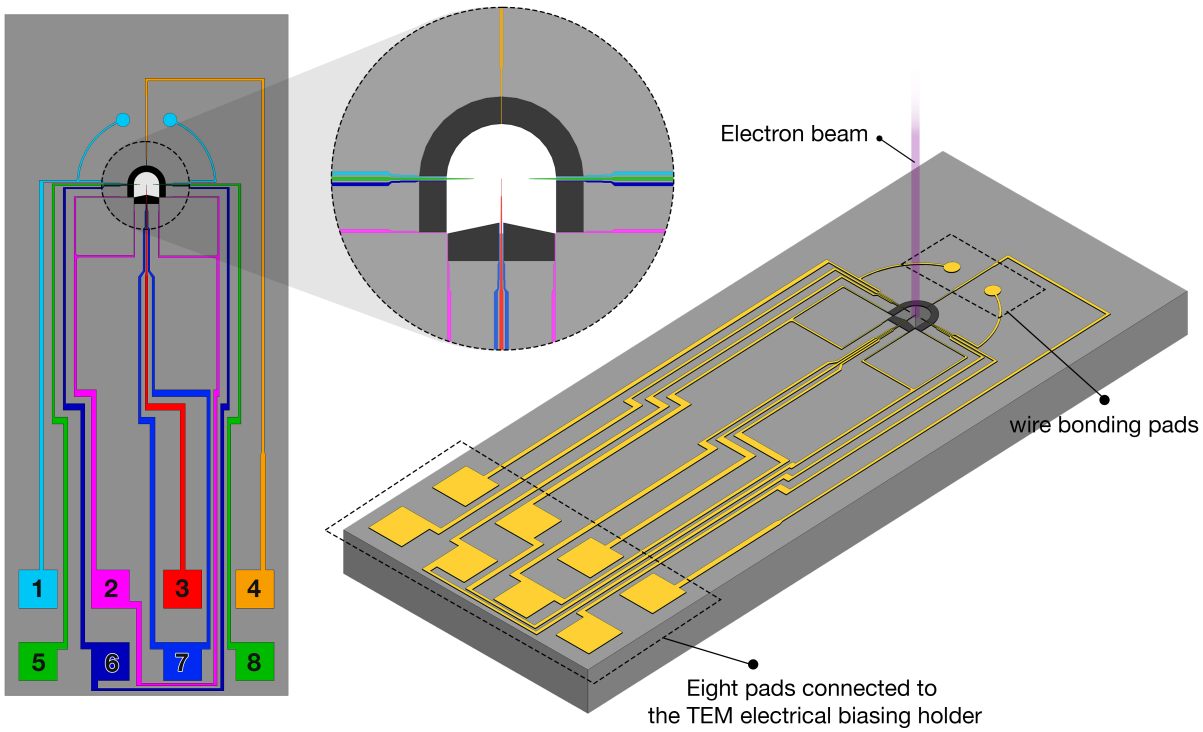


Figure 7.20. Current-controlled electrostatic OAM sorter design and chip layout.

Part III

Experimental results

8 Fabrication Process

Contents

8.1 Material and Substrate Selection 142
8.2 Lithography and Pattern Transfer 143
8.3 Device Fabrication Results 146
8.4 Electrical Characterization and Breakdown Testing 148

8.1. Material and Substrate Selection

One of the key advantages of silicon in electronic devices is that its electrical conductivity can be tuned in a controlled and predictable way by adjusting the impurity concentration. Intrinsic (undoped) silicon contains very few free carriers and therefore exhibits high resistivity, making it unsuitable for carrying significant current. To overcome this, impurity atoms are intentionally introduced into the lattice in a process known as *doping* [164].

Group III atoms—most commonly boron, but also aluminum or gallium—act as acceptors and generate p-type semiconductors, while Group V atoms, such as phosphorus, arsenic, or antimony, act as donors and create n-type semiconductors. These impurity atoms introduce shallow energy levels in the bandgap that ionize easily at room temperature, increasing the free carrier density by several orders of magnitude [164].

A more precise description of how doping affects conductivity is obtained from the drift current density,

$$J_{\text{drift}} = q\left(\mu_n n + \mu_p p\right) E, \tag{8.1}$$

which can be written in the form

$$J_{\text{drift}} = \sigma E, \tag{8.2}$$

where σ is the conductivity. The conductivity of a semiconductor is therefore

$$\sigma = q\left(\mu_n n + \mu_p p\right), \tag{8.3}$$

and the corresponding resistivity is the reciprocal,

$$\rho = \frac{1}{\sigma} = \frac{1}{q\left(\mu_n n + \mu_p p\right)}. \tag{8.4}$$

Here q is the elementary charge, n and p are the electron and hole concentrations, and μ_n and μ_p are their mobilities, respectively [164].

Equations (8.3)–(8.4) show that conductivity depends not only on the free carrier concentrations, but also on their mobilities, which themselves are functions of the impurity concentration. As the doping level increases, the carrier concentration rises while the mobility gradually decreases due to enhanced scattering from ionized impurities. The net result is a nonlinear dependence of σ (or ρ) on the donor or acceptor doping level [164].

In current-controlled electrostatic phase plates, the silicon wafer must balance electrical conduction and Joule heating. When a current I flows through a resistance R, the dissipated

power $P = I^2R$ heats the device and can change its electrical properties. A moderate temperature rise can help reduce contamination, but excessive heating risks damaging or breaking the structure at high voltage. Very low-resistivity (heavily doped) silicon reduces R, but for a given applied voltage, it allows large currents and thus strong Joule heating. In addition, large currents in the active area generate magnetic fields that introduce an extra phase-shift artifact in the electron beam. Intrinsic or lightly doped silicon, in contrast, has high resistivity and tends to form wide depletion regions and Schottky barriers at the metal–silicon interface, which limit current injection and make the contact behavior difficult to control [164]. For this reason, in current-controlled phase plates, we deliberately use moderately doped silicon: the resistivity is low enough that the contacts behave close to Ohmic, but high enough to limit I and keep the Joule heating and resulting temperature rise within a safe range. For purely electrostatic phase plates, where a thin gold layer covers the electrodes and only small leakage currents are expected, it is generally advantageous to use highly doped, highly conductive silicon to minimize voltage drops through the electrodes.
In this work, we used different types of four inch silicon-on-insulator (SOI) wafers with different doping levels in the device and handle layers to investigate the electrical behavior of the phase-plate. The main wafer parameters are summarized in Table 8.1.

Table 8.1. SOI wafers used in this work.

	Device layer			**BOX**	**Handle layer**		
	Thickness (µm)	Resistivity (Ω cm)	Orient	Thickness (µm)	Thickness (µm)	Resistivity (Ω cm)	Orient
S1	10	0.003–0.005 (P)	⟨100⟩	1	400	0.003–0.005 (P)	⟨100⟩
S2	11	3–5 (P)	⟨100⟩	4	400	0.005–0.1 (Sb)	⟨100⟩
S3	20	0.3–0.5 (B)	⟨100⟩	0.5	400	0.3–0.5 (B)	⟨100⟩

8.2. Lithography and Pattern Transfer

The chip layouts were designed using computer-aided design (CAD) software and exported as a GDSII (`.gds`) file, the standard format used in nanofabrication. This format is compatible with most lithography systems, ensuring that the electrode geometries and alignment features are transferred accurately to the photomasks used in the fabrication process. Figure 8.1 shows the complete process flow for the device fabrication, which consists of twelve main steps:

1. **Starting wafer.** The process begins with a 4-inch silicon-on-insulator (SOI) wafer.

2. **Oxide deposition.** A thin layer of silicon dioxide is deposited on both sides of the wafer using low-pressure chemical vapor deposition (LPCVD) from SiH_4 and O_2 at 180 mTorr and 420 °C. This oxide acts as the hard mask for subsequent etching steps. The thickness is 600

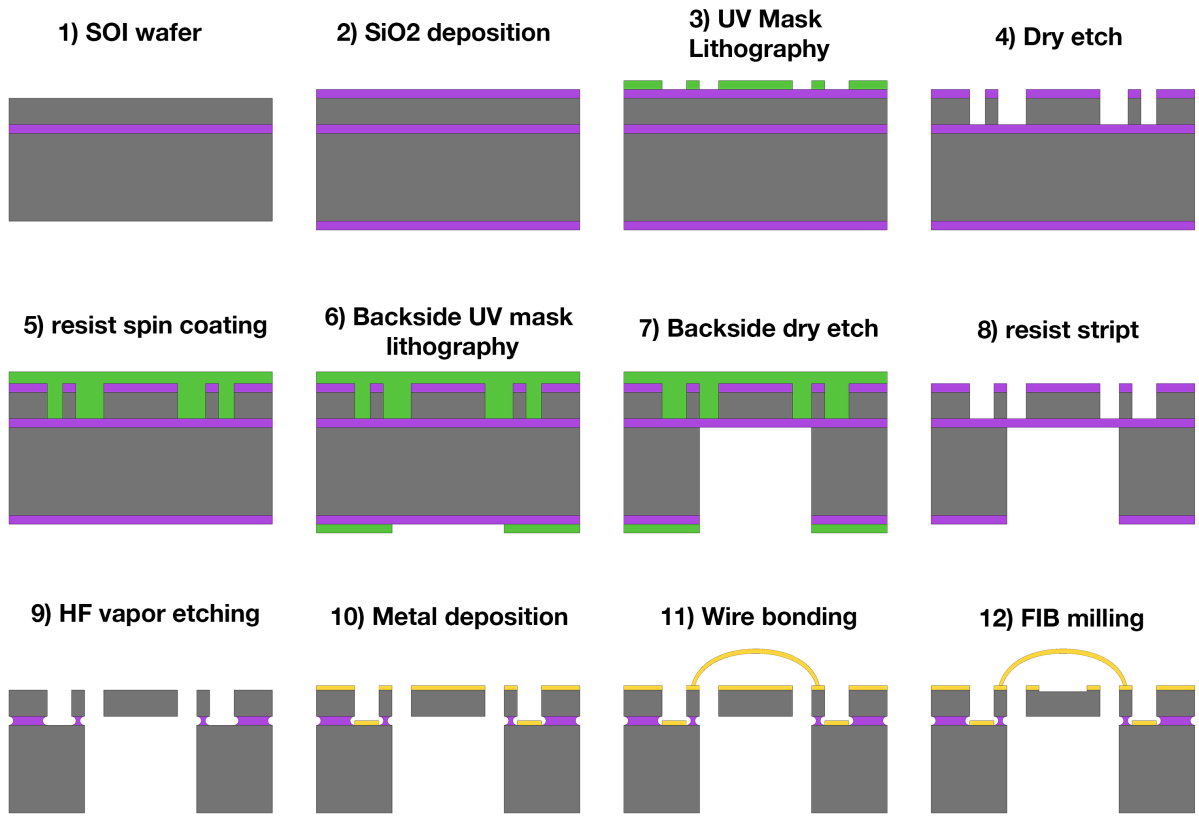


Figure 8.1. Process flow for chips fabrication

nm on the front side and 2 µm on the back side.

3. **Front side lithography.** A positive deep-UV 2.2 μm thick photoresist (UV26 from Fujifilm) is spin-coated onto the front surface, patterned using UV contact lithography, and developed to transfer the first pattern onto the resist.

4. **Front side etching.** The oxide and the device silicon layer are etched using dry plasma etching. Silicon oxide etching is performed first by C_4F_8 inductively coupled plasma (ICP) etching. The resist is not stripped after the SiO_2 patterning, and the subsequent silicon etching is carried out in a self-aligned fashion using the same resist pattern. Silicon etching is then performed using a Bosch deep reactive ion etching (DRIE) process, based on alternating SF_6 and C_4F_8 ICP steps, to ensure vertical sidewalls and maintain high aspect ratio features. The process used is a low-scallop one, with an etching rate of around 3 μm/min on large features. Both etching steps are carried out in the same tool (Oxford Cobra). The oxide layer (BOX layer) serves as the etch-stop.

5. **Front side protection.** A thick layer of photoresist (AZ resist from Fujifilm, 10 μm thick after spinning) is deposited on the front side to protect the patterned structures during backside processing.

6. **Back side lithography.** A contact lithography procedure is also performed on the backside of the wafer, including resist spin coating (4 μm thick near-UV resist OiR from

Fujifilm), UV mask lithography, and development.

7. **Back side etching.** The backside oxide is etched using dry etching, followed by DRIE Bosch etching of the handle layer of the wafer. Again, the BOX layer acts as an etch stop. The same processes described in step 4 are used, but this time a medium-scallop DRIE Bosch process is adopted, providing an etch rate of around 5 μm/min on large features.

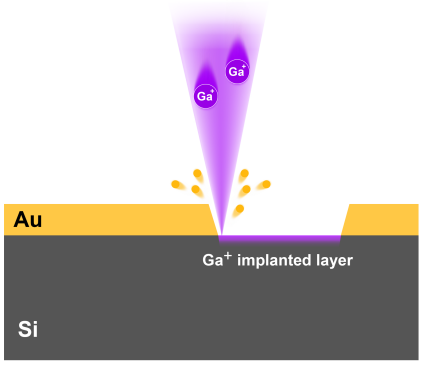
8. **Resist removal.** All remaining photoresist is stripped from both sides of the wafer by O_2 plasma ashing.

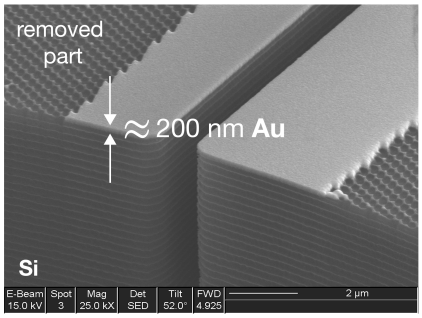
9. **BOX removal.** The BOX layer is removed using HF vapor etching, performed by suspending the sample over a cold 47% wt. aqueous HF solution and heating the sample to 38 °C for about one hour. This process is isotropic, resulting in a lateral under-etch of the oxide beneath the device silicon.

10. **Dicing and metallization.** After dicing the wafer into individual chips, a thin layer of gold is deposited on the entire chip using thermal evaporation. This technique provides highly directional coating, allowing metal deposition primarily on horizontal surfaces while leaving the vertical electrode sidewalls largely uncoated. This directional deposition is crucial for preventing electrical short circuits between neighboring electrodes, ensuring that only the intended conductive paths are formed.

11. **Wire bonding.** Electrical connections between selected pads are made using wire bonding. We used 50 μm aluminum wire, a standard choice for ultrasonic wire bonding, which was the bonding technique employed here.

12. **FIB milling (current-controlled design only).** For the current-controlled electrostatic phase plate, the gold layer remaining on the active electrode surfaces must be removed. This is achieved using focused ion beam (FIB) milling with Ga^+ ions. Although Xe-based plasma FIB avoids ion implantation, the Ga^+-based process introduces only a very thin (∼20 nm) implanted layer [165]. The gallium concentration introduced by FIB milling is significantly lower than that obtained through standard metal deposition methods, resulting in minimal change to the electrode conductivity. This maintains the required conductivity contrast between the doped silicon electrodes and the gold-coated connections.





8.3. Device Fabrication Results

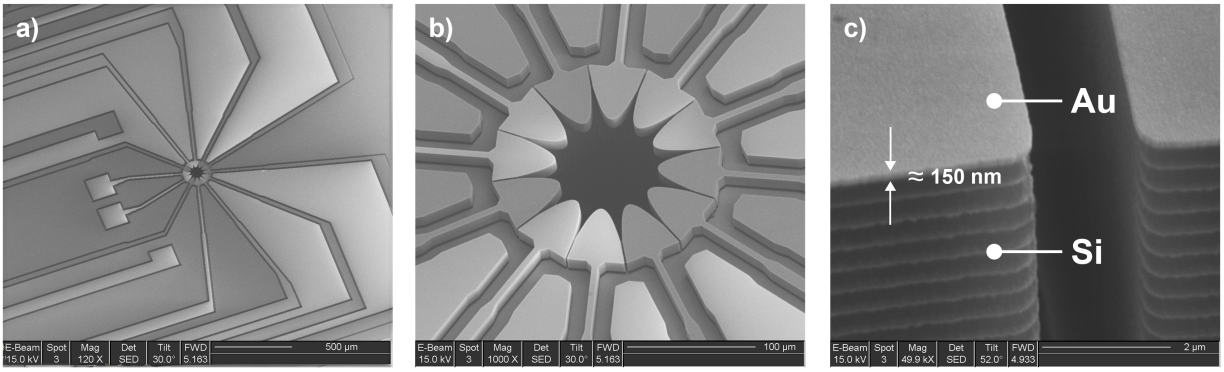


Figure 8.2. SEM images of the dodecapole hyperbola phase-plate device. (a) Overview of the fabricated device showing the four contact pads used for wire bonding. (b) Active area of the device. (c) Zoom-in on the electrode edge, highlighting the $\sim 150\,\mathrm{nm}$ Au layer and the step-like sidewall profile produced by the Bosch DRIE process.

Figure 8.2 shows SEM images of the fabricated dodecapole electrostatic phase plate (Hyperbola phase plate). The device is fabricated on wafer S1, with a $10\,\mu\mathrm{m}$-thick device layer (see Table 8.1), which provides low resistivity suitable for electrostatic phase plates. An approximately $150\,\mathrm{nm}$-thick gold layer is deposited over the entire device. As shown in Figure 8.2(c), the Boersch process yields nearly vertical sidewalls, allowing for very small gaps between adjacent electrodes, which is crucial for the phase-plate design.

