Quantitative imaging and alternative proposals on image simulation and reconstruction in atomic resolution transmission electron microscopy

A Thesis Submitted for the Degree of

Doctor of Philosophy

in the Faculty of Science

by

Usha Manjunath Bhat

CHEMISTRY AND PHYSICS OF MATERIALS UNIT
INTERNATIONAL CENTRE FOR MATERIALS SCIENCE
JAWAHARLAL NEHRU CENTRE FOR ADVANCED SCIENTIFIC RESEARCH
(A Deemed University)
Bangalore – 560064, INDIA.
April 2021

Dedicated to

My Family

DECLARATION

I hereby declare that the thesis entitled "**Quantitative imaging and alternative proposals on image simulation and reconstruction in atomic resolution transmission electron microscopy**" is an authentic record of research work carried out by me at the Chemistry and Physics of Materials Unit, International Centre for Materials Science, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India under the supervision of **Prof. Ranjan Datta** and that it has not been submitted elsewhere for the award of any degree or diploma.

In keeping with the general practice in reporting scientific observations, due acknowledgment has been made whenever the work described is based on the findings of other investigators. Any omission that might have occurred due to oversight or error in judgment is regretted.

Usha Manjunath Bhat

CERTIFICATE

Certified that the work described in this thesis titled "**Quantitative imaging and alternative proposals on image simulation and reconstruction in atomic resolution transmission electron microscopy**" has been carried out by Mrs. **Usha Manjunath Bhat** at the Chemistry and Physics of Materials Unit, International Centre for Materials Science, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India under my supervision and that it has not been submitted elsewhere for the award of any degree or diploma.

Prof. Ranjan Datta

(Research Supervisor)

Acknowledgments

I take this opportunity to sincerely thank my research supervisor, Prof. Ranjan Datta, for his constant motivation, guidance, and encouragement throughout my Ph.D. His endless enthusiasm for science and hardworking nature has inspired and motivated me at various times.

I sincerely thank Prof. C. N. R. Rao for creating world-class facilities at International Centre for Materials Science (ICMS). I specially thank him for providing an aberration-corrected transmission electron microscope facility at this center that gave me an excellent opportunity to learn and carry out research on advanced microscopy techniques.

I would like to express my gratitude to Prof. Arunava Gupta, Prof. Chandrabhas Narayana, and Prof. Sebastian Peter for fruitful collaborations.

I would like to thank all the researchers Dr. P. Silwal, Dr. U.Palchoudhury, Dr. Shreya Sarkar, Dr. N. M. Batra, Dr. M. B. Sreedhara, Dr. V. Rajaji, Dr. S. C. Sarma, and Dr. S. Roy, for collaborating and for their valuable discussions.

I would like to take this opportunity to thank Prof. Sebastian Peter, Prof. N.S.Vidhyadhiraja, Prof. Umesh V. Waghmare, and Prof. Shobhana Narasimhan, Prof. S. M. Shivaprasad for the excellent teaching and the courses.

I thank the past and present chairman of the Chemistry and Physics of Materials Unit, Prof. Chandrabhas Narayana, and Prof. A Sundaresan for their invaluable support.

I would like to thank the entire faculty at JNCASR for being a constant source of inspiration with their advanced research. I take this opportunity to thank all my past teachers at the school and college level for sharing their knowledge, providing their support and encouragement, without which it is impossible for me to be here.

I take this opportunity to thank my past and present labmates, Dr. Loukya, Dr. Dileep, Dr. D. S. Negi, Dr. Rajib, Dr. Badri, Ms. Sharona, Mr. Rajendra, Mr. Ankit Sharma, and Mr. Arindam, for their useful discussions and for creating a cheerful environment.

I extend my gratitude to all the efficient staff at the Library, Academic and Administration sections, and non-academic staff for their timely help and cooperation.

I would like to thank the Hostel staff, Mess workers for providing food and accommodation during my stay at JNCASR.

I thank all my friends inside and outside JNCASR for their immense support.

I sincerely thank all my family members for their unconditional love, support and encouragement. I am lucky to have wonderful parents Manjunath Bhat and Sumangala Bhat, who has been the pillar of my strength and guiding me during difficult times. I especially thank my sibling, Mr. Chinmay, for being a constant support, and my grandparents Mrs. Lakshmi and Mr. Narasimha Shastri, for unconditional love.