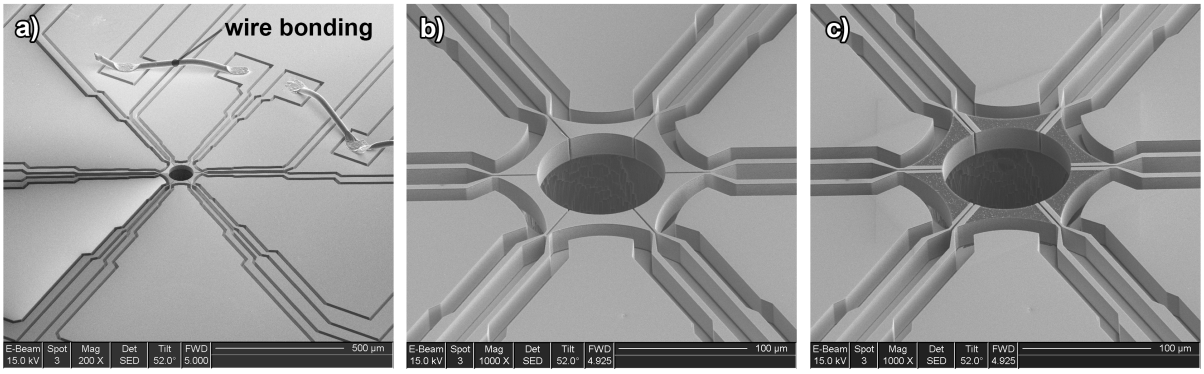


Figure 8.3. SEM images of the Helios phase-plate device. (a) Overview of the fabricated chip after wire bonding. (b) Active area of the device before FIB milling. (c) Active area of the device after FIB milling.

Figure 8.3 presents SEM images of the hexapole current-controlled electrostatic phase plate (Helios phase plate). As shown in Figure 8.3(a), the bonding pads are connected to the chip by $50\,\mu\mathrm{m}$ Al wires using ultrasonic wire bonding. Figure 8.3(b) shows the active area of the device coated with a $200\,\mathrm{nm}$ Au layer, and Figure 8.3(c) shows the same area after FIB milling, where selected regions of the Au layer have been removed. The device is fabricated on wafer S3 with a $20\,\mu\mathrm{m}$-thick device layer.

The third class of phase plates reported in this thesis consists of vortex beam generators and OAM sorters. Figure 8.4 shows SEM images of these devices. Figure 8.4(c) highlights the fabrication quality, displaying nearly vertical sidewalls on one of the OAM sorter needles with an approximate width of 1 µm.

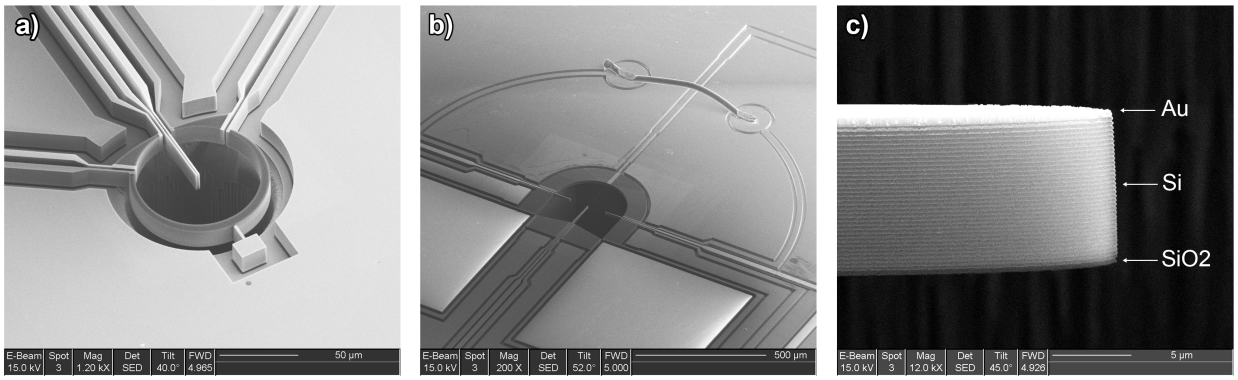


Figure 8.4. SEM images of vortex and OAM sorter phase plates: (a) vortex phase plate with Au-removed region; (b) OAM sorter with three needles; (c) close-up of one needle showing the fabrication quality and nearly vertical sidewalls.

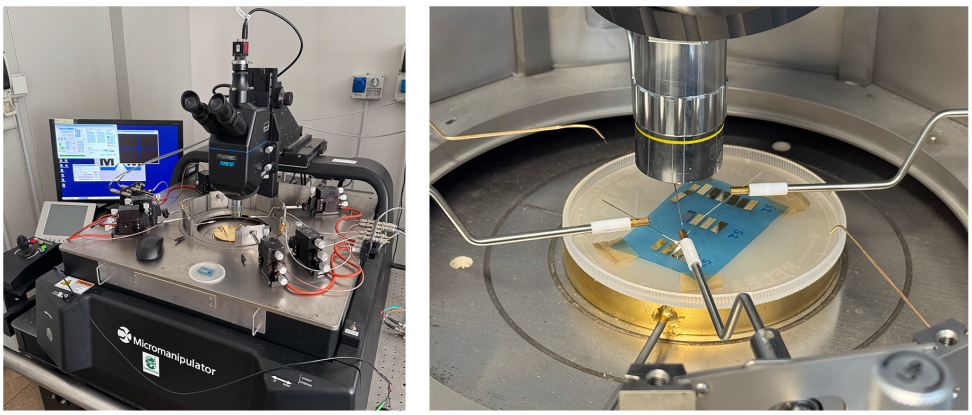


Figure 8.5. Electrical characterization using a probe station.

8.4. Electrical Characterization and Breakdown Testing

One of the crucial steps in the fabrication of silicon-based electronic devices is electrical characterization. For this purpose, two main procedures are required:

1. **Short-circuit inspection.** This test ensures that no unintended electrical connections exist between pads. We performed the measurements using a P300A semi-automatic probe station (Micromanipulator) at the CNR-ISMN facility in Bologna (Figure 8.5) , which provides high-precision electrical testing with nanoampere sensitivity. During this inspection, each pad is measured against all other non-connected pads and against ground regions to confirm the absence of short circuits. This procedure is necessary for both the electrostatic and current-controlled designs.

2. **Resistivity measurement.** For our work, resistivity verification is essential, especially for the current-controlled electrostatic phase plate. Measurements must be performed in two stages: after gold metallization and after FIB milling. This ensures that the gold removal produces the required conductivity contrast between the doped silicon electrodes and the gold-coated connections. These measurements are also important for determining the experimental material parameters used in the COMSOL simulations, enabling accurate modeling of the device behavior.

Electrical characterization revealed variations between devices, mainly due to fluctuations in the thin-film Au deposition, uncertainties in the silicon characterization, probe-tip contamination, and the complexity of the routing and wire-bond layout. For this reason, I do not report a complete set of numerical values here, but instead provide a representative example from the chips used in the experiments. For the Helios phase plate fabricated on wafer S3 with a 200 nm Au layer, the measured resistance between pads 3 and 7 (see Figure 7.7) increased from an average of 254 Ω (over three identical chips) before FIB milling to 3.1 kΩ after the removal of Au from the electrodes. This increase confirms that selective Au removal raises the electrode resistance.

9 Experimental result

Contents

9.1 Phase Plate Mounting and Electrical Connections 150

9.1.1 Mounting the MEMS device on the biasing holder 151

9.1.2 External wiring and voltage control 152

9.1.3 Insertion and in-column alignment workflow 152

9.2 Experimental result : Hyperbola and Helios phase plates 153

9.3 Phase Retrieval via Transport of Intensity 158

9.1. Phase Plate Mounting and Electrical Connections

In this section, the experimental integration of the MEMS-based electrostatic phase plate into the (S)TEM is described, including mechanical mounting, electrical interfacing, microscope insertion, and alignment procedures used to achieve reproducible phase modulation.
Most experiments were performed on a Thermo Fisher Scientific (S)TEM Spectra 300 (SPE-QTEM) equipped with an X-FEG source operating from 30 to 300 kV. When operated with the monochromator, the instrument can achieve an energy resolution of tens of meV. A schematic overview of the instrument configuration is shown in Figure 9.1.

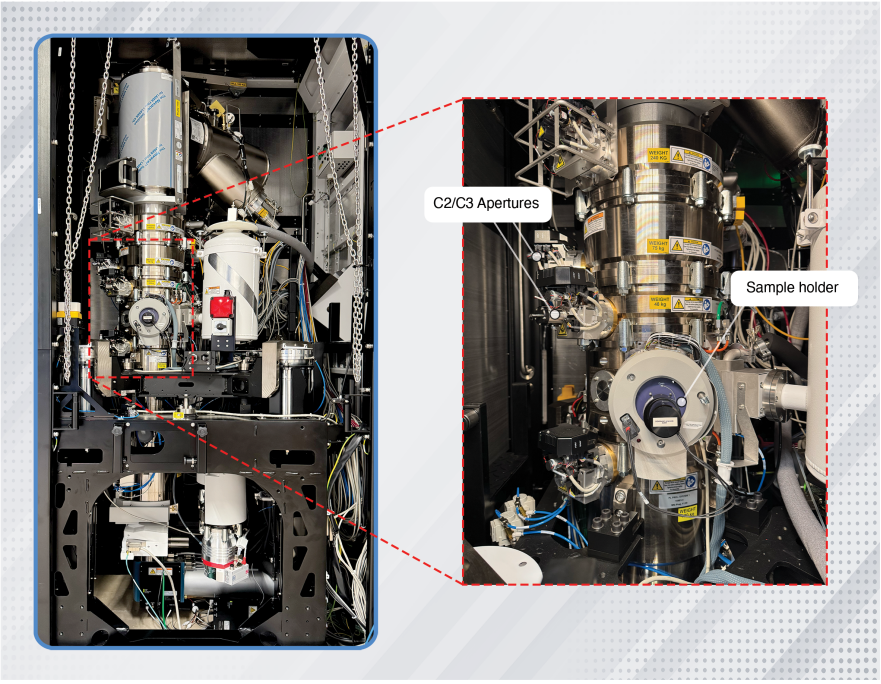


Figure 9.1. A schematic overview of the Thermo Fisher Scientific (S)TEM Spectra 300 (SPE-QTEM) microscope

The Spectra 300 condenser system includes three condenser lenses upstream of the specimen, and the column provides aperture planes that can host dedicated in-column devices. This enables two distinct experimental configurations for phase modulation: (i) insertion of the phase plate in the condenser aperture system (pre-sample phase manipulation of the illumination), and (ii) insertion of the phase plate in the specimen holder (device positioned at the specimen level). In the first configuration, the phase plate acts on the illumination wavefront before it forms the probe at the specimen; in the second configuration, the device operates

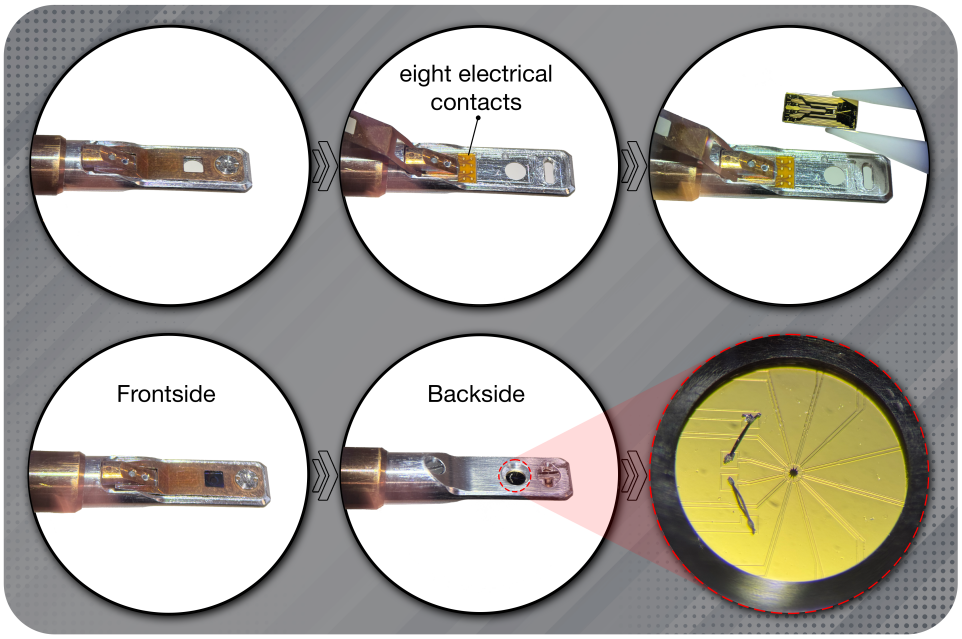


Figure 9.2. Mounting MEMS on the Thermo Fisher Scientific NanoEx-i/v TEM holder

as a specimen-mounted electrostatic element. Although the present device is intended to compensate for threefold astigmatism by acting on the illumination in the condenser system (i.e., to correct the STEM probe), the experiments reported in this thesis were performed with the MEMS inserted at the specimen position. This configuration allowed us to characterize the device response in both imaging mode and diffraction mode, where the Fourier transform of the shaped beam can be recorded and directly compared with simulation results. In addition, using the sample position avoids venting or reconfiguring the condenser-aperture hardware, enabling fast and safe exchange of different phase-plate chips without breaking the microscope vacuum.

9.1.1. Mounting the MEMS device on the biasing holder

To mount the MEMS device in the microscope, a Thermo Fisher Scientific NanoEx-i/v TEM holder for electrical biasing (and heating capability) was used [166]. The complete mounting sequence is illustrated in Figure 9.2. The holder provides eight electrical contacts routed through an 8-pin vacuum-compatible connector to the external power supply.

The MEMS chip was mounted face-down, with the electron beam passing through the device from the backside. The holder geometry requires this orientation. After placing the chip onto the insulated holder seat, a copper clamping cap was installed to enclose and mechanically secure the device; the cap was tightened using the dedicated screw interface. Due to the holder's encapsulated design, the wire-bonding geometry is constrained. As shown in Figure 9.2(c), only the circular opening (approximately 2 mm in diameter) provides sufficient out-of-plane access for wire bonding on the MEMS pads. Therefore, the chip layout used for the experiments was designed such that all wire bonding pads are located within the

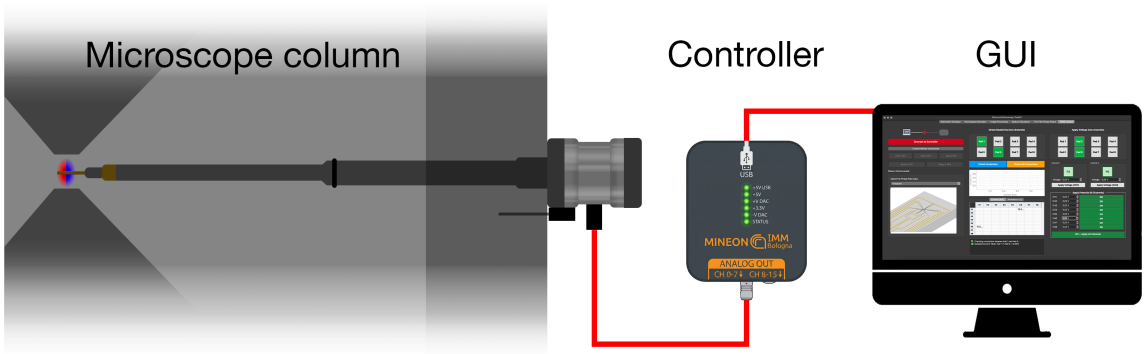


Figure 9.3. Holder connections: the holder is connected to a controller via an 8-pin connector, and the controller is connected to the computer via USB. The custom GUI controls the controller during the experiments.

accessible region.

Before microscope insertion, the device was verified outside the microscope by (i) optical inspection and (ii) resistance/leakage checks between biased electrodes and ground using a Keithley instrument. These steps are essential to prevent in-column failures.

9.1.2. External wiring and voltage control

The holder was connected to a dedicated multi-channel controller (Figure 9.3) through the 8-pin connector. The controller, developed in our group, provides up to 16 independently addressable output channels for biasing the device. In USB-powered mode, the output range is -10 to $+10$ V; with an external power supply, it extends to -20 to $+20$ V (within the board's compliance limits). The controller is interfaced via a Python library ("mineon"), which enables scripted voltage sequences, logging, and repeatable experimental protocols. For convenient operation, a graphical user interface (GUI) was used as a front-end to the same control layer, enabling real-time adjustments while maintaining traceability of the applied voltages.

9.1.3. Insertion and in-column alignment workflow

After mounting and electrical verification, the holder was inserted into the microscope following the standard load-lock procedure. Once vacuum stability was reached, a baseline dataset was acquired at 0 V on all electrodes to establish the reference probe/diffraction signature and to check for charging or contamination. The phase plate was then aligned relative to the illumination by centering the beam within the device's active region and by optimizing condenser and aperture alignment to ensure that the phase plate acts on the intended portion of the wavefront. Voltage-dependent measurements were performed by applying a known bias configuration, recording the resulting changes in the diffraction pattern and/or probe shape, and confirming reversibility by returning to the 0 V state. This baseline–bias–return sequence was repeated to verify stability and to identify hysteresis, drift, or unstable electrical contact.

9.2. Experimental result : Hyperbola and Helios phase plates

In this section, we report two phase plate implementations introduced in Chapter 7: the electrostatic **Hyperbola** phase plate and the current-controlled **Helios** phase plate. Both devices are designed to compensate three-fold astigmatism (A_3) and provide two key functionalities: (i) *tunable* correction strength through the applied electrode biases, and (ii) *rotatable* correction by programming the orientation of the bias pattern. For the measurements reported here, the MEMS chips were operated at the specimen plane. This configuration allows for characterization in both imaging conditions and in the objective back focal plane (BFP), where the diffraction pattern of the shaped beam can be recorded and directly compared with simulated predictions. The electrodes were driven using the optimized bias sets obtained from the numerical workflow in Chapter 7 (see Figurs 7.5 and 7.10).

The first set of results concerns the Hyperbola phase plate. Figure 9.4 compares the experimental diffraction patterns obtained for five optimized orientations of the A_3 correction axis, spanning a relative rotation from 0° to 30°. Intermediate orientations can be achieved by applying the corresponding optimized electrode potentials, demonstrating continuous angular control rather than discrete operation. By reversing the polarity of the applied potentials and using the corresponding reversed bias sequence, the full 0°–360° angular range is accessible.

Figure 9.5 demonstrates three-fold astigmatism compensation using a hyperbola phase plate. The 3-fold astigmatism is naturally present in the diffraction lens of the microscope and cannot be compensated in standard operations. We aim to evaluate it and compensate through the device located in the MEMS position. Direct evaluation using an in-focus diffraction pattern is experimentally challenging because the relevant features are tightly localized and highly sensitive to small alignment changes. Moreover, our neural-network-based aberration predictor is designed to operate on defocused probe images. Therefore, to robustly visualize the correction and enable AI-assisted alignment, the compensation procedure was performed under controlled defocus values of 200 nm and 400 nm. For each defocus condition, Figure 9.5 shows the following sequence: phase plate off (all electrodes connected to ground), AI-based estimation of the microscope A_3 orientation, selection of the corresponding phase plate bias rotation, and compensation by applying the optimized bias set (with $V_{\max} = 0.35\,\mathrm{V}$, which corresponds in simulation to an approximately 40 rad phase modulation at a beam radius of 30 μm). In both defocus conditions, the probe becomes significantly more circular and symmetric after compensation, confirming successful threefold astigmatism correction. The maximum applied potential required for this compensation is only $0.35V$, highlighting the low-voltage operation of the device.

The second set of experimental results concerns the Helios phase plate. Figure 9.6 compares the simulated and experimental diffraction patterns for five orientations between 0° and 30°. The white ghost feature visible in both datasets arises from inserting a $60\mu m$ objective aperture into the column to reduce device charging. Figure 9.7 demonstrates threefold-astigmatism compensation at $V_{\max} = 0.3\,\mathrm{V}$ in an in-focus diffraction pattern, showing a clear reduction of aberration-related parasitic distortions.

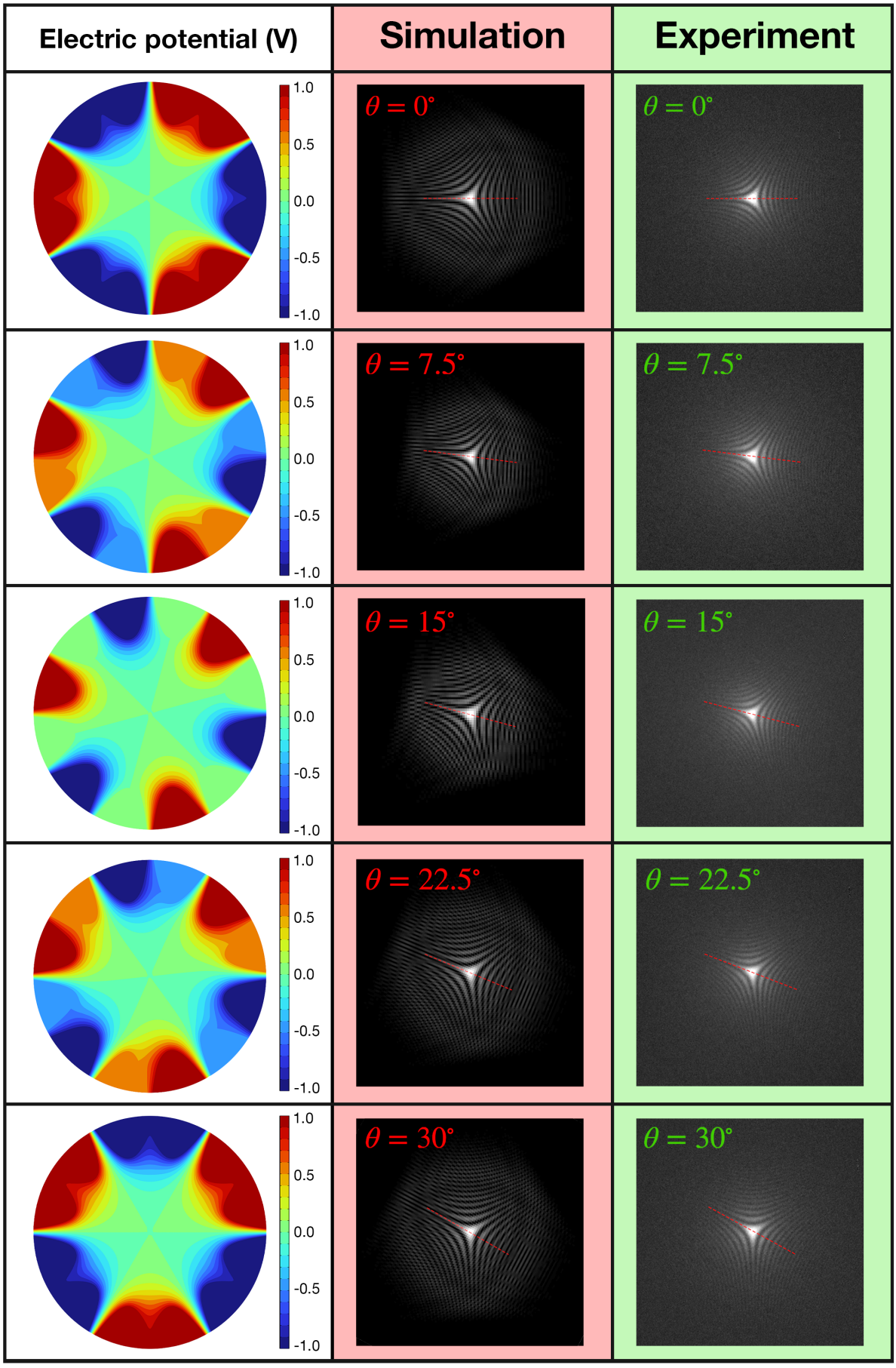


Figure 9.4. Hyperbola phase plate: comparison between experimental and simulated diffraction patterns for five orientations of the three-fold phase between 0° and 30°, demonstrating controlled rotation of the phase profile.

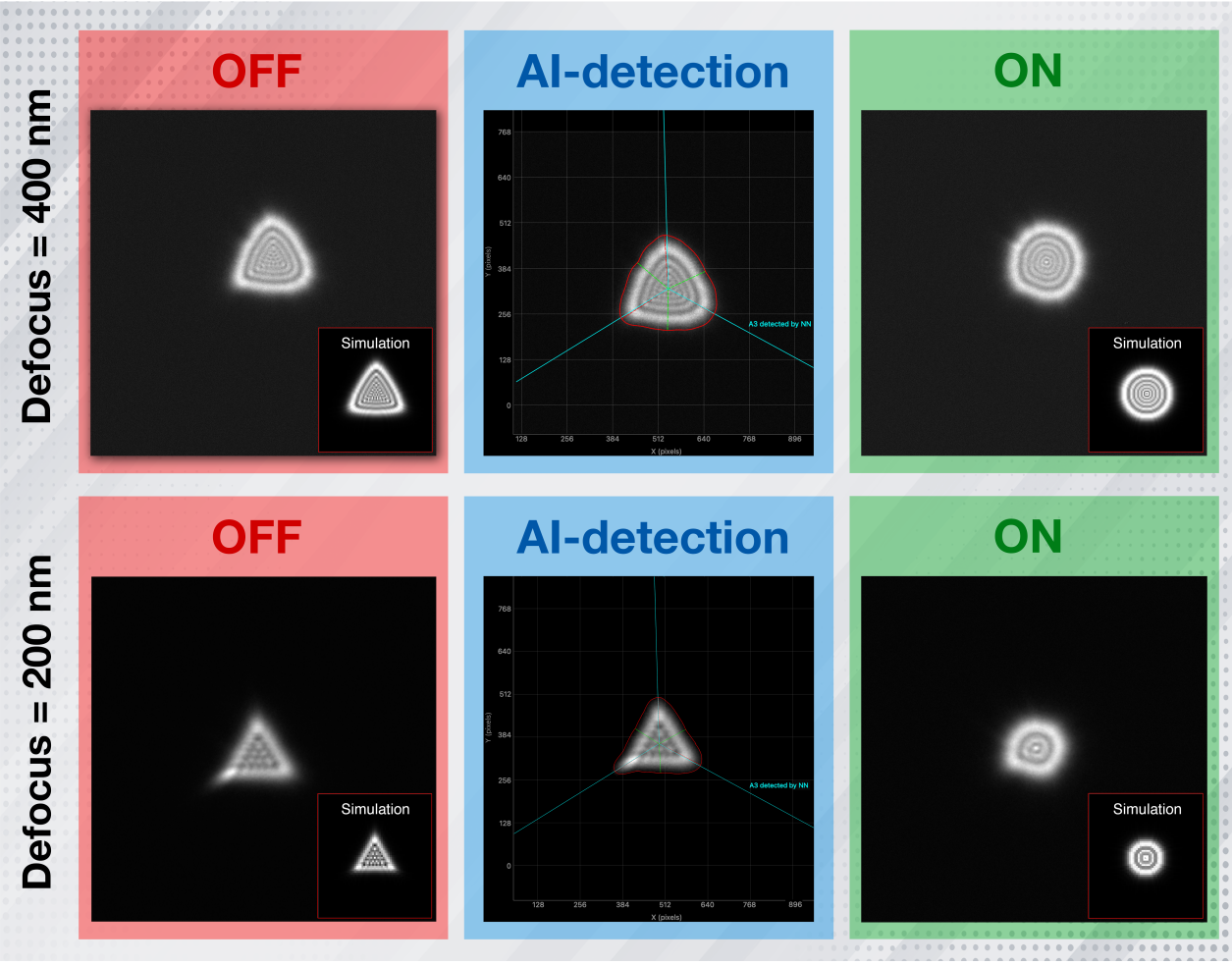


Figure 9.5. Compensation of the microscope threefold astigmatism using the Hyperbola phase plate for defocused diffraction patterns at 200 nm and 400 nm. In the "OFF" state, the phase plate is grounded. The middle column shows the orientation estimated by the developed deep neural network, which selects the optimized electrode potentials and automatically aligns the phase plate to the microscope threefold orientation. In the "ON" state, the optimized potentials are applied via the GUI.

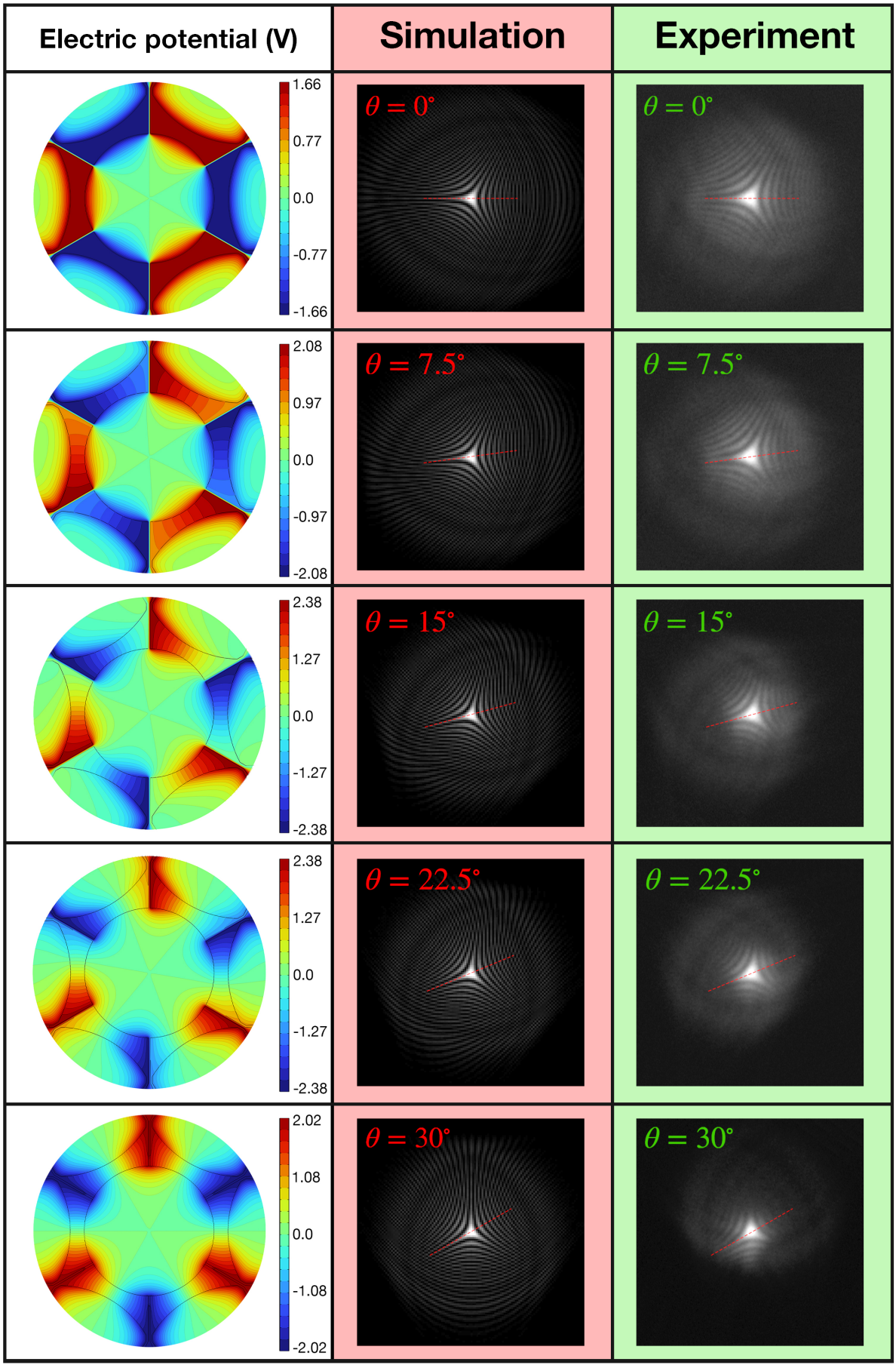


Figure 9.6. Helios phase plate: comparison between experimental and simulated diffraction patterns for five orientations of the three-fold phase between 0° and 30°, demonstrating controlled rotation of the phase profile.

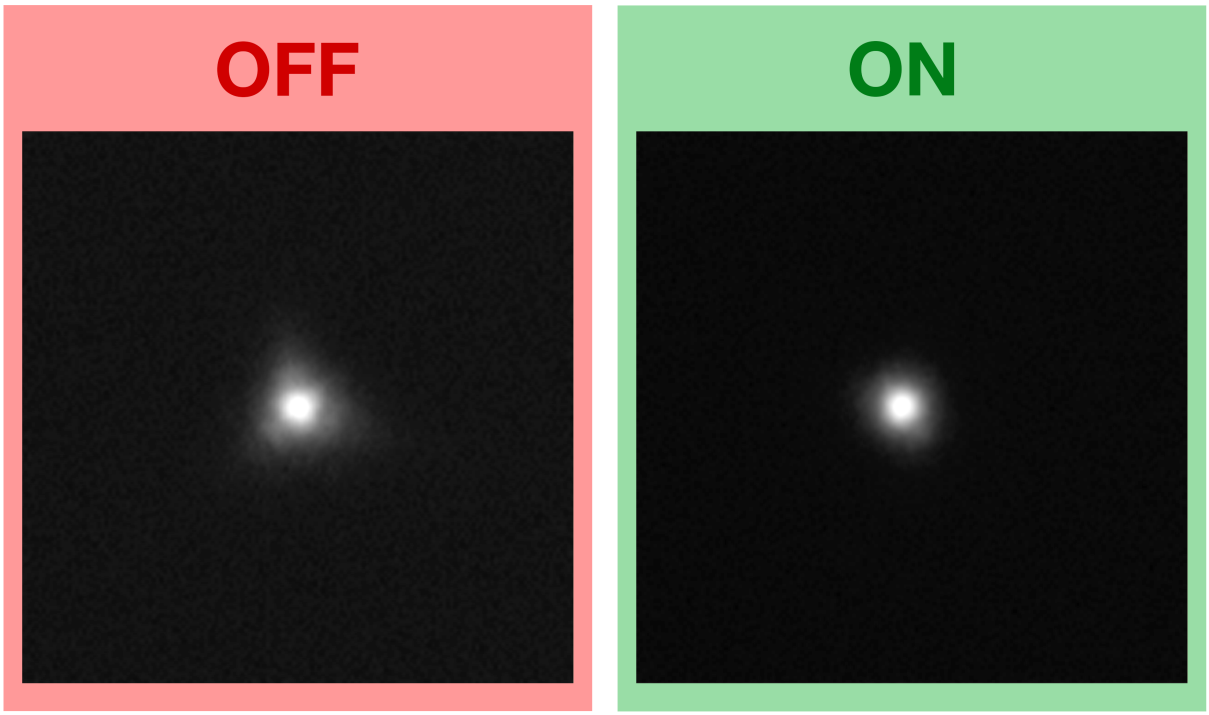


Figure 9.7. Threefold-astigmatism compensation in the in-focus diffraction pattern using the Helios phase plate shows a reduction of parasitic aberrations when the phase plate is biased with the optimized electrode potentials.