I am grateful to my husband, Mr. Sumant Hegde, for his immense support, love, and encouragement.

Synopsis

Quantitative imaging is an active area of research in atomic resolution transmission electron microscopy. Both aberration-corrected high-resolution transmission electron microscopy (HRTEM) and atomic resolution off-axis electron holography provide a unique opportunity to study phase information at the atomic and sub-atomic length scale. These methods can be utilized to extract a wide range of information on materials, e.g., atom types, distribution, electronic bonding, etc. The thesis starts with an introduction to aberration-corrected atomic resolution HRTEM and off-axis electron holography techniques.

In chapter 2, two different atomic resolution techniques: Off-axis electron holography and HRTEM, have been employed to count atoms on individual columns of Zn and O in ZnO epitaxial thin film. Results show that the reconstructed phase from both the side and the central bands and the corresponding number of Zn (Z = 30) and O (Z = 8) atoms are in close agreement with the systematic increase in the number of atoms for a sample area less than the extinction distance. However, complete disagreement is observed for the sample area more than the extinction distance. On the other hand, the reconstructed phase obtained via in-line holography shows no systematic change with thickness for the same sample. Phase detection limits and the atomic model used to count the atoms are also presented.

In chapter 3, an alternative approach to the image simulation in HRTEM is introduced after a comparative analysis of the existing image simulation methods. The alternative method is based on considering the atom center as an electrostatic interferometer akin to the conventional off-axis electron biprism within few nanometers of focus variation. Simulation results are compared with the experimental images of 2D materials of MoS_2 and BN recorded under the optimum combination of third-order spherical aberration $C_s = -35$ μm and defocus $\Delta f = 1$, 4, and 8 nm and are found to be in good agreement.

In chapter 4, An alternative reconstruction method is proposed for retrieving the object exit wave function (OEW) directly from the recorded image intensity pattern in HRTEM. The method is based on applying a modified intensity equation representing the HRTEM image. A comparative

discussion is provided between the existing methodologies of the reconstruction of OEW in HRTEM, off-axis electron holography, and the present proposal. Phase shift extracted from the experimental images of MoS_2, BN, and ZnO are found to be in excellent agreement with the theoretical reference values. Additionally, it is shown that the Fourier series expansion of diffraction pattern is effective in retrieving the isolated and periodic image functions of a specific form. However, for aperiodic object information, e.g., defects, dopants, edges, etc., the first method works its entirety. A future perspective is provided based on the results presented in Chapter 5 of the thesis.

List of Publications

1. **U. Bhat**, and R. Datta. "Image simulation in high resolution transmission electron microscopy considering atom as an electrostatic charge center". *Journal of Physics Communication, https://doi.org/10.1088/2399-6528/ac1839 (2021).*
2. **U. Bhat**, and R. Datta. "Alternative proposals on the phase retrieval in high resolution transmission electron microscopy". *arXiv:2103.00538 (2021).*
3. **U. Bhat**, and R. Datta. "Quantitative counting of Zn and O atoms by atomic resolution off-axis and in-line electron holography". *Journal of Applied Physics 125, 154902 (2019).*
4. **U. Bhat**, R. Singh, B. Vishal, A. Sharma, S. Horta, R. Sahu, and R. Datta. "Distinct Photoluminescence in Multilayered van der Waals Heterostructures of $MoS_2/WS_2/ReS_2$ and BN". *Phys. Status Solidi B 1700691(2018).*
5. B. Vishal, **U. Bhat**, H. Sharona, R. Datta, "Heterostructures of hetero-stack of 2D TMDs (MoS_2, WS_2, and ReS_2) and BN". *arXiv:2009.11495[cond-mat.mtrl-sci] (2020).*
6. B. Vishal, **U. Bhat**, H. Sharona, A. Mukherjee, S. Roy, S.C. Peter and R. Datta. "Microstructural and magnetic properties of epitaxial $Ni_{50}Mn_{37/35}Sn_{13/15}$ Heusler alloy thin films grown by pulsed laser deposition". *Journal of Crystal Growth 546, 125772 (2020).*
7. H. Sharona, B. Vishal, **U. Bhat**, A. Paul, A. Mukherjee, S.C. Sarma, S.C. Peter and R. Datta. "Rich diversity of crystallographic phase formation in 2D $Re_xMo_{1-x}S_2$ ($x<0.5$) alloy". *Journal of Applied Physics 126, 224302 (2019).*
8. N. S. Farheen, M. B. Sreedhara, A. Soni, **U. Bhat**, R. Datta, A. J. Bhattacharyya, and C. N. R. Rao. "Li and Na-ion diffusion and intercalation characteristics in vertically aligned TiS_2 nanowall network grown using atomic layer deposition". *Mater. Res. Express 6, 115549 (2019).*
9. B. Vishal, H. Sharona, **U. Bhat**, A. Paul, M.B. Shreedhara, V. Rajaji, S.C. Sarma, C. Narayana, S.C. Peter and R. Datta."Growth of ReS_2 thin films by pulsed laser deposition". *Thin Solid Films 685, 81 (2019).*