9.3. Phase Retrieval via Transport of Intensity

As discussed in the previous chapter, the current-controlled electrostatic phase plate developed in this work enables the generation of well-defined electric potential gradients across the electrode structure. Although the operating principle of the Helios phase plate was introduced earlier, the current-controlled concept must still be experimentally validated by directly assessing the phase imprinted on the electron wavefront. In a transmission electron microscope, however, the detector records only the beam intensity, while the phase information is not directly accessible in a conventional image. Therefore, dedicated phase-retrieval techniques are required to quantitatively evaluate the phase plate performance and to verify the intended wavefront modulation.

Several methods are available for phase retrieval in electron microscopy. Off-axis electron holography is a powerful and well-established technique, but it requires a coherent reference wave and a biprism-based interferometric setup, which can be experimentally demanding. An alternative, non-interferometric approach is provided by the Transport of Intensity Equation (TIE). This method reconstructs the phase by relating the axial derivative of the image intensity to the transverse phase gradients of the electron wave within the paraxial approximation. In practice, the axial intensity derivative is obtained from two symmetrically defocused images acquired at Δz, and the phase is then recovered by solving a Poisson-type equation subject to appropriate boundary conditions. Under the paraxial approximation [167]:

$$\nabla \cdot [I(\mathbf{r}, 0)\, \nabla \varphi(\mathbf{r})] = -\frac{2\pi}{\lambda} \left. \frac{\partial I(\mathbf{r}, z)}{\partial z} \right|_{z=0}, \tag{9.1}$$

showing that the measurable axial intensity derivative provides the source term for reconstructing the phase $\varphi(\mathbf{r})$.

In our implementation, this axial intensity-derivative term is obtained without changing the objective-lens defocus. Instead, two images are recorded under identical microscope defocus while applying equal-magnitude voltages with opposite polarity to the phase plate, $+V$ and $-V$. Reversing the phase-plate bias flips the sign of the phase modulation, which, in the small-defocus regime, produces the same intensity response as reversing the defocus. Thus, the two bias states provide the intensity variation required by TIE without changing the microscope focus [167]. The resulting intensity difference is then used as an effective estimate of the derivative term in Eq. 9.1, enabling phase reconstruction while keeping all lens excitations fixed.

Figure 9.8 shows the three images used for phase retrieval. The in-focus image is acquired with the phase plate in the OFF state, where all electrodes are connected to ground. The other two images are acquired at a defocus of 5 mm while applying equal-magnitude voltages with opposite polarity, $+4\,\mathrm{V}$ and $-4\,\mathrm{V}$. The corresponding electrode potential distribution is shown for each biased state.

Figure 9.9 shows the TIE phase reconstruction of the Helios phase plate for two bias configurations. In Configuration 1, opposite-polarity biases are applied to the electrode terminals,

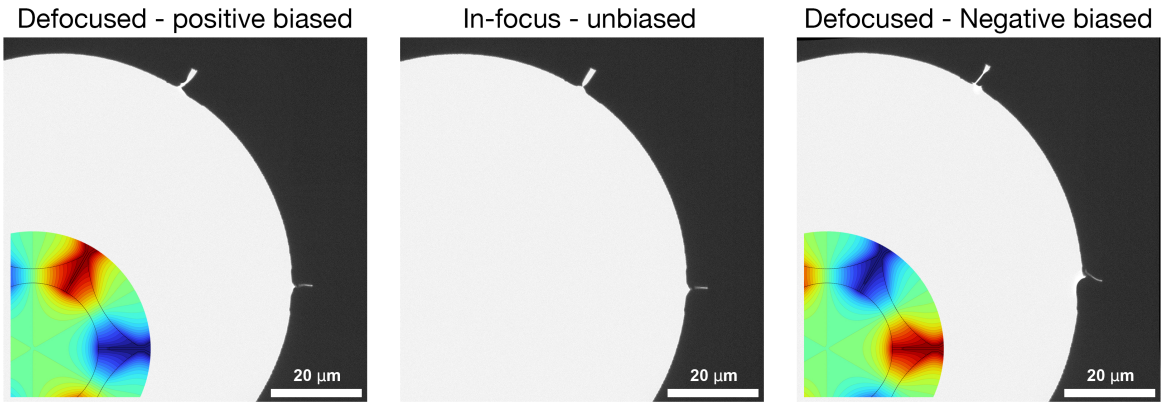


Figure 9.8. Three images used for TIE reconstruction: one in-focus image with the phase plate grounded, and two images recorded at $\Delta f = 5\,\mathrm{mm}$ with equal-magnitude voltages of opposite polarity.

which drive a current and generate a controlled potential gradient along the boundary electrode. The analysis region is extracted from the dashed square in the first panel. The middle panel shows the intensity difference, $\Delta I = I_+ - I_-$, between the two opposite-bias images. The reconstructed phase map exhibits the expected phase gradient, consistent with the simulations of the current-controlled phase plate. A localized artifact is also visible (red arrow), which breaks the expected symmetry and introduces an additional phase contribution.
Configuration 2 corresponds to the electrostatic mode. Here, both ends of the electrode are connected to the same bias, so no current flows and the electrode remains at an approximately uniform potential. The reconstructed phase exhibits an approximately 30° rotation relative to Configuration 1.

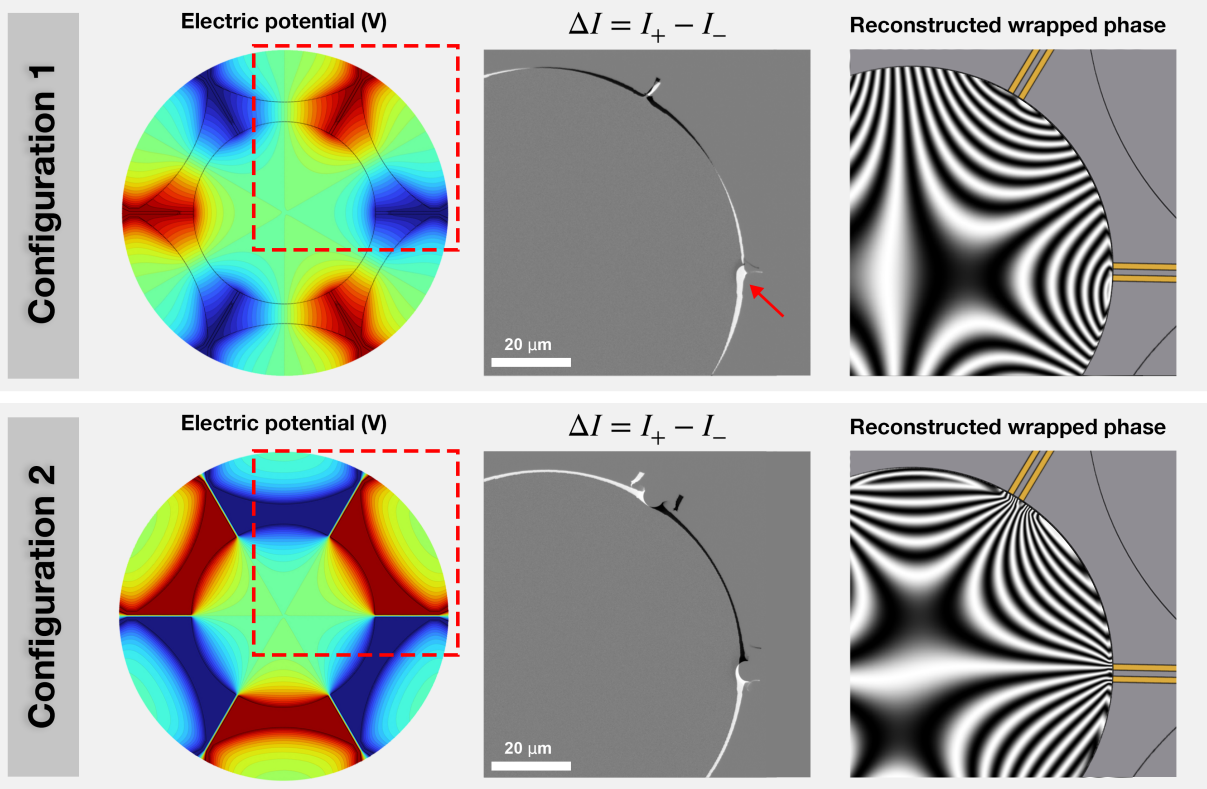


Figure 9.9. Phase reconstruction for two operating configurations: (1) current-controlled mode (electric current flows through the phase plate), and (2) electrostatic mode (no current flow).

Part IV

Discussion and Outlook

10 Conclusions and Future Outlook

Contents

10.1 Summary of Contributions and Future Directions 163

10.1. Summary of Contributions and Future Directions

This thesis demonstrated that MEMS-based electrostatic phase plates provide a practical and powerful platform for controlled electron-beam shaping in transmission electron microscopy. An end-to-end methodology—device design, multiphysics simulation, numerical optimization, and cleanroom fabrication—was implemented across the reported phase-plate architectures. Experimental validation is presented for two tunable multipole phase plates for aberration compensation: the Hyperbola electrostatic hexapole phase plate and the Helios current-controlled multipole phase plate. The results show that strong, well-defined phase modulation can be achieved with low voltages by placing electrodes in close proximity to the electron beam. The fabricated devices exhibited stable electrical behavior, reproducible phase response, and good agreement with electrostatic and wave-optical simulations. Experimental diffraction measurements confirmed controllable rotation, tunability, and symmetry of the induced phase profiles, demonstrating effective compensation of threefold astigmatism and flexible phase control within the sample plane. Importantly, the current-controlled designs introduced here open a new design strategy for enhancing the performance of electrostatic phase plates by enabling more robust and uniform phase control through engineered current paths and resistive networks.

Beyond the device demonstrations, this thesis delivered a numerical and software framework that supports the design and operation of electrostatic phase plates in TEM. An aberration-simulation and image-processing toolkit was developed, along with a graphical user interface for fast, user-friendly wave-optics simulations and data analysis. The code is provided in an open, modular form, allowing users to adapt the workflow and interface to their own phase-plate geometries, microscope settings, and optimization targets. In addition, a neural-network model was trained to estimate the orientation of twofold and threefold astigmatism directly from beam or diffraction images. This capability enables rapid aberration diagnosis and can support automated correction by linking the prediction output to the phase-plate controller and microscope software, forming a practical route toward closed-loop alignment and adaptive beam shaping.

Although the two phase plates reported in this thesis (Hyperbola and Helios), which were designed for threefold astigmatism compensation, showed acceptable performance, several limitations remain and can be addressed through new designs. A practical issue was charging, which prevented full-beam operation and required the use of an objective aperture. A future improvement is to refine the current architectures by adding a thick metallic layer on the back

side of the phase-plate chip, so that it also acts as an integrated aperture and helps suppress charging-related artifacts. In addition, further optimization of the fabrication process and the use of thicker devices (e.g., $\sim$50 μm) could improve mechanical stability, leading to smoother phase profiles. A larger thickness also increases the interaction length, thereby relaxing voltage requirements for a given phase stroke.

As discussed throughout this thesis, a major application of hexapole elements in electron microscopy is spherical aberration correction. An important future direction is therefore the design of a MEMS-based double-hexapole corrector for C_s compensation. Such a concept would require transfer optics between the two hexapole elements, which adds complexity. Still, it is a promising route toward cheap, compact, and tunable aberration correction. This approach will likely require higher voltages and stronger phase modulation than the present devices. These requirements could be met by thicker phase plates and, potentially, by stacking multiple phase-plate layers to increase the effective interaction strength while preserving the desired field symmetry.

As a further potential outlook and direction for future work, the tunable electrostatic phase plates developed in this thesis can be extended beyond beam shaping to support computational imaging techniques such as *ptychography* and *computational ghost imaging* in electron microscopy. Both of these techniques allow overcoming aberration limits in resolution. Ptychography , for example, has allowed us to reach the highest lateral resolution (0.02 nm) so far [168]; it is therefore one of the most important developments in modern electron microscopy. In *ptychography*, a tunable phase plate provides controlled probe diversity, which can improve the stability and quality of the phase reconstruction [169]. In *computational ghost imaging*, a tunable phase plate is required to generate a large set of known structured illumination patterns used for image reconstruction [170].

Acknowledgements

The PhD journey is one of the most formative periods of an academic career. Over the past three years, I have benefited from the guidance, support, and generosity of many people, and I have learned far more than what is written in this thesis. I was fortunate to pursue this work in a genuinely friendly environment—one where colleagues treated each other as collaborators and friends, not simply as "students" and "supervisors." I am deeply grateful to everyone who contributed to making this experience both scientifically rigorous and personally meaningful.

I am deeply grateful to **Prof. Marco Beleggia** and **Dr. Vincenzo Grillo** for their continuous guidance, scientific rigor, and encouragement throughout this project. They gave me the freedom to explore my own ideas and, importantly, they never discouraged curiosity—even when those ideas were still rough, or not yet the most realistic version. Their trust and constructive feedback helped me grow into a more independent researcher, while keeping the work grounded in strong physics. I am also grateful to the Quantum Optics group for the supportive and friendly environment. In particular, I thank **Dr. Paolo Rosi** and **Dr. Enzo Rotunno** for their guidance and help during the experiments. I also feel fortunate to have worked with **Prof. Giulio Pozzi**, whose insight and passion for electron microscopy have been truly inspiring, and whose advice and perspective have been invaluable.

I would like to thank **Dr. Alberto Roncaglia** and his team at **CNR-ISMN** Bologna for their essential support during the fabrication work. Their expertise, practical help, and willingness to troubleshoot made a real difference at every critical step. In addition, I am grateful to colleagues at **Forschungszentrum Jülich**, in particular **Dr. Amir Tavabi** and **Prof. Rafal E. Dunin-Borkowski**, for their insightful discussions and valuable guidance throughout these years. I also thank **Dr. Peter Tiemeijer** (Thermo Fisher Scientific) for his support from the industrial side and for helpful technical discussions.

I am grateful for the opportunity to complete an external research stay at **DTU Nanolab**, and I would like to thank **Prof. Ada Ioana Bunea** for her advice and for giving me the opportunity to work in a high-level cleanroom environment. I also thank all **DTU Nanolab** staff and colleagues for their support, professionalism, and assistance during my time there. That period was a significant part of my PhD training and strongly shaped the way I approach experimental work.

I would like to acknowledge the financial support provided by the **PRIN** program (Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale) through the project "AI-TEM—Artificial Intelligence Enhanced Transmission Electron Microscopy

for Advanced Imaging," which enabled this research.
I would like to express my deepest gratitude to my family for their love and encouragement throughout my life and studies. Although my parents are no longer with me, their presence remains in the strength they gave me and the values they left behind. I am deeply thankful to my grandmother, my brother Puya, and my aunts for their warmth, compassion, and unwavering support, which has carried me through life and made this journey possible.
Finally, I would like to thank **the University of Modena and Reggio Emilia**, the **FIM Department**, and all collaborators and staff at **CNR nano** who helped make this project possible.

A

Code and Simulation Availability

The numerical codes developed in this thesis, including Python scripts for electrostatic phase-plate design, wave-optics simulations, data analysis, and graphical user interfaces, are available in a public GitHub repository:
https://github.com/SPEQTEM/TEMtoolkit
These codes were used to generate the analytical models, numerical results, and figures presented throughout the thesis.
Due to software licensing restrictions and file-size limitations, the finite-element simulations performed in COMSOL Multiphysics are not publicly distributed. The corresponding COMSOL models are available from the author upon request.