10. S. V. Amit, S. Abhishek, J.B. Mohammadi, S.Regmi, **U. Bhat**, R. Datta, T. Mewes, *and* A. Gupta. "Studies of electrical and magnetic properties across the Verwey transition in epitaxial magnetite thin films". *Journal of Applied Physics 126, 093902 (2019).*
11. B. Vishal, R. Singh, A. Chaturvedi, A. Sharma, M.B. Sreedhara, R. Sahu, **U. Bhat**, U. Ramamurty, and R. Datta. "Chemically stabilized epitaxial wurtzite-BN thin film". *Superlattices and Microstructures 115, 197 (2018).*
12. H. Sharona, B. Loukya, **U. Bhat**, R. Sahu, B. Vishal, P. Silwal, A. Gupta, and R. Datta. "Coexisting nanoscale inverse spinel and rock salt crystallographic phases in NiCo2O4 epitaxial thin films grown by pulsed laser deposition". *Journal of Applied Physics 122, 225301 (2017).*
13. Negi D. S., Sharona H., **U. Bhat**, Palchoudhury S., Gupta A., & Datta R. "Surface spin canting in Fe3O4 and CoFe2O4 nanoparticles probed by high resolution electron energy loss spectroscopy". *Physical Review B 95, 174444 (2017).*
14. Sahu, R., **U. Bhat**, Batra, N.M., Sharona, H., Vishal, B., Sarkar, S., Assa Aravindh, S., Peter, S.C., Roqan, I.S., Costa, P.M.F.J. and Datta, R. "Nature of low dimensional structural modulations and relative phase stability in $Re_xMo(W)_{1-x}S_2$ transition metal dichalcogenide alloys". *Journal of Applied Physics, 121, 105101 (2017).*
15. B. Vishal, R. Sahu, **U. Bhat** and R. Datta. "Investigation of microstructural details in low thermal conducting thermoelectric $Sn_{1-x}Sb_xTe$ alloy". *Journal of Applied Physics, 122, 055102 (2017).*

Contents

Acknowledgments v

Synopsis ...……vii

List of Publications ……...……...ix

1. Introduction……..…….……1

1.1 Phase-contrast microscopy……....…….……....2

1.2 Aberration corrected high-resolution transmission electron microscopy ……..…....…3

1.2.1 Concept of resolution ……....3

1.2.2 Principle of image formation ……..……...5

1.2.3 Aberration correction in TEM ……..6

1.3 Electron holography: A brief introduction...……..……..…….....9

1.3.1 In-line electron holography ……..9

1.3.2 Off-axis electron holography at medium resolution ……9

1.3.3 Off-axis electron holography at atomic resolution ……11

1.4 Quantitative imaging ……..…….....13

1.5 Overview of the thesis ……15

1.6 Bibliography….....16

2. Quantitative counting of Zn and O atoms by atomic resolution off-axis and in-line electron holography ……..……..…19

2.1 Introduction ……...…….……....20

2.2 Experimental details and data analysis …….…….....21

2.2.1 Thin film crystal growth method ……21

2.2.2 TEM sample preparation ……22

2.2.3 Off-axis electron holography ……..22

2.2.4 In-line electron holography ……26
2.3 Results and Discussion ...……..…….....28
2.3.1 Phase detection limit …….....28
2.3.2 Atomic potential model …….32
2.3.3 Atomic resolution off-axis electron holography...……..……...38
2.3.4 In-line electron holography (HRTEM) …….....42
2.4 Conclusion ...…….……...43
2.5 Bibliography….....43