B

Python Code: Functions

B.0.1. Center Detection

Once the intensity distribution is computed, the first step in the analysis pipeline is to determine the center of the beam spot. This center is required for all subsequent radial operations such as radius estimation, edge detection, and circularity analysis.
The intensity image $I(x, y)$ is first normalized and converted into an 8-bit grayscale representation. To isolate the bright region of the beam, we use Otsu's thresholding method, which selects the optimal threshold by maximizing the inter-class variance between the foreground and background. The resulting binary mask $B(x, y) \in \{0, 1\}$ contains the region of interest corresponding to the main beam.
The beam center (c_x, c_y) is computed as the centroid of this region using the first-order image moments:

$$c_x = \frac{\sum_{x,y} x\, B(x, y)}{\sum_{x,y} B(x, y)}, \qquad c_y = \frac{\sum_{x,y} y\, B(x, y)}{\sum_{x,y} B(x, y)}.$$

If the mask contains no valid foreground pixels (i.e., $\sum B(x, y) = 0$), the algorithm reports an error, ensuring that subsequent analysis steps are not performed on invalid or empty data. The corresponding implementation is shown in Listing **??**.

```python
import cv2
import numpy as np

class CenterFinder:
    @staticmethod
    def find_center(image: np.ndarray):
        # Normalize to 8-bit
        img_norm = cv2.normalize(image, None, 0, 255, cv2.NORM_MINMAX)
        img_uint8 = img_norm.astype(np.uint8)

        # Automatic thresholding (Otsu)
        _, binary = cv2.threshold(
            img_uint8, 0, 255, cv2.THRESH_BINARY + cv2.THRESH_OTSU
        )
```

```
15
16          # Compute centroid via image moments
17          M = cv2.moments(binary)
18          if M["m00"] == 0:
19              raise ValueError("No valid bright region detected for center
                ↪  finding.")
20
21          cx = M["m10"] / M["m00"]
22          cy = M["m01"] / M["m00"]
23
24          return (cx, cy), img_uint8
```

B.0.2. Radius Estimation

After the beam center is identified, the next step is to estimate the radius of the beam spot. This quantity provides a global measure of the beam size and is also used as an initial value for more advanced analyses such as angular edge tracking and circularity computation.
To extract the radius, the normalized 8-bit image is first binarized using Otsu's thresholding, producing a foreground mask that outlines the bright region of the beam. From this mask, we compute the external contour using OpenCV's `findContours` function. Among all detected contours, only the one with the largest area is retained, ensuring robustness against noise and spurious bright pixels.
Given a center location (c_x, c_y) and a contour represented by a set of points $\{(x_i, y_i)\}$, the radius is computed as the median of the Euclidean distances between the center and each contour point:

$$r_i = \sqrt{(x_i - c_x)^2 + (y_i - c_y)^2}, \qquad R = \text{median}\{r_i\}.$$

Using the median instead of the mean provides improved stability in the presence of local irregularities, weak noise, or small distortions in the contour.
The implementation is shown in Listing **??**.

```
1   import cv2
2   import numpy as np
3
4   class RadiusEstimator:
5       @staticmethod
6       def estimate_from_contour(img_uint8: np.ndarray, center: tuple[float,
        ↪  float]) -> float:
7           cx, cy = center
8
9           # Automatic thresholding
10          _, binary = cv2.threshold(
11              img_uint8, 0, 255, cv2.THRESH_BINARY + cv2.THRESH_OTSU
12          )
13
```

```
        # Extract contours
        contours, _ = cv2.findContours(
            binary, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_NONE
        )
        if not contours:
            raise ValueError("No contours found.")

        # Largest contour (beam outline)
        contour = max(contours, key=cv2.contourArea)

        # Distance of each contour point to the center
        distances = [
            np.hypot(pt[0][0] - cx, pt[0][1] - cy) for pt in contour
        ]

        # Median radius for robustness
        return float(np.median(distances))
```

B.0.3. Angular Edge Detection and Radial Profile Analysis

With the beam center (c_x, c_y) and an initial estimate of its radius, the next step is to characterize the beam boundary as a function of angle. This allows a full reconstruction of the beam shape, the extraction of circularity metrics, and the identification of angular maxima and minima associated with aberrations such as twofold or threefold astigmatism.
For a given angle θ, a one–dimensional radial profile is sampled along the ray originating at the center. The sampling is implemented by tracing a line

$$(x(t), y(t)) = \big(c_x + t\cos\theta,\ c_y - t\sin\theta\big), \qquad t \in [0, L],$$

where L is chosen based on the estimated radius or, if unavailable, on the diagonal of the image. Linear interpolation is used to extract intensity values along the ray, producing a one–dimensional profile $I_\theta(t)$.
To improve robustness, each profile is smoothed using a Gaussian filter. The radial position of the beam edge is then determined using one of two criteria:

1. **Sharp-drop method:** The edge is defined as the first point where the normalized intensity drops below a fixed threshold (typically 20% of the peak value). This approach is useful for beams with steep intensity transitions.

2. **Gradient method:** The edge is taken as the point of maximum negative slope, i.e. the minimum of $\nabla I_\theta(t)$. This is more appropriate for smoothly decaying beams.

Repeating this procedure for all angles $\theta \in [0°, 360°]$ produces a set of radii $\{r(\theta)\}$. From these, we compute the mean radius, standard deviation, and the relative circular deviation

(RCD):

$$\mathrm{RCD} = 100 \times \frac{\sigma_r}{\bar{r}},$$

which provides a quantitative measure of circularity. When the deviation exceeds a chosen threshold (e.g. 2%), local extrema in $r(\theta)$ are identified using peak-detection methods. These angular locations often correspond to principal aberration axes.
The full implementation is provided in Listing **??**.

```python
import numpy as np
from scipy.ndimage import gaussian_filter1d, map_coordinates
from scipy.signal import find_peaks

class RadiusAnalyzer:

    @staticmethod
    def extract_radial_profile(
        img, cx, cy, angle_deg, method,
        radius=None, length=None, smooth_sigma=1
    ):
        angle_rad = np.deg2rad(angle_deg)

        # Choose sampling length
        if length is not None:
            L = length
        elif radius is not None:
            L = int(2 * radius)
        else:
            h, w = img.shape
            L = int(np.hypot(w, h))

        # Sample coordinates along the ray
        x = np.linspace(cx, cx + L * np.cos(angle_rad), L)
        y = np.linspace(cy, cy - L * np.sin(angle_rad), L)

        # Extract profile via interpolation
        profile = map_coordinates(img, [y, x], order=1, mode='constant',
        ↪  cval=0.0)
        smoothed = gaussian_filter1d(profile, sigma=smooth_sigma)

        # Edge detection
        if method == 'sharp_drop':
            normed = smoothed / (np.max(smoothed) + 1e-8)
            drop_indices = np.where(normed < 0.2)[0]
            edge_index = drop_indices[0] if len(drop_indices) > 0 else
            ↪  np.argmin(smoothed)

        elif method == 'gradient':
```

```
            gradient = np.gradient(smoothed)
            edge_index = np.argmin(gradient)

        return smoothed, edge_index

    @staticmethod
    def analyze(img, cx, cy, method, radius=None, angle_step=1, smooth_sigma=2):
        angles = np.arange(0, 361, angle_step)
        radii = []

        for ang in angles:
            try:
                _, r = RadiusAnalyzer.extract_radial_profile(
                    img, cx, cy, ang, method,
                    radius=radius, smooth_sigma=smooth_sigma
                )
                radii.append(r)
            except:
                radii.append(np.nan)

        radii = np.array(radii)
        valid = radii[~np.isnan(radii)]

        mean_r = np.mean(valid)
        std_r = np.std(valid)
        rcd = 100 * std_r / mean_r
        max_dev = np.max(np.abs(valid - mean_r))

        # Detect peaks only when strong deviation is present
        if rcd > 2:
            peaks, _ = find_peaks(radii, prominence=0.01 * np.nanmax(radii))
            peak_angles = angles[peaks]
            peak_radii = radii[peaks]
        else:
            peak_angles, peak_radii = [], []

        return angles, radii, mean_r, std_r, rcd, max_dev, peak_angles,
        ↪ peak_radii
```

B.0.4. Aberration Detection and Circularity Analysis

Once the full angular radius profile $r(\theta)$ is obtained, the next step is to identify characteristic features in the profile that correspond to specific probe aberrations. In an ideal circular beam, all radii are equal and the profile is constant. In practice, lens imperfections or intentional aberration terms introduce angular modulations, producing maxima and minima in $r(\theta)$. Their number, spacing and symmetry contain direct information about the underlying

aberration: for example, twofold astigmatism produces two maxima and two minima, while threefold astigmatism produces a threefold modulation.

To identify these extrema, the radius profile is first smoothed using a Gaussian filter to suppress high-frequency noise. Because the profile is periodic over 0°–360°, it is extended by concatenation to avoid boundary artifacts during peak detection. Extrema are then located using a prominence-based peak-finding algorithm. A secondary check identifies "flat" extrema, i.e. regions where the gradient is nearly zero over multiple consecutive samples, ensuring robustness against plateau-like peaks.

Once the sets of maxima and minima are extracted, the detector computes the *circularity* of the beam, defined as

$$C = 1 - \frac{\max(r) - \min(r)}{\overline{r}}, \qquad 0 \leq C \leq 1,$$

where $\overline{r}$ is the mean radius. A perfectly circular beam yields $C = 1$, whereas larger deviations reduce the score. Because the radii are converted to micrometers via the physical pixel size, this metric directly reflects the physical beam distortion.

Finally, the detected maxima and minima are classified into aberration types by examining both their multiplicity and their symmetry. For example, the presence of two maxima and two minima with comparable values indicates pure twofold astigmatism, whereas asymmetric pairs indicate a mixture of second- and third-order aberrations. A similar criterion is applied for threefold astigmatism.

The computational implementation is shown in Listing **??**.

```python
import numpy as np
from scipy.signal import find_peaks
from scipy.ndimage import gaussian_filter1d

class AberrationDetector:
    """
    Detects maxima and minima in the angular radius profile and uses these
    extrema to compute circularity and classify aberration type.
    """

    def __init__(self, radii, dx, angles, angles_rad, cx, cy, prominence=5):
        self.radii = np.array(radii, dtype=float)
        self.dx = dx
        self.angles = angles
        self.angles_rad = angles_rad
        self.cx = cx
        self.cy = cy
        self.prominence = prominence

        self.peaks_max = None
        self.peaks_min = None
```

```
22
23      # ------------------------------------------------------------
24      # Peak detection in the angular radius profile
25      # ------------------------------------------------------------
26      def detect_peaks(self):
27          N = len(self.radii)
28
29          # Smooth profile to reduce noise
30          smoothed = gaussian_filter1d(self.radii, sigma=1.0)
31          self.radii = smoothed.copy()
32
33          # Periodic extension to avoid boundary issues
34          wrapped = np.concatenate([smoothed, smoothed, smoothed])
35          offset = N
36
37          # Primary maxima/minima detection
38          raw_max, _ = find_peaks(wrapped, prominence=self.prominence)
39          raw_min, _ = find_peaks(-wrapped, prominence=self.prominence)
40
41          peaks_max = raw_max[(raw_max >= offset) & (raw_max < 2 * offset)] -
            ↪ offset
42          peaks_min = raw_min[(raw_min >= offset) & (raw_min < 2 * offset)] -
            ↪ offset
43
44          # ------------------------------------------------------------
45          # Additional detection of flat extrema = near-zero gradient regions
46          # ------------------------------------------------------------
47          wrapped_grad = np.gradient(wrapped)
48          flat = np.where(np.abs(wrapped_grad) < 1e-3)[0]
49
50          suspicious_max = []
51          suspicious_min = []
52
53          if len(flat) > 0:
54              groups = np.split(flat, np.where(np.diff(flat) != 1)[0] + 1)
55
56              for g in groups:
57                  if len(g) == 0:
58                      continue
59
60                  start, end = g[0], g[-1]
61                  center = (start + end) // 2
62
63                  if not (offset <= center < 2 * offset):
64                      continue
65
66                  value = wrapped[center]
```

```
                left = wrapped[start - 2] if start - 2 >= 0 else value
                right = wrapped[end + 2] if end + 2 < len(wrapped) else value

                if value > left and value > right:
                    suspicious_max.append(center - offset)
                elif value < left and value < right:
                    suspicious_min.append(center - offset)

        # Merge normal and suspicious extrema
        peaks_max = np.array(sorted(set(peaks_max.tolist() + suspicious_max)))
        peaks_min = np.array(sorted(set(peaks_min.tolist() + suspicious_min)))

        # ------------------------------------------------------------
        # Sanity check: reject extrema if too symmetric (false modulation)
        # ------------------------------------------------------------
        if len(peaks_max) == 0 or len(peaks_min) == 0:
            self.peaks_max = np.array([], dtype=int)
            self.peaks_min = np.array([], dtype=int)
            return

        max_vals = self.radii[peaks_max]
        min_vals = self.radii[peaks_min]
        mean_val = np.mean([np.max(max_vals), np.min(min_vals)])
        std_val = np.std([np.max(max_vals), np.min(min_vals)])

        if std_val / mean_val < 0.02:
            # No reliable modulation → treat as circular
            self.peaks_max = np.array([], dtype=int)
            self.peaks_min = np.array([], dtype=int)
            return

        self.peaks_max = peaks_max
        self.peaks_min = peaks_min

    # ------------------------------------------------------------
    # Circularity
    # ------------------------------------------------------------
    def compute_circularity(self):
        if len(self.radii) == 0:
            return None

        radii_um = self.radii * self.dx * 1e6
        mean_r = np.mean(radii_um)
        deviation = np.max(radii_um) - np.min(radii_um)

        if mean_r == 0:
            return 0
```

```
114
115         C = 1.0 - (deviation / mean_r)
116         return float(np.clip(C, 0, 1))
117
118     # ------------------------------------------------------------
119     # Aberration classification
120     # ------------------------------------------------------------
121     def aberration_status(self):
122         if self.peaks_max is None or self.peaks_min is None:
123             self.detect_peaks()
124
125         peaks_max = np.array(self.peaks_max, dtype=int)
126         peaks_min = np.array(self.peaks_min, dtype=int)
127
128         if len(peaks_max) == 0 or len(peaks_min) == 0:
129             return "No Aberration detected"
130
131         max_r = (self.radii[peaks_max] * self.dx) * 1e6
132         min_r = (self.radii[peaks_min] * self.dx) * 1e6
133
134         mean_val = np.mean([np.max(max_r), np.min(min_r)])
135         threshold = 0.05 * mean_val
136
137         # Classification based on extrema multiplicity
138         if len(max_r) == 2 and len(min_r) == 2:
139             return "2-fold astigmatism"
140         if len(max_r) == 3 and len(min_r) == 3:
141             return "3-fold astigmatism"
142
143         # Mixed patterns
144         if len(max_r) == 3 or len(min_r) == 3:
145             return "Combination of astigmatism detected - 3-fold dominated"
146         if len(max_r) == 2 or len(min_r) == 2:
147             return "Combination of astigmatism detected - 2-fold dominated"
148
149         return "No Aberration detected"
```

List of Figures

1.1 Early optics milestones. (a) The Nimrud lens (8th c. BC), a rock-crystal magnifier (Image credit: British Museum). (b) Medieval Latin illustration of a parabolic burning mirror igniting ships (Archimedes legend; Alhazen tradition). (c) Tommaso da Modena's 1352 fresco at San Nicolò, Treviso — earliest known depiction of spectacles on Cardinal Hugh of Saint-Cher. 6

1.2 Early microscopes. (a) Janssen-type compound microscope (c. 1590s). (b) Antoni van Leeuwenhoek single-lens microscope (1670s). (c) Robert Hooke's compound microscope with an image from his book *Micrographia* (1665). (Images credit: Science Museum Group, London. (website).) 6

1.3 Nature of light. (a) Isaac Newton with *Opticks* (1704): light as particles. (b) Christiaan Huygens with *Traité de la lumière* (1690): light as waves. (Image credit: Smithsonian Libraries, Washington, website). 7

1.4 Young's double-slit experiment 8

1.5 Wave propagation in the direction of $\mathbf{E} \times \mathbf{B}$ [6]. 9

1.6 Historical milestones in the development of the electron microscope. (a) Ernst Ruska (right) and Max Knoll (left) with the first magnetic electron microscope, Berlin, 1932. (b) Sketch by Ernst Ruska, dated 9 March 1931, depicting the cathode ray tube used to test one-stage and two-stage electron-optical imaging with magnetic electron lenses ("electron microscope"). (c) Ernst Ruska adjusting a Siemens electron microscope, 1949. 11

1.7 Three early commercial electron microscopes. (a) Siemens Übermikroskop, 1939 (Germany); (b) Toronto TEM, 1939 — E. F. Burton, J. Hillier, A. F. Prebus (Canada); (c) RCA Model B, 1940 — V. K. Zworykin, J. Hillier (USA). 12

2.1 Electron velocity and de Broglie wavelength as a function of accelerating voltage. The left plot compares classical and relativistic velocity; the right plot shows the corresponding electron wavelength. 16

2.2 (a) Overview of electron–specimen interactions in a thin sample, including elastic and inelastic processes. (b) Elastic-scattering regimes by angle: low-angle forward (coherent diffraction), high-angle forward (large-momentum-transfer), and backscattered (>90°). 21
2.3 Illustration of a two-lens ($4f$) Fourier imaging system. The first lens performs a Fourier transform of the object field, producing its spatial-frequency spectrum in the back focal plane (BFP). The second lens performs an inverse Fourier transform, reconstructing the image in its back focal plane. The intermediate BFP allows direct access to the frequency domain, where filters, apertures, or phase plates can modify selected spatial frequencies. The final image is reproduced with unit magnification ($M = 1$) and a 180° rotation relative to the object, not a mirror inversion. 23
2.4 Illustration of the diffraction limit set by a finite numerical aperture. (a) Wave propagation in the xz plane showing the probe formation. (b) Airy pattern in the focal xy plane. (c) Three-dimensional representation of the Airy-pattern intensity in the focal plane. (d–f) Summed intensity of two identical Airy patterns as their lateral separation decreases. (f) Diffraction-limited condition (Rayleigh criterion), where the two probes are just resolved and cannot be reliably distinguished in intensity. 24
2.5 Effect of wavelength and numerical aperture on the focused Airy beam. (a–c) Increasing the electron energy (shorter wavelength) reduces the probe size. (d–f) Increasing the numerical aperture (larger aperture angle) further decreases the Airy spot size but also reduces the depth of focus. 25
2.6 Effect of convergence semi–angle α on the point–spread function (PSF) at focus and defocus $\Delta f = 20\ nm$. At exact focus, increasing α (i.e., larger NA $\simeq \sin\alpha \approx \alpha$) reduces diffraction blur, shrinking the Airy core (first zero $r_1 \approx 0.61\,\lambda/\mathrm{NA}$). Under defocus (either sign), larger α makes the PSF broaden and rings strengthen. Thus, a wider aperture yields a smaller in-focus spot but greater sensitivity to Δf. 27
2.7 Schematic illustration of aberration in an electron optical system (Image credit: [41], modified by the author). 28
2.8 Phase profiles of defocus, spherical aberration, and twofold and threefold astigmatism shown in 2D (wrapped phase) and corresponding 3D surface representations. 29
2.9 Effect of spherical aberration on the point-spread function (PSF). The transverse PSF is taken along the red dashed line. 31
2.10 Chromatic aberration 32
2.11 Schematic illustration of two-fold astigmatism. Anisotropic focusing produces two orthogonal line foci, indicated by the x-plane (green) and y-plane (red). Intensity cross-sections are shown at underfocus, at the circle of least confusion, and at overfocus, illustrating the spot evolution along the optical axis. 33