3. Image simulation in HRTEM considering atom as an electrostatic charge center …….47
3.1 Introduction ……..……...48
3.2 Coherent mage formation at near and far field ……..…….…..49
3.2.1 Fresnel diffraction pattern ……..50
3.2.2 Frauhofer diffraction pattern ……..53
3.2.3 Fourier transform-based method ……54
3.2.4 Abbe's theory of imaging ……..58
3.3 Existing method of image simulation ...…….…….....59
3.3.1 Image simulation based on Zernike phase object and WPOA …..…….....…….…...60
3.3.2 Image simulation based on scattering factor ……..…….....…….…...67
3.4 New method: Atom as an electrostatic charge center …..……...…….73
3.4.1 Image intensity pattern in off-axis electron holography …..…….....…….…...75
3.4.2 Image contrast in off-axis electron holography with electron biprism …....…….…...75
3.4.3 Image contrast due to atom as a charge center equivalent to biprism …....…….…...76
3.4.4 Effect of spherical aberration and defocus on the image ……...….....…….…...84
3.5 Conclusion ……...……89
3.6 Bibliography ……..…….…….89

4. Insights and alternative methods on the phase retrieval in HRTEM ……..……93
4.1 Introduction ……94
4.2 Experimental techniques …...…..…….……...96

4.3 Existing methods of exit wave reconstruction ……..…….…….96
4.3.1 Various schemes on HRTEM focal series reconstruction ……..97
4.3.1.1 Focal variation method ……..97
4.3.1.2 Paraboloid method ……...101
4.3.1.3 Partial coherence theory ……...102
4.3.1.4 Partial coherence theory-based formulations ……...103
4.3.1.5 Schiske filter function ……..……...103
4.3.2 Reconstruction scheme in off-axis electron holography ……..104
4.3.3 Comparison between the HRTEM and off-axis electron holography reconstruction ..107
4.4 Alternative methods of phase retrieval …….……112
4.4.1 Recovering phase from the HRTEM image intensity pattern …….112
4.4.2 Recovering image function from diffracted intensity ……..……121
4.5 Conclusion …….……126
4.6 Bibliography …….…….……126

5. Summary and Future perspective ……129
5.1 Summary of the thesis ……....……...130
5.2 Future perspectives …..…….....131
5.3 Bibliography …….…….……134

Chapter 1

Introduction

This chapter reviews the basic principles behind atomic resolution phase-contrast imaging by two different techniques: HRTEM and off-axis electron holography. The principle of aberration correction and its role in improving resolution and contrast is provided. The general applications in materials science and its usefulness at the extreme levels to understand materials by these techniques are outlined. The brief background provided in this chapter will serve as a guide for the results and contributions made in subsequent chapters in terms of atom counting in ZnO in chapter 2, alternative methods of image simulation, and reconstructions in high-resolution phase-contrast transmission electron microscopy in chapter 3 and 4, respectively.

1.1 Phase-contrast microscopy

Based on transmission electron microscopy (TEM), numerous imaging and analytical methods are used in various areas of physical, chemical, and biological sciences. As the name suggests, the fast probe electrons transmit through the electron transparent specimen and carry information about the structure of the sample in the form of an image or diffraction pattern. There are various imaging and spectroscopic techniques available in TEM to investigate materials properties at the atomic and nanometer length scale. For example, diffraction/strain contrast imaging avails information on image dislocations and other types of defects in crystals, phase contrast, and Z-contrast microscopy to resolve material structure at the atomic level. For weak scattering objects, phase-contrast techniques come in handy. Phase variation of the incident beam itself is not visible but becomes visible when shown as a contrast variation. Zernike discovered phase contrast in the year 1933 in the optical microscope. He separated direct and diffracted light from biological tissue using a special disk in the condenser [1]. He increased the phase difference between the direct and diffracted light using a special plate in the back focal plane of the objective lens. This resulted in the enhanced phase-contrast in the intensity pattern without any chemical staining. He received the Nobel prize in 1953 for this work.

High-resolution transmission electron microscopy (HRTEM) or phase contrast TEM is the most widely used technique to reveal the atomic organization in crystalline materials along high symmetry orientations at a sub-angstrom resolution [2–4]. Most laboratories utilize HRTEM to resolve atomic planes and atoms in a crystal or distinguish areas with amorphous and crystalline phases, image defects, e.g., dislocations, stacking faults, nanosized particles and its lattice matching with the host lattice, etc. HRTEM image is an interference pattern of the electron wave function with itself after the interaction with the specimen. HRTEM demands the performance of the objective below the desired level of aberrations and mechanical, electrical, and environmental stability of the entire setup. In principle, the phase-contrast microscopy can provide information on the potential and fields at the atomic and subatomic length scale (other than what was mentioned just before), which is important to understand a material's response or properties in many applications.

1.2 Aberration corrected high-resolution transmission electron microscopy

1.2.1 Concept of resolution

Human eyes can distinguish objects only up to a fraction of mm, within a certain distance. With an optical microscope, 'invisible' objects for the unaided eyes are magnified and made visible. Optical microscopes use light as illumination, and it has a limit up to which it can resolve. Electron microscope offers a thousand times better resolution than the light microscope. In the conventional imaging theory, the probe illumination is considered a wave with a specific wavelength. Wave has a special characteristic of manifesting interference or diffraction phenomena. The image of a point object is blurred due to diffraction of wave known as Airy disk [Figure 1.01 (a)]. Rayleigh's criteria [Figure 1.01 (b)] determines the resolution limit in terms of distinguishing two such blurred disks, which depend on the wavelength of the illumination (λ) and diameter of the circular aperture (D), and is given by $\theta=1.22\lambda/D$, where θ is the angular resolution. For an electron microscope, λ can be varied with accelerating voltage, and resolution can be tuned. Thus, an electron microscope can overcome the resolution limit of the optical microscope imposed due to the wavelength of illumination. At 300 kV accelerating voltage, the λ of the electron is 1.97 pm. However, the achievable resolution in an aberration-corrected microscope is only 0.8 Å, which is much inferior to the limit set by the diffraction criteria. This is because of the aberration and the stability of the electromagnetic lens, limiting the electron microscope's resolution. The most dominant axial aberration of a TEM is a third-order spherical aberration and is parameterized by spherical aberration coefficient (Cs). Therefore, the resolution and optimum contrast in HRTEM in the presence of aberration is defined by Scherzer criteria [5]. The Scherzer relation is obtained by counterbalancing the Cs to the disk of least confusion from the Gaussian image plane corresponding to a phase shift of $-3\pi/4 < W(\Delta k) < 0$.

$$\text{Scherzer resolution} = \sqrt[4]{\frac{Cs}{6}\lambda^3} \tag{1.01}$$

But Zernike-like λ/4 or π/2 phase plate is not available in TEM. In HRTEM, a combination of Cs and defocus is used as a phase plate. Defocus has the effect of reducing the effect of Cs of the objective lens. The Scherzer criteria set the optimum balance between Cs and defocus, and the resolution is expressed regarding the phase contrast transfer function (PCTF) function [Figure 1.02].

Scherzer defocus is given by $= \sqrt{\frac{3Cs}{2}\lambda}$ (1.02)

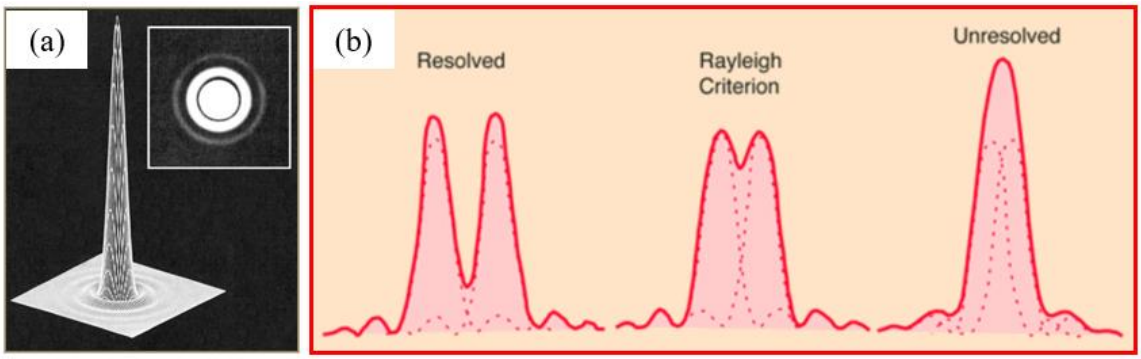


Figure. 1.01. (a) Blurring of the point object from Airy disk. Copyright (2009) by Springer [6]. (b) Rayleigh's criteria of resolution.