3.1 Overview of Transmission Electron Microscope (TEM) . . . 37
3.2 Electron gun and common electron sources. (a) Electron gun with a Wehnelt cylinder. (b) Tungsten filament. (c) LaB_6 cathode. (d) Field-emission (FEG) tip [41]. . . . 38
3.3 Electrostatic electron lens configurations and simulations. (a) Image of a three-electrode Einzel lens device. (b) Electrostatic simulation showing the electric potential distribution generated by the three cylindrical electrodes. (c) Simulation of electron trajectories through the lens; red arrows represent the local electric field vectors. . . . 41
3.4 Magnetic lens structure and field distribution . . . 41
3.5 Electron trajectories in a magnetic lens . . . 42
3.6 Examples of magnetic lenses in TEM . . . 43
3.7 Schematics comparing TEM and STEM imaging modes. . . . 46
3.8 Imaging vs. Diffraction modes. . . . 47
3.9 TEM images and selected-area electron diffraction (SAED) patterns of (a,b) Al_2Sm, (c,d) Al_2Ca, (e,f) lamellar $(Mg,Al)_2Ca$, and (g,h) $Mg_{17}Al_{12}$ [62]. . . . 49
3.10 Schematic layout of STEM imaging. A focused, raster-scanned electron probe is formed by the condenser–objective system and steered by the scan (deflector) coils across the specimen. Transmitted electrons are collected by the bright-field (BF), annular dark-field (ADF), and high-angle annular dark-field (HAADF) detectors. . . . 50
3.11 Bright-field (left) and annular dark-field (right) STEM images of Pt catalyst particles on a graphite support, showing dark contrast in BF and bright (reversed) contrast in ADF due to high-angle scattering collected by the annular detector. . . . 51
3.12 Atomic-resolution HAADF-STEM image of a BFO/LSMO superlattice grown on an STO substrate, with EELS elemental maps (right) of La $M_{4,5}$, Mn $L_{2,3}$, Fe $L_{2,3}$, and Ti $L_{2,3}$, and a composite overlay [65]. . . . 52
3.13 Device, magnetic field, and electron beam trajectory simulations for (a) a quadrupole and (b) a hexapole [67]. . . . 53
3.14 Hexapole Cs correctors [50]. . . . 55
3.15 Quadrupole-Octopole C_s corrector. The corrector consists of three octupoles (O2, O4 and O6) and four quadrupoles (Q1, Q3, Q5 and Q7). The radial elongation of the beam through the corrector is indicated along two orthogonal axes x and y (Adapted from Ref. [80] and redesigned by the author). . . . 58
3.16 Zemlin tableau recorded at a beam tilt of 10.8 mrad. (a) Uncorrected but aligned microscope showing tilt-dependent variations dominated by residual spherical aberration. (b) Corrected microscope showing minimal changes in defocus and astigmatism [81]. . . . 60

3.17 Simulated Ronchigrams illustrating the effect of common aberrations in STEM. The reference Ronchigram shows a large central constant-phase region, while added twofold astigmatism, coma, and threefold astigmatism introduce characteristic distortions and symmetry changes [84]. 61

4.1 Spatial-frequency decomposition of an image. (a) Original image. (b) Fourier-magnitude spectrum (log scale) with three annular masks marking low, mid, and high frequency bands. (c) Low-pass frequencies: slowly varying components that set overall contrast and large-scale structure. (d) Mid-band frequencies: edges and mesoscale textures. (e) High-pass frequencies: finest details and high-frequency texture (also most sensitive to noise). 65

4.2 Schematic of image formation pathway of Weak Phase Object (WPO) 66

4.3 Defocus phase contrast in TEM: in-focus ($\Delta f = 0$) versus Scherzer defocus ($\Delta f = \Delta f_S$), with pathway schematic, CTF, and example phase object/images illustrating contrast changes. 68

4.4 Zernike and HFPP phase plates. 69

4.5 Phase-contrast TEM using thin-film phase plates. (a–c) Negatively stained horse spleen ferritin: (a) phase TEM image acquired with a Zernike phase plate, and (b,c) conventional TEM images recorded at defocus values of 2550 nm and 540 nm, respectively [107]. (d–f) Carbon nanotube sample: (d) conventional TEM near focus, (e) underfocus, and (f) with a Hole-free phase plate (HFPP) providing a central phase shift of $\phi_{\mathrm{PP}}(0) = -0.8\pi$ [108]. 71

4.6 Hilbert (top) and Schlieren (bottom) phase plates. 72

4.7 Hilbert phase plate on a biological sample. 72

4.8 Three examples of electrostatic Boersch phase plates: (a,b) einzel-lens configurations formed by five alternating layers of Au and either Al_2O_3 [121] (a) or Si_3N_4 [120] (b); (c) a modified design in which the central electrodes are removed by FIB milling [120]. 74

4.9 Zach phase plate [123, 124] 75

4.10 Vortex beams with different vortex charges ℓ, showing helical phase fronts (cyan), circulating probability current (orange), and the transverse amplitude–phase distribution at z=0 with the corresponding radial profiles [128]. . . 76

4.11 MEMS based Vortex beam generator with chopstick (left) and Vortex beams (recorded in the diffraction plane of the aperture) shown as a function of bias voltage applied to one of the chopsticks (right) [98]. 78

4.12 SEM images of electrostatic OAM sorter elements (a) S1 (unwrapper) and (b) S2 (corrector). (c) Schematic diagram of part of an electron microscope column that contains a sorter, with experimental images showing the evolution of the electron wave function [135]. 79

4.13 48-element programmable phase plate (left) and the three rows of figures on the right. From top to bottom, the phase configurations set on the phase plate, the simulated probe shapes, and the resulting experimental probe images, respectively [141]. 79

4.14 Magnetic-material phase plates. (a) Magnetized Ni needle generating an electron vortex. (b) CoFe rod in the condenser forming a vortex probe. (c) The samerod in the objective yielding spiral-type phase contrast (Zernike–Hilbert) [144, 145]. 81

4.15 Ampère (current-driven) phase plate. (a) Electron micrograph of the tunable Ampère phase plate (TAPP). (b) Off-axis holography phase maps of the surrounding vacuum region for a 4 mA current. (c) Simulated out-of-focus Fresnel images of an NCP at 0.5 µm defocus. (d) Simulated TAPP images for a 0.6 mA current. 82

4.16 Laser-based control of the electron phase in a TEM: a resonantly enhanced standing laser wave in an optical cavity imparts a phase shift to the unscattered beam in the objective-lens back focal congugate plane, which is then recombined with the scattered beams to form phase-contrast images [157]. . . 84

5.1 Three commonly used electrode layouts in conventional electrostatic phase-plate designs. (a) An open-circuit electrode that provides only a uniform potential. (b) A multi-electrode arrangement used to approximate a desired potential gradient. (c) A closed-circuit electrode in which the applied current creates a distributed voltage drop. 92

5.2 Current-controlled electrostatic phase plate. (a) Geometry and simulated electric potential distribution, showing a localized drop at the active area enabled by combining two materials with different electrical conductivity (high-conductivity leads and a lower-conductivity active area). (b) Line profile of the electric potential along the circuit, demonstrating that the voltage drop occurs almost exclusively at the active area. (c) Conventional circuit with uniform conductivity, in which the voltage drop is distributed along the entire path according to the local resistance. 92

5.3 Control of a monotonic potential drop by optimizing the shape of the low-conductivity region. (a) Rectangular shape producing an approximately linear voltage drop. (b) Triangular shape. (c,d) Tapered geometries enabling tailored non-linear potential profiles. 94

5.4 Ideal 96

5.5 Comparison between a conventional single-chip design and the sandwich design. Field localization in the sandwich configuration reduces parasitic distortions from external connections and results in a more symmetric phase profile. 96

5.6 Sandwich-design fabrication process flow for the thin-film phase plate. The device is fabricated as two separate chips (top and bottom), which are then aligned and bonded to form the final assembled chip. 97
5.7 Full sandwich-chip layout. (a) Protochips electrical-biasing holder with an open-top design (left) and the sandwich-chip design with mechanical alignment (right). (b) Fabricated devices: (left) SEM image of the mechanical alignment steps (dashed square in the middle image), (middle) final chip formed by stacking the two chips, and (right) optical micrographs of the top- and bottom-side apertures showing accurate alignment achieved by the mechanical aligners. . . 98

6.1 COMSOL simulation domain for the electrostatic phase plate: the device is placed in a vacuum region enclosed by a grounded infinite-element domain; electric potentials are applied via Terminal boundary conditions, the silicon substrate is grounded, and an electron-path subdomain is used to compute the phase shift using the General Projection function 103
6.2 Vertical evaluation geometry used in COMSOL to extract the electron phase shift. The purple cylinder defines the projection volume used with COMSOL's General Projection operator to compute the phase map $\varphi(x, y)$ on a plane by integrating along z. The dashed line indicates the integration path at a fixed lateral position (x_0, y_0), yielding $\varphi(x_0, y_0)$. The same approach can be applied to tilted illumination by orienting the projection direction. 105
6.3 Single projection vs Multislice. 108
6.4 Graphical user interface (GUI) developed in Python for interactive control of the multislice simulation, wave–optics propagation, and phase-plate analysis. The interface provides real-time parameter adjustment and visualisation of wavefronts, potentials, and simulation outputs. 111
6.5 Graphical user interface (GUI) developed in Python for interactive Simulation 114
6.6 Graphical user interface (GUI) developed in Python for interactive Simulation 117
6.7 Simulation of an abberated probe shape (left) and corresponding radial profile used as ML training dataset (right) 119
6.8 (left) Neural-network architecture with preprocessing, harmonic augmentation, fully connected layers, and a two-neuron output encoding the aberration orientation; (right) training and validation loss history for the same model. . . 119
6.9 Experimental probe images with deliberately applied twofold astigmatism at different magnifications. Overlaid markers indicate the neural-network–predicted orientations of the twofold and threefold astigmatism. 120

7.1 Phase plate design process flow 122
7.2 Comparison between two hexapole configurations used for generating threefold astigmatism: a conventional 6-pole design (Hexapole: fixed orientation) and an extended 12-pole (dodecapole: tunable orientation). 123
7.3 hyperbola-layout 124

7.4 Simulated phase profiles generated with the dodecapole phase plate, including quadrupole, hexapole, and dodecapole modes, shown together with the corresponding analytical Zernike polynomials. 125
7.5 Hyperbola phase plate: optimized electrode potentials for rotating the three-fold phase from 0° to 30°, together with the corresponding simulated phase and Fourier-transform (in-focus diffraction patterns). 126
7.6 Effect of device radius and thickness on the total phase shift. The phase is evaluated along a line profile through the center of the phase profile. (a) Phase shift for an 80 μm device diameter at different applied voltages. (b) Phase shift for a 120 μm device diameter under the same voltage conditions. (c) Phase shift for different device thicknesses. 128
7.7 Helios layout and design configuration 129
7.8 Helios current-controlled electrostatic phase plate: (a) geometry layout showing the materials and overall design; (b) electrostatic mode, where both ends of the electrode are connected to the same potential; (c) current-controlled mode, where different potentials are applied to the two ends of the electrode. . 129
7.9 Optimized electrode shape (black dashed line) and electric potential at $z = 0$ along the circular contour near the boundary electrodes (red dashed line). . . . 130
7.10 Helios phase plate: optimized electrode potentials for rotating the three-fold phase from 0° to 30°, together with the corresponding simulated phase and Fourier-transform (in-focus diffraction patterns). 131
7.11 Helios phase plate: simulated phase shift showing (a) the angular phase profile along a circular contour for five orientations and (b) the phase line profile along a central line through the center of the phase profile. 132
7.12 Analytical vortex phase in wrapped and unwrapped form, and the phase profile extracted along the dashed circular contour, showing the linear azimuthal dependence. 133
7.13 Simulation geometry for parameter optimization, including 41 discretized boundary electrodes. 133
7.14 Boundary-electrode parameter optimization. Left: initial linear voltage ramp used as the initial guess. Right: optimized solution. The red profile indicates the required boundary-electrode voltage trend. 134
7.15 Shape-optimization workflow for designing a current-controlled electrostatic vortex-beam generator. 135
7.16 Current-controlled electrostatic vortex beam generator phase plate design and chip layout. 136
7.17 OAM sorter analytical target. 137
7.18 Current-controlled OAM sorter: (a) optimized device geometry (300 μm size) with the 70 μm region of interest (ROI) indicated by a red dashed square; (b) unwrapped phase shift in the ROI; (c) wrapped phase shift in the ROI. 138

7.19 Comparison between (a) the analytical target and (b) the simulated phase of the OAM sorter phase plate, together with (c) the relative phase error between simulation and target (percentage). 138
7.20 Current-controlled electrostatic OAM sorter design and chip layout. 139

8.1 Process flow for chips fabrication 144
8.2 SEM images of the dodecapole hyperbola phase-plate device. (a) Overview of the fabricated device showing the four contact pads used for wire bonding. (b) Active area of the device. (c) Zoom-in on the electrode edge, highlighting the $\sim 150\,\mathrm{nm}$ Au layer and the step-like sidewall profile produced by the Bosch DRIE process. 146
8.3 SEM images of the Helios phase-plate device. (a) Overview of the fabricated chip after wire bonding. (b) Active area of the device before FIB milling. (c) Active area of the device after FIB milling. 146
8.4 SEM images of vortex and OAM sorter phase plates: (a) vortex phase plate with Au-removed region; (b) OAM sorter with three needles; (c) close-up of one needle showing the fabrication quality and nearly vertical sidewalls. 147
8.5 Electrical characterization using a probe station. 148

9.1 A schematic overview of the Thermo Fisher Scientific (S)TEM Spectra 300 (SPEQTEM) microscope 150
9.2 Mounting MEMS on the Thermo Fisher Scientific NanoEx-i/v TEM holder . . 151
9.3 Holder connections: the holder is connected to a controller via an 8-pin connector, and the controller is connected to the computer via USB. The custom GUI controls the controller during the experiments. 152
9.4 Hyperbola phase plate: comparison between experimental and simulated diffraction patterns for five orientations of the three-fold phase between 0° and 30°, demonstrating controlled rotation of the phase profile. 154
9.5 Compensation of the microscope threefold astigmatism using the Hyperbola phase plate for defocused diffraction patterns at 200 nm and 400 nm. In the "OFF" state, the phase plate is grounded. The middle column shows the orientation estimated by the developed deep neural network, which selects the optimized electrode potentials and automatically aligns the phase plate to the microscope threefold orientation. In the "ON" state, the optimized potentials are applied via the GUI. 155
9.6 Helios phase plate: comparison between experimental and simulated diffraction patterns for five orientations of the three-fold phase between 0° and 30°, demonstrating controlled rotation of the phase profile. 156
9.7 Threefold-astigmatism compensation in the in-focus diffraction pattern using the Helios phase plate shows a reduction of parasitic aberrations when the phase plate is biased with the optimized electrode potentials. 157

9.8 Three images used for TIE reconstruction: one in-focus image with the phase plate grounded, and two images recorded at $\Delta f = 5\,\mathrm{mm}$ with equal-magnitude voltages of opposite polarity. 159
9.9 Phase reconstruction for two operating configurations: (1) current-controlled mode (electric current flows through the phase plate), and (2) electrostatic mode (no current flow). 160

List of Tables

2.1 Electron parameters at various accelerating voltages: Lorentz factor γ, velocity ratio β, relativistic and classical velocities, and the de Broglie wavelength λ. . 17
2.2 Interaction constant C_E as a function of acceleration voltage. 20

3.1 Comparison of electron sources [41]. 39

4.1 Mean inner potential V_0 and corresponding $\pi/2$ phase-shift thickness $t_{\pi/2}$ for selected elements at 100 kV and 300 kV accelerating voltages [107]. 70

8.1 SOI wafers used in this work. 143

Bibliography

[1] The British Museum. Rock crystal object, assyrian, possibly a magnifying lens. https://www.britishmuseum.org/collection/object/W_-90959?selectedImageId=564340001. Accessed: 7 July 2025.

[2] William J. Croft. *Under the Microscope: A Brief History of Microscopy*. World Scientific, Singapore, 2006.

[3] Ibn al Haytham. *The Book of Optics*. c. 1021. Originally written in Arabic as *Kitāb al-Manāir*; also known in Latin as *De Aspectibus* or *Perspectiva*.