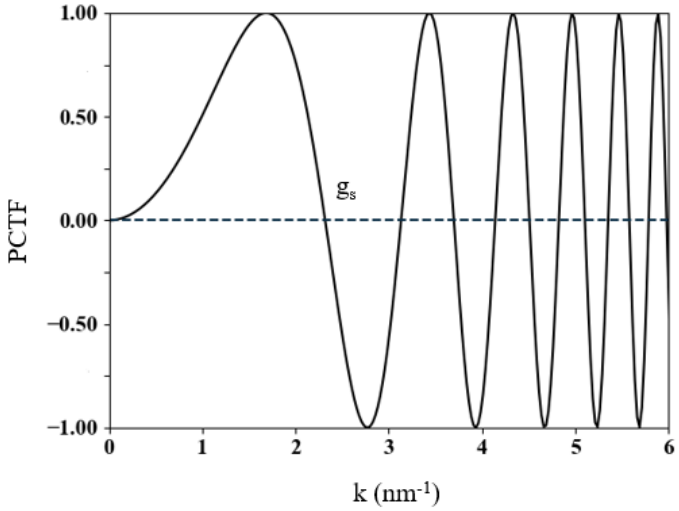


Figure. 1.02. PCTF showing Scherzer resolution of 0.43nm for a microscope with a fixed Cs = 1 mm with optimum Scherzer defocus of 82.4nm for a 200 keV electron microscope.

The first crossing of PCTF defines the microscope's resolution, and the last crossing (not shown in the above Figure) represents the information limit. Parasitic aberrations, e.g., vibrations, instability in the lens, power supply, come as an envelope. Their role is to lower the microscope's information limit and reduce the resolution limit if their magnitude is significant. Correcting the aberration is discussed in sec. 1.2.3.

1.2.2 Principle of image formation

The principle of image formation, according to various exiting theories and the alternative method proposed by us, is described in detail in chapter 3. Briefly, a crystal is made of a periodic arrangement of atoms. The positively charged nucleus is highly localized in space (1 to 15 fm) surrounded by the delocalized electron clouds (~ Å). When a fast electron plane wave travels through the crystalline specimen along the high symmetry zone axis (Z.A.), the wave undergoes a strong phase shift by the localized positively charged nucleus. The self-interference of this phase-shifted part of the wave with itself forms the exit wave function. The exit wave function is further modified by the lens aberration function and finally recorded by the detector, which modifies the signal through the modulation transfer function.

Incident electron wave after passing through the sample generates exit wave function, which is given by

$$\psi = A(x,y)e^{i\phi(x,y)} \tag{1.03}$$

Where $A(x,y)$ is the amplitude and $\Phi(x,y)$ is the phase representing the 2D modulation of the respective parameters. The object wave is modified by the lens aberration function or lens transfer function to form an image wave and is given by

$$b(u) = exp(i\chi) \tag{1.04}$$

Where χ consists of dominant aberration, i.e., Cs and defocus, which plays a role in reducing the effect of Cs to some extent. Self-interference of the wave at the image plane will give the image intensity pattern.

An interference pattern considering phase-contrast component valid within weak phase object approximation (WPOA) is given by [6],

$$I(x,y) = 1 + 2\phi(x,y) \otimes \sin(\chi) \tag{1.05}$$

The object wave function must be determined from the reconstructed image wave first by determining the lost phase and then by deconvolution of the aberrations [Figure 1.03].