[4] Clara Sue Ball. The early history of the compound microscope. *Bios*, 37(2):51–60, 1966.

[5] Isaac Newton. *Opticks: or, a Treatise of the Reflexions, Refractions, Inflexions and Colours of Light*. Printed for Sam. Smith, and Benj. Walford, London, 1704. First edition.

[6] Eugene Hecht. *Optics*. Pearson Education, Harlow, England, 5th edition, 2017.

[7] Thomas Young. The bakerian lecture:: Experiments and calculations relative to physical optics. *Philosophical transactions of the Royal Society of London, 94(94):1–16,1804. 25*, 2016.

[8] Francesco Maria Grimaldi. *Physico mathesis de lumine, coloribus, et iride, aliisque annexis libri duo*. Vittorio Bonati, Bologna ("Bonomia"), Italy, 1665. pp. 1–11, in Latin.

[9] Pier Giorgio Merli, Gian Franco Missiroli, and Giulio Pozzi. On the statistical aspect of electron interference phenomena. *Am. J. Phys*, 44(3):306–307, 1976.

[10] Stefano Frabboni, Gian Carlo Gazzadi, Giulio Pozzi, et al. Young's double-slit interference experiment with electrons. *American Journal of Physics*, 75(11):1053–1055, 2007.

[11] S. Frabboni, C. Frigeri, G. C. Gazzadi, and G. Pozzi. Young-feynman experiment with electrons: Double-slit interference pattern and wave-particle duality. *American Journal of Physics*, 79(6):615–618, 2011.

[12] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Vol. III: Quantum Mechanics*. Addison-Wesley, Reading, Massachusetts, 1965.

[13] Frank L. Pedrotti, Leno S. Pedrotti, and Leno M. Pedrotti. *Introduction to Optics*. Pearson Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2007.

[14] James Clerk Maxwell. *A treatise on electricity and magnetism*, volume 1. Clarendon press, 1873.

[15] Richard P Feynman, Robert B Leighton, Matthew Sands, and Everett M Hafner. The feynman lectures on physics; vol. i. *American Journal of Physics*, 33(9):750–752, 1965.

[16] David J Griffiths. *Introduction to electrodynamics*. Cambridge University Press, 2023.

[17] Robert Eisberg and Robert Resnick. *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. Wiley, New York, 2nd edition, 1985.

[18] Arthur Beiser. *Concepts of modern physics*. 2003.

[19] Ahmed H. Zewail. Micrographia of the twenty-first century: from camera obscura to 4d microscopy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1914):1191–1204, Mar 2010.

[20] Friedel Weinert. Davisson–germer experiment. In Daniel Greenberger, Klaus Hentschel, and Friedel Weinert, editors, *Compendium of Quantum Physics*, pages 150–152. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[21] G. P. Thomson and A. Reid. Diffraction of cathode rays by a thin film. *Nature*, 119(3007):890–890, 1927.

[22] The Nobel Prize Organization. The nobel prize in physics, 2025. Accessed: 2025-07-17.

[23] Ernst Ruska. The development of the electron microscope and of electron microscopy. *Rev. Mod. Phys.*, 59:627–638, Jul 1987.

[24] Manfred Von Ardenne. Das elektronen-rastermikroskop. praktische ausführung. *Zeitschrift für technische Physik*, 19(11):407–416, 1938.

[25] Manfred Von Ardenne. Das elektronen-rastermikroskop: Theoretische grundlagen. *Zeitschrift für Physik*, 109(9):553–572, 1938.

[26] Gottfried Möllenstedt. My early work on convergent-beam electron diffraction. *physica status solidi (a)*, 116(1):13–22, 1989.

[27] Richard P Feynman. B there's plenty of room at the bottom. 1959.

[28] Otto Scherzer. Über einige fehler von elektronenlinsen. *Zeitschrift für Physik*, 101(9):593–603, 1936.

[29] PW Hawkes. Aberration correction past and present. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1903):3637–3664, 2009.

[30] Harald H Rose. Historical aspects of aberration correction. *Journal of electron microscopy*, 58(3):77–85, 2009.

[31] Harald Rose. Aberration correction in electron microscopy. *International journal of materials research*, 97(7):885–889, 2022.

[32] Max Haider, Harald Rose, Stephan Uhlemann, Eugen Schwan, Bernd Kabius, and Knut Urban. A spherical-aberration-corrected 200 kv transmission electron microscope. *Ultramicroscopy*, 75(1):53–60, 1998.

[33] Philip E Batson, Niklas Dellby, and Ondrej L Krivanek. Sub-ångstrom resolution using aberration corrected electron optics. *Nature*, 418(6898):617–620, 2002.

[34] The Kavli Prize. The 2020 kavli prize in nanoscience. https://www.kavliprize.org/prizes/nanoscience/2020, 2020. Accessed: 2025-12-12.

[35] Jacques Dubochet, Marc Adrian, Jiin-Ju Chang, Jean-Claude Homo, Jean Lepault, Alasdair W McDowall, and Patrick Schultz. Cryo-electron microscopy of vitrified specimens. *Quarterly reviews of biophysics*, 21(2):129–228, 1988.

[36] Richard Henderson, Joyce M Baldwin, Thomas A Ceska, Friedrich Zemlin, Erich Beckmann, and Kenneth H Downing. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *Journal of molecular biology*, 213(4):899–929, 1990.

[37] Richard Henderson. Realizing the potential of electron cryo-microscopy. *Quarterly reviews of biophysics*, 37(1):3–13, 2004.

[38] Joachim Frank. Single-particle reconstruction of biological molecules–story in a sample. *The Nobel Prizes 2017*, page 236, 2020.

[39] Peter W. Hawkes and Erwin Kasper. *Principles of Electron Optics, Volume 1: Basic Geometrical Optics*. Elsevier, Amsterdam, 2 edition, 2017.

[40] David J Smith. Progress and perspectives for atomic-resolution electron microscopy. *Ultramicroscopy*, 108(3):159–166, 2008.

[41] David B. Williams and C. Barry Carter. *Transmission Electron Microscopy: A Textbook for Materials Science*. Springer, 2009.

[42] Viraht Sahni. *Schrödinger Theory of Electrons: Complementary Perspectives*, volume 285. Springer Nature, 2022.

[43] Ludwig Reimer and Helmut Kohl. *Transmission Electron Microscopy: Physics of Image Formation*. Springer, 2008.

[44] Yakir Aharonov and David Bohm. Significance of electromagnetic potentials in the quantum theory. *Physical review*, 115(3):485, 1959.

[45] Simon Wächter. The aharonov-bohm effect. *Proseminar on Algebra, Topology and Group theory in physics, ETH Zürich (March 2018)*, 2018.

[46] Ray F Egerton. *Electron energy-loss spectroscopy in the electron microscope*. Springer Science & Business Media, 2011.

[47] Elisha Huggins. Introduction to fourier optics. *The Physics Teacher*, 45(6):364–368, 2007.

[48] Joseph W. Goodman. *Introduction to Fourier Optics*. Roberts and Company Publishers, Englewood, Colorado, 3rd edition, 2005.

[49] Stephane Perrin and Paul Montgomery. Fourier optics: basic concepts. *arXiv preprint arXiv:1802.07161*, 2018.

[50] Earl J. Kirkland. *Advanced Computing in Electron Microscopy*. Springer, Cham, 3 edition, 2020.

[51] OL Krivanek, N Dellby, and AR Lupini. Towards sub-å electron beams. *Ultramicroscopy*, 78(1-4):1–11, 1999.

[52] M Haider, S Uhlemann, and J Zach. Upper limits for the residual aberrations of a high-resolution aberration-corrected stem. *Ultramicroscopy*, 81(3-4):163–175, 2000.

[53] O Scherzer. The theoretical resolution limit of the electron microscope. *Journal of Applied Physics*, 20(1):20–29, 1949.

[54] Koji Kimoto. Practical aspects of monochromators developed for transmission electron microscopy. *Journal of Electron Microscopy*, 63(5):337–344, 2014.

[55] Kevin L Jensen. A tutorial on electron sources. *IEEE Transactions on Plasma Science*, 46(6):1881–1899, 2018.

[56] Saul Dushman. Electron emission from metals as a function of temperature. *Physical Review*, 21(6):623, 1923.

[57] Peter W. Hawkes and Erwin Kasper. *Principles of Electron Optics, Volume Two: Applied Geometrical Optics*. Elsevier, 2 edition, 2018.

[58] Giuseppe Mario Caruso, Florent Houdellier, Pierre Abeilhou, and Arnaud Arbouet. Development of an ultrafast electron source based on a cold-field emission gun for ultrafast coherent tem. *Applied Physics Letters*, 111(2), 2017.

[59] Vladimir Shiltsev. Electron lenses: historical overview and outlook. *Journal of Instrumentation*, 16(03):P03039, 2021.

[60] AE Ennos. The origin of specimen contamination in the electron microscope. *British Journal of Applied Physics*, 4(4):101, 1953.

[61] Ted Pella, Inc. Electron microscope aperture overview, n.d. Accessed: 2025-12-08.

[62] Yanhong Chen, Liping Wang, Yicheng Feng, Erjun Guo, Sicong Zhao, and Lei Wang. Effect of ca and sm combined addition on the microstructure and elevated-temperature mechanical properties of mg-6al alloys. *Journal of Materials Engineering and Performance*, 28(5):2892–2902, 2019.

[63] Leonid A Bendersky and Frank W Gayle. Electron diffraction using transmission electron microscopy. *Journal of research of the National Institute of Standards and Technology*, 106(6):997, 2001.

[64] Peter D Nellist. Scanning transmission electron microscopy. In *Springer Handbook of Microscopy*, pages 49–99. Springer, 2019.

[65] Jaume Gázquez, Gabriel Sánchez-Santolino, Neven Biškup, Manuel A Roldán, M Cabero, Stephen J Pennycook, and María Varela. Applications of stem-eels to complex oxides. *Materials Science in Semiconductor Processing*, 65:49–63, 2017.

[66] Ray F Egerton et al. *Physical principles of electron microscopy*, volume 56. Springer, 2005.

[67] Heiko Müller, Stephan Uhlemann, Peter Hartel, and Maximilian Haider. Advancing the hexapole cs-corrector for the scanning transmission electron microscope. *Microscopy and Microanalysis*, 12(6):442–455, 2006.

[68] Harald Rose. Chapter 1 - history of direct aberration correction. In Peter W. Hawkes, editor, *Aberration-Corrected Electron Microscopy*, volume 153 of *Advances in Imaging and Electron Physics*, pages 3–39. Elsevier, 2008.

[69] R Seeliger. Die spharische korrektur von elektronenlinsen mittels nichtrotationssymmetrischer abbildungselemente. *Optik*, 8:311–317, 1951.

[70] Maximilian Haider, Herald Rose, Stephan Uhlemann, Bernd Kabius, and Knut Urban. Towards 0.1 nm resolution with the first spherically corrected transmission electron microscope. *Microscopy*, 47(5):395–405, 1998.

[71] G Mollenstedt. Elektronenmikroskopische bilder mit einem nach o. scherzer spharisch korrigierten objektiv. In *Proc. Int. Conf. on Electr. Micr.*, pages 694–698, 1954.

[72] Maximilian Haider, Gerhard Braunshausen, and Eugen Schwan. Correction of the spherical aberration of a 200 kv tem by means of a hexapole-corrector. *Optik (Stuttgart)*, 99(4):167–179, 1995.

[73] OL Krivanek, GJ Corbin, N Dellby, BF Elston, RJ Keyse, MF Murfitt, CS Own, ZS Szilagyi, and JW Woodruff. An electron microscope for the aberration-corrected era. *Ultramicroscopy*, 108(3):179–195, 2008.

[74] Harald Rose. Outline of a spherically corrected semiaplanatic medium-voltage transmission electron microscope. *Optik*, 85:19–24, 1990.

[75] Maximilian Haider, Stephan Uhlemann, Eugen Schwan, Harald Rose, Bernd Kabius, and Knut Urban. Electron microscopy image enhanced. *Nature*, 392(6678):768–769, 1998.

[76] H Rose. Correction of aberrations, a promising means for improving the spatial and energy resolution of energy-filtering electron microscopes. *Ultramicroscopy*, 56(1-3):11–25, 1994.

[77] Fumio Hosokawa, Hidetaka Sawada, Yukihito Kondo, Kunio Takayanagi, and Kazutomo Suenaga. Development of cs and cc correctors for transmission electron microscopy. *Microscopy*, 62(1):23–41, 2013.

[78] F Zemlin, K Weiss, P Schiske, W Kunath, and K-H Herrmann. Coma-free alignment of high resolution electron microscopes with the aid of optical diffractograms. *Ultramicroscopy*, 3:49–60, 1978.

[79] CEOS GmbH. Corrected electron optical systems. https://www.ceos-gmbh.de/en. Accessed: December 09, 2025.

[80] Rik Brydson and Rik Brydson. *Aberration-corrected analytical transmission electron microscopy*, volume 280. Wiley Online Library, 2011.

[81] Stephan Uhlemann and Maximilian Haider. Residual wave aberrations in the first spherical aberration corrected transmission electron microscope. *Ultramicroscopy*, 72(3-4):109–119, 1998.

[82] Giovanni Bertoni, Enzo Rotunno, Daan Marsmans, Peter Tiemeijer, Amir H Tavabi, Rafal E Dunin-Borkowski, and Vincenzo Grillo. Near-real-time diagnosis of electron optical phase aberrations in scanning transmission electron microscopy using an artificial neural network. *Ultramicroscopy*, 245:113663, 2023.

[83] Andrew R Lupini. The electron ronchigram. *Scanning Transmission Electron Microscopy: Imaging and Analysis*, pages 117–161, 2010.

[84] Noah Schnitzer, Suk Hyun Sung, and Robert Hovden. Introduction to the ronchigram and its calculation with ronchigram. com. *Microscopy Today*, 27(3):12–15, 2019.

[85] H Sawada, T Sannomiya, F Hosokawa, T Nakamichi, T Kaneyama, T Tomita, Y Kondo, T Tanaka, Y Oshima, Y Tanishiro, et al. Measurement method of aberration from ronchigram by autocorrelation function. *Ultramicroscopy*, 108(11):1467–1475, 2008.

[86] Andrew R Lupini, Peng Wang, Peter D Nellist, Angus I Kirkland, and Stephen J Pennycook. Aberration measurement using the ronchigram contrast transfer function. *Ultramicroscopy*, 110(7):891–898, 2010.

[87] Takumi Sannomiya, Hidetaka Sawada, Tomohiro Nakamichi, Fumio Hosokawa, Yoshio Nakamura, Yasumasa Tanishiro, and Kunio Takayanagi. Determination of aberration center of ronchigram for automated aberration correctors in scanning transmission electron microscopy. *Ultramicroscopy*, 135:71–79, 2013.

[88] Svetlana N Khonina, Nikolay L Kazanskiy, Roman V Skidanov, and Muhammad A Butt. Advancements and applications of diffractive optical elements in contemporary optics: A comprehensive overview. *Advanced Materials Technologies*, 10(4):2401028, 2025.

[89] ZS Sacks, D Rozas, and GA Swartzlander Jr. Holographic formation of optical-vortex filaments. *Journal of the Optical Society of America B*, 15(8):2226–2234, 1998.

[90] Yiqian Yang, Andrew Forbes, and Liangcai Cao. A review of liquid crystal spatial light modulators: devices and applications. *Opto-Electronic Science*, 2(8):230026–1, 2023.

[91] Robert K Tyson and Benjamin West Frazier. *Principles of adaptive optics*. CRC press, 2022.

[92] Nanfang Yu and Federico Capasso. Flat optics with designer metasurfaces. *Nature materials*, 13(2):139–150, 2014.

[93] Marek Malac, Simon Hettler, Misa Hayashida, Emi Kano, Ray F Egerton, and Marco Beleggia. Phase plates in the transmission electron microscope: operating principles and applications. *Microscopy*, 70(1):75–115, 2021.

[94] Kuniaki Nagayama. Development of phase plates for electron microscopes and their biological application. *European Biophysics Journal*, 37(4):345–358, 2008.

[95] Kuniaki Nagayama and Radostin Danev. Phase contrast electron microscopy: development of thin-film phase plates and biological applications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1500):2153–2162, 2008.

[96] Frits Zernike. Phase contrast, a new method for the microscopic observation of transparent objects. *Physica*, 9(7):686–698, 1942.

[97] Hans Boersch. Über die kontraste von atomen im elektronenmikroskop. *Zeitschrift für Naturforschung A*, 2(11-12):615–633, 1947.

[98] AH Tavabi, P Rosi, A Roncaglia, E Rotunno, M Beleggia, P-H Lu, L Belsito, G Pozzi, S Frabboni, P Tiemeijer, et al. Generation of electron vortex beams with over 1000 orbital angular momentum quanta using a tunable electrostatic spiral phase plate. *Applied Physics Letters*, 121(7), 2022.

[99] Paolo Rosi, Federico Venturi, Giacomo Medici, Claudia Menozzi, Gian Carlo Gazzadi, Enzo Rotunno, Stefano Frabboni, Roberto Balboni, Mohammadreza Rezaee, Amir H Tavabi, et al. Theoretical and practical aspects of the design and production of synthetic holograms for transmission electron microscopy. *Journal of Applied Physics*, 131(3), 2022.

[100] Jeremy J Axelrod, Jessie T Zhang, Petar N Petrov, Robert M Glaeser, and Holger Mueller. Modern approaches to improving phase contrast electron microscopy. *Current Opinion in Structural Biology*, 86:102805, 2024.