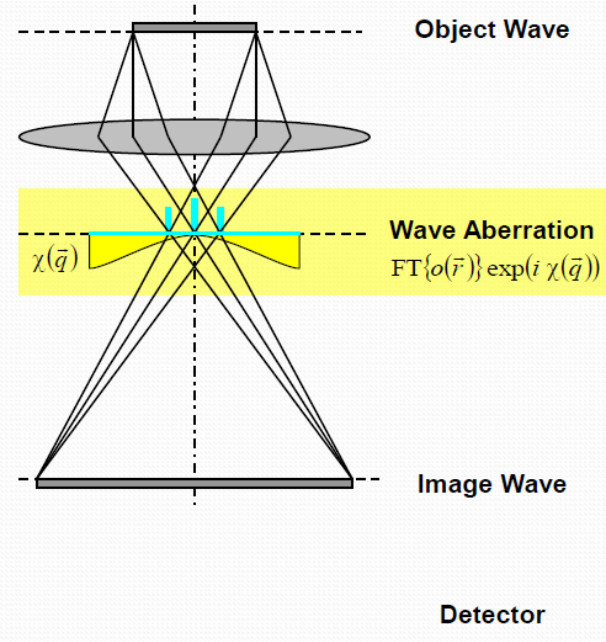


Figure 1.03. The schematic shows the plane wave electron propagation through the object to the exit wave function. The lens aberration function then modifies the exit wave function to the image wave function, and this is recorded by the detector. Copyright (2007) by Institute of Physics publishing group [7].

1.2.3 Aberration correction in TEM

The objective lens in the microscope deviates the rays from the ideal Gaussian plane, causing image aberration [Figure 1.03]. There are various types of aberrations in which third-order spherical aberration Cs is the limiting resolution for a round magnetic lens. The development of non-round electromagnetic lenses, e.g., hexapole correctors, opens up new possibilities of improving the resolution and opportunities for quantitative imaging at atomic and sub-atomic length scales [8]. The corrector contains a pair of hexapoles and two transfer round-lens doublets. The primary aberrations of the first hexapole are compensated by the second. These hexapoles additionally induce a secondary residual aberration, known as third-order spherical aberration. The sign of the spherical aberration coefficient is the opposite of the objective lens. Hexapoles can then be excited to the degree required to eliminate the spherical aberration of the entire system. This corrector system, which contains two hexapoles and a pair of round lenses, is proposed by Rose [9]. Hence it is known as Rose corrector. An example of such a microscope is given in Figure

1.04. It is a double aberration-corrected FEI TITAN 80-300 kV transmission electron microscope equipped with a Rose corrector and additionally gun monochromator with a point resolution better than 0.8 Å.

With aberration-corrected HRTEM and a combination of negative Cs and positive defocus, a single vacancy of boron (B) can be imaged in monolayer boron nitride (BN) [Figure 1.05 (a)] [10]. Even the edge structure of 2D materials can be imaged [Figure 1.05 (b)]. With the help of negative Cs imaging, atoms placed closely in the periodic table can be distinguished. BN imaging is one such example, where the N atom appears slightly brighter than the B atom [Figure 1.05 (a)] [10].

Figure 1.04. FEI TITAN 80-300 kV transmission electron microscope with a double aberration corrector.

Negative spherical aberration (Cs) avails imaging of small atoms like oxygen in the vicinity of heavy atoms in the crystal [11]. In this study by Jia et al. [11], successful imaging of all types of atomic columns in the dielectric $SrTiO_3$ (Strontium titanate) and the superconductor $YBa_2Cu_3O_7$ (Yttrium barium copper oxide) is carried out. Oxygen atoms were imaged using the negative Cs. Otherwise, they are not visible due to the low scattering power. With this, they were successful in detecting the nonstoichiometric or the degree of the oxygen vacancy ordering. Figure 1.06 shows the oxygen atoms clearly in the presence of yttrium, barium, and copper. Negative Cs also help improve the image contrast [12], which is essential to the quantitative study.