[101] Robert M Glaeser. Invited review article: Methods for imaging weak-phase objects in electron microscopy. *Review of Scientific Instruments*, 84(11), 2013.

[102] Simon Hettler, Manuel Dries, Peter Hermann, Martin Obermair, Dagmar Gerthsen, and Marek Malac. Carbon contamination in scanning transmission electron microscopy and its impact on phase-plate applications. *Micron*, 96:38–47, 2017.

[103] Simon Hettler, Emi Kano, Manuel Dries, Dagmar Gerthsen, Lukas Pfaffmann, Michael Bruns, Marco Beleggia, and Marek Malac. Charging of carbon thin films in scanning and phase-plate transmission electron microscopy. *Ultramicroscopy*, 184:252–266, 2018.

[104] Radostin Danev, Robert M Glaeser, and Kuniaki Nagayama. Practical factors affecting the performance of a thin-film phase plate for transmission electron microscopy. *Ultramicroscopy*, 109(4):312–325, 2009.

[105] Hong-Wei Wang and Xiao Fan. Challenges and opportunities in cryo-em with phase plate. *Current Opinion in Structural Biology*, 58:175–182, 2019.

[106] Marek Malac, Marco Beleggia, Ray Egerton, and Yimei Zhu. Imaging of radiation-sensitive samples in transmission electron microscopes equipped with zernike phase plates. *Ultramicroscopy*, 108(2):126–140, 2008.

[107] Radostin Danev and Kuniaki Nagayama. Transmission electron microscopy with zernike phase plate. *Ultramicroscopy*, 88(4):243–252, 2001.

[108] Rebecca Pretzsch, Manuel Dries, Simon Hettler, Martin Spiecker, Martin Obermair, and Dagmar Gerthsen. Investigation of hole-free phase plate performance in transmission electron microscopy under different operation conditions by experiments and simulations. *Advanced Structural and Chemical Imaging*, 5(1):5, 2019.

[109] Marek Malac, Marco Beleggia, Masahiro Kawasaki, Peng Li, and Ray F Egerton. Convenient contrast enhancement by a hole-free phase plate. *Ultramicroscopy*, 118:77–89, 2012.

[110] Radostin Danev, Bart Buijsse, Maryam Khoshouei, Jürgen M Plitzko, and Wolfgang Baumeister. Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proceedings of the National Academy of Sciences*, 111(44):15635–15640, 2014.

[111] Maryam Khoshouei, Stefan Pfeffer, Wolfgang Baumeister, Friedrich Förster, and Radostin Danev. Subtomogram analysis using the volta phase plate. *Journal of structural biology*, 197(2):94–101, 2017.

[112] Shoh Asano, Yoshiyuki Fukuda, Florian Beck, Antje Aufderheide, Friedrich Förster, Radostin Danev, and Wolfgang Baumeister. A molecular census of 26 s proteasomes in intact neurons. *Science*, 347(6220):439–442, 2015.

[113] Yasuko Kaneko, Radostin Danev, Koji Nitta, and Kuniaki Nagayama. In vivo subcellular ultrastructures recognized with hilbert differential contrast transmission electron microscopy. *Journal of electron microscopy*, 54(1):79–84, 2005.

[114] Bastian Barton, Friederike Joos, and Rasmus R Schröder. Improved specimen reconstruction by hilbert phase contrast tomography. *Journal of structural biology*, 164(2):210–220, 2008.

[115] Tyler R Harvey, Jordan S Pierce, Amit K Agrawal, Peter Ercius, Martin Linck, and Benjamin J McMorran. Efficient diffractive phase optics for electrons. *New Journal of Physics*, 16(9):093039, 2014.

[116] Dolev Roitman, Roy Shiloh, Peng-Han Lu, Rafal E Dunin-Borkowski, and Ady Arie. Shaping of electron beams using sculpted thin films. *ACS photonics*, 8(12):3394–3405, 2021.

[117] Roy Shiloh, Roei Remez, Peng-Han Lu, Lei Jin, Yossi Lereah, Amir H Tavabi, Rafal E Dunin-Borkowski, and Ady Arie. Spherical aberration correction in a scanning transmission electron microscope using a sculpted thin film. *Ultramicroscopy*, 189:46–53, 2018.

[118] Martin Linck, Peter A Ercius, Jordan S Pierce, and Benjamin J McMorran. Aberration corrected stem by means of diffraction gratings. *Ultramicroscopy*, 182:36–43, 2017.

[119] Takao Matsumoto and Akira Tonomura. The phase constancy of electron waves traveling through boersch's electrostatic phase plate. *Ultramicroscopy*, 63(1):5–10, 1996.

[120] Andreas Walter, Siegfried Steltenkamp, Sam Schmitz, Peter Holik, Edvinas Pakanavicius, Roland Sachser, Michael Huth, Daniel Rhinow, and Werner Kühlbrandt. Towards an optimum design for electrostatic phase plates. *Ultramicroscopy*, 153:22–31, 2015.

[121] E Majorovits, Bastian Barton, K Schultheiss, F Perez-Willard, D Gerthsen, and Rasmus R Schröder. Optimizing phase contrast in transmission electron microscopy with an electrostatic (boersch) phase plate. *Ultramicroscopy*, 107(2-3):213–226, 2007.

[122] Pooja Thakkar, Vitaliy A Guzenko, Peng-Han Lu, Rafal E Dunin-Borkowski, Jan Pieter Abrahams, and Soichiro Tsujino. Fabrication of low aspect ratio three-element boersch phase shifters for voltage-controlled three electron beam interference. *Journal of Applied Physics*, 128(13), 2020.

[123] Nicole Frindt, Marco Oster, Simon Hettler, Björn Gamm, Levin Dieterle, Wolfgang Kowalsky, Dagmar Gerthsen, and Rasmus R Schröder. In-focus electrostatic zach phase plate imaging for transmission electron microscopy with tunable phase contrast of frozen hydrated biological samples. *Microscopy and Microanalysis*, 20(1):175–183, 2014.

[124] Simon Hettler, Björn Gamm, Manuel Dries, Nicole Frindt, Rasmus R Schröder, and Dagmar Gerthsen. Improving fabrication and application of zach phase plates for phase-contrast transmission electron microscopy. *Microscopy and Microanalysis*, 18(5):1010–1015, 2012.

[125] J. Zach. Phase plate, image producing method and electron microscope. PCT patent WO 2008/061603 (WIPO), 2008.

[126] Simon Hettler, Manuel Dries, J Zeelen, Marco Oster, Rasmus R Schröder, and Dagmar Gerthsen. High-resolution transmission electron microscopy with an electrostatic zach phase plate. *New Journal of Physics*, 18(5):053005, 2016.

[127] Konstantin Yu Bliokh, Yury P Bliokh, Sergey Savel'Ev, and Franco Nori. Semiclassical dynamics of electron wave packet states with phase vortices. *Physical Review Letters*, 99(19):190404, 2007.

[128] Konstantin Y Bliokh, Igor P Ivanov, Giulio Guzzinati, Laura Clark, Ruben Van Boxem, Armand Béché, Roeland Juchtmans, Miguel A Alonso, Peter Schattschneider, Franco Nori, et al. Theory and applications of free-electron vortex states. *Physics Reports*, 690:1–70, 2017.

[129] Les Allen, Marco W Beijersbergen, RJC Spreeuw, and JP Woerdman. Orbital angular momentum of light and the transformation of laguerre-gaussian laser modes. *Physical review A*, 45(11):8185, 1992.

[130] OV Angelsky, A Ya Bekshaev, MV Vasnetsov, C Yu Zenkova, PP Maksimyak, and Jun Zheng. Optical phase singularities: Physical nature, manifestations and applications. *Frontiers in Physics*, 10:1060787, 2022.

[131] Sophia Marriot Lloyd, Mohamed Babiker, Gnanavel Thirunavukkarasu, and Jun Yuan. Electron vortices: Beams with orbital angular momentum. *Reviews of Modern Physics*, 89(3):035004, 2017.

[132] Vincenzo Grillo, Tyler R Harvey, Federico Venturi, Jordan S Pierce, Roberto Balboni, Frédéric Bouchard, Gian Carlo Gazzadi, Stefano Frabboni, Amir H Tavabi, Zi-An Li, et al. Observation of nanoscale magnetic fields using twisted electron beams. *Nature communications*, 8(1):689, 2017.

[133] Alexander Edström, Axel Lubk, and Ján Rusz. Elastic scattering of electron vortex beams in magnetic matter. *Physical Review Letters*, 116(12):127203, 2016.

[134] Giulio Pozzi, Paolo Rosi, Amir H Tavabi, Ebrahim Karimi, Rafal E Dunin-Borkowski, and Vincenzo Grillo. A sorter for electrons based on magnetic elements. *Ultramicroscopy*, 231:113287, 2021.

[135] Amir H Tavabi, Paolo Rosi, Enzo Rotunno, Alberto Roncaglia, Luca Belsito, Stefano Frabboni, Giulio Pozzi, Gian Carlo Gazzadi, Peng-Han Lu, Robert Nijland, et al. Experimental demonstration of an electrostatic orbital angular momentum sorter for electron beams. *Physical review letters*, 126(9):094802, 2021.

[136] Benjamin J McMorran, Amit Agrawal, Peter A Ercius, Vincenzo Grillo, Andrew A Herzing, Tyler R Harvey, Martin Linck, and Jordan S Pierce. Origins and demonstrations of electrons with orbital angular momentum. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2087):20150434, 2017.

[137] Benjamin J McMorran, Amit Agrawal, Ian M Anderson, Andrew A Herzing, Henri J Lezec, Jabez J McClelland, and John Unguris. Electron vortex beams with high quanta of orbital angular momentum. *science*, 331(6014):192–195, 2011.

[138] Giulio Pozzi, Peng-Han Lu, Amir H Tavabi, Martial Duchamp, and Rafal E Dunin-Borkowski. Generation of electron vortex beams using line charges via the electrostatic aharonov-bohm effect. *Ultramicroscopy*, 181:191–196, 2017.

[139] Amir H Tavabi, Hugo Larocque, Peng-Han Lu, Martial Duchamp, Vincenzo Grillo, Ebrahim Karimi, Rafal E Dunin-Borkowski, and Giulio Pozzi. Generation of electron vortices using nonexact electric fields. *Physical Review Research*, 2(1):013185, 2020.

[140] Giulio Pozzi, Vincenzo Grillo, Peng-Han Lu, Amir H Tavabi, Ebrahim Karimi, and Rafal E Dunin-Borkowski. Design of electrostatic phase elements for sorting the orbital angular momentum of electrons. *arXiv preprint arXiv:1906.02344*, 2019.

[141] Chu-Ping Yu, Francisco Vega Ibañez, Armand Béché, and Johan Verbeeck. Quantum wavefront shaping with a 48-element programmable phase plate for electrons. *SciPost Physics*, 15(6):223, 2023.

[142] Jo Verbeeck, Armand Béché, Knut Müller-Caspary, Giulio Guzzinati, Minh Anh Luong, and Martien Den Hertog. Demonstration of a 2× 2 programmable phase plate for electrons. *Ultramicroscopy*, 190:58–65, 2018.

[143] Akira Tonomura, Nobuyuki Osakabe, Tsuyoshi Matsuda, Takeshi Kawasaki, Junji Endo, Shinichiro Yano, and Hiroji Yamada. Evidence for aharonov-bohm effect with magnetic field completely shielded from electron wave. *Physical review letters*, 56(8):792, 1986.

[144] AM Blackburn and JC Loudon. Vortex beam production and contrast enhancement from a magnetic spiral phase plate. *Ultramicroscopy*, 136:127–143, 2014.

[145] Armand Béché, Ruben Van Boxem, Gustaaf Van Tendeloo, and Jo Verbeeck. Magnetic monopole field exposed by electrons. *Nature Physics*, 10(1):26–29, 2014.

[146] CJ Edgcombe, A Ionescu, JC Loudon, AM Blackburn, H Kurebayashi, and CHW Barnes. Characterisation of ferromagnetic rings for zernike phase plates using the aharonov–bohm effect. *Ultramicroscopy*, 120:78–85, 2012.

[147] Amir H Tavabi, Marco Beleggia, Vadim Migunov, Alexey Savenko, Ozan Öktem, Rafal E Dunin-Borkowski, and Giulio Pozzi. Tunable ampere phase plate for low dose imaging of biomolecular complexes. *Scientific reports*, 8(1):5592, 2018.

[148] Radostin Danev and Kuniaki Nagayama. Phase plates for transmission electron microscopy. In *Methods in enzymology*, volume 481, pages 343–369. Elsevier, 2010.

[149] F Javier García de Abajo and Andrea Konečná. Optical modulation of electron beams in free space. *Physical Review Letters*, 126(12):123901, 2021.

[150] Holger Mueller, Jian Jin, Radostin Danev, John Spence, Howard Padmore, and Robert M Glaeser. Design of an electron microscope phase plate using a focused continuous-wave laser. *New journal of physics*, 12(7):073011, 2010.

[151] Franz Mandl and Graham Shaw. *Quantum field theory*. John Wiley & Sons, 2013.

[152] William Kruer. *The physics of laser plasma interactions*. crc Press, 2019.

[153] KAH van Leeuwen, Walter J Schaap, B Buijsse, Simona Borrelli, Stefan T Kempers, Wouter Verhoeven, and OJ Luiten. Feasibility of a pulsed ponderomotive phase plate for electron beams. *New Journal of Physics*, 25(2):023031, 2023.

[154] Daniel X Du and Anthony WP Fitzpatrick. Design of an ultrafast pulsed ponderomotive phase plate for cryo-electron tomography. *Cell Reports Methods*, 3(1), 2023.

[155] Beatrice Matilde Ferrari, Cameron James Richard Duncan, Michael Yannai, Raphael Dahan, Paolo Rosi, Irene Ostroman, Maria Giulia Bravi, Arthur Niedermayr, Tom Lenkiewicz Abudi, Yuval Adiv, et al. Realization of a pre-sample photonic-based free-electron modulator in ultrafast transmission electron microscopes. *ACS photonics*, 12(11):5864–5873, 2025.

[156] Andrea Konečná and F Javier García de Abajo. Electron beam aberration correction using optical near fields. *Physical Review Letters*, 125(3):030801, 2020.

[157] Osip Schwartz, Jeremy J Axelrod, Sara L Campbell, Carter Turnbaugh, Robert M Glaeser, and Holger Müller. Laser phase plate for transmission electron microscopy. *Nature methods*, 16(10):1016–1020, 2019.

[158] Ori Reinhardt and Ido Kaminer. Theory of shaping electron wavepackets with light. *ACS Photonics*, 7(10):2859–2870, 2020.

[159] Wen-Ming Zhang, Guang Meng, and DI Chen. Stability, nonlinearity and reliability of electrostatically actuated mems devices. *Sensors*, 7(5):760–796, 2007.

[160] Yun-Yu Wang, Kenton Burns, Shixin Wang, Zhouguang Wang, Hongqi Li, Mojtaba Asadirad, Mike Smith, and Qiang Jin. Direct observation of trap charge in sio2 by electron holography and differential phase contrast (dpc). *Microscopy Today*, 33(5):47–54, 11 2025.

[161] P Goodman and Alexander Forbes Moodie. Numerical evaluations of n-beam wave functions in electron scattering by the multi-slice method. *Foundations of Crystallography*, 30(2):280–290, 1974.

[162] Giulio Pozzi, Vincenzo Grillo, Peng-Han Lu, Amir H Tavabi, Ebrahim Karimi, and Rafal E Dunin-Borkowski. Design of electrostatic phase elements for sorting the orbital angular momentum of electrons. *Ultramicroscopy*, 208:112861, 2020.

[163] Benjamin J McMorran, Tyler R Harvey, and Martin PJ Lavery. Efficient sorting of free electron orbital angular momentum. *New Journal of Physics*, 19(2):023053, 2017.

[164] Donald A. Neamen. *Semiconductor Physics and Devices*. McGraw-Hill Education, New York, 4th edition, 2012.

[165] M Tamura, S Shukuri, M Moniwa, and M Default. Focused ion beam gallium implantation into silicon. *Applied Physics A*, 39(3):183–190, 1986.

[166] Thermo Fisher Scientific. Nanoex-i/v tem holder: Datasheet, 2025. Technical datasheet.

[167] Marco Beleggia, MA Schofield, VV Volkov, and Y Zhu. On the transport of intensity technique for phase retrieval. *Ultramicroscopy*, 102(1):37–49, 2004.

[168] Zhen Chen, Yi Jiang, Yu-Tsun Shao, Megan E Holtz, Michal Odstrčil, Manuel Guizar-Sicairos, Isabelle Hanke, Steffen Ganschow, Darrell G Schlom, and David A Muller. Electron ptychography achieves atomic-resolution limits set by lattice vibrations. *Science*, 372(6544):826–831, 2021.

[169] Stephanie M Ribet, Steven E Zeltmann, Georgios Varnavides, Roberto Dos Reis, Vinayak P Dravid, and Colin Ophus. Phase diversity in ptychographic reconstructions with a programmable phase plate, 2023.

[170] P Rosi, L Viani, E Rotunno, S Frabboni, AH Tavabi, RE Dunin-Borkowski, A Roncaglia, and V Grillo. Increasing the resolution of transmission electron microscopy by computational ghost imaging. *Physical Review Letters*, 133(12):123801, 2024.