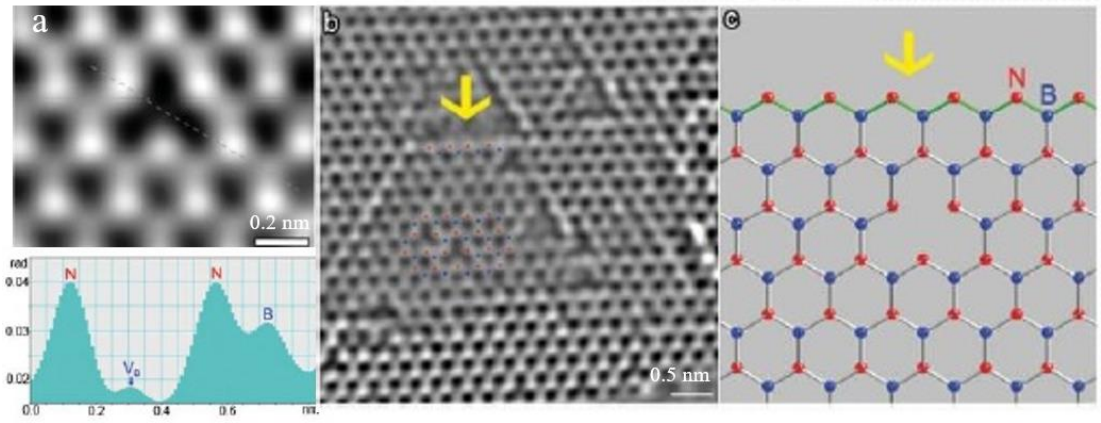


Figure 1.05. Phase image of (a) boron monovacancy with the corresponding line profile (b) edge structure terminated with the N atom. (c) A model for the edge structure. Copyright (2009) by American Physical Society [10].

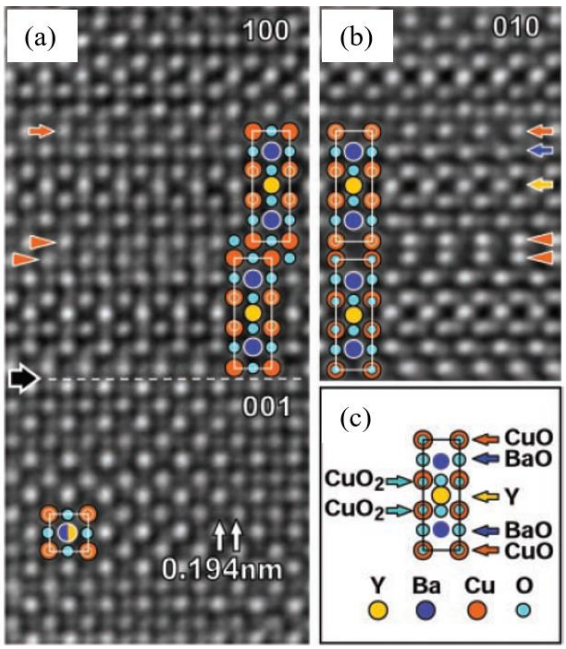


Figure 1.06. Experimental images of $YBa_2Cu_3O_7$ in different crystal orientations. (a) A 90° tilt boundary (black arrow) separating two crystal domains in the upper part, the crystal orientation is parallel to [100], whereas it is parallel to [001] in the lower part of the image. (b) The crystal is imaged along the [010] direction, i.e., tilted with respect to (a) by 90° around the vertical axis. The [010] Cu-O chains are now seen end-on. (c) The structure model indicates all atomic positions of the cations Ba, Y, Cu, and oxygen. Copyright (2003) by American Association for the Advancement of Science [11].

1.3 Electron holography: A brief introduction

1.3.1 In-line electron holography

Holography is a Greek word that means it contains the complete information (both amplitude and phase). Dennis Gabor initially proposed electron holography while improving the resolution limit of a conventional TEM [13]. Gabor recorded both amplitude and phase by a two-step process, first writing and then reading. In the writing step, the amplitude and phase information is written on a film. The recorded hologram is illuminated with a reference wave similar to the writing step in the reading step. Hence the single hologram contains the full information of the object. But the image quality was poor due to the overlap of the virtual and the real image, known as the twin image problem.

1.3.2 Off-axis electron holography at medium resolution

Off-axis electron holography was introduced only after the invention of electron biprism by Mollenstedt and Duker in 1955 [14]. Leith and Upatnieks later developed off-axis geometry in light optics to eliminate the twin image [15]. The principle of the technique is shown schematically in Figure 1.07. Electron wave is split into two halves; one goes through the sample, and another through vacuum, which serves as the reference wave. The interference between these waves produces an image called an electron hologram.

Medium resolution holography concentrates majorly on the large area phase structure. It is increasingly used in the analysis of 2D dopant profiling and dopant distribution. The study of doping profiles at a p-n junction in transistors, magnetic nanomaterials, and imaging of nanoparticles, ferroelectric polarization, etc. [16–21] are a few examples.

Ref. [19] is one such example of a dopant profile study in the semiconductor at the nanoscale. Here nMOS (n-type metal oxide semiconductor), pMOS (n-type metal oxide semiconductor), and CMOS (complementary metal oxide semiconductor) devices, grown on bulk silicon and silicon-on-insulator type device, electrostatic potential profiles in the presence of active dopants are investigated using the technique. Measurement of phase is carried out across the p-n junction.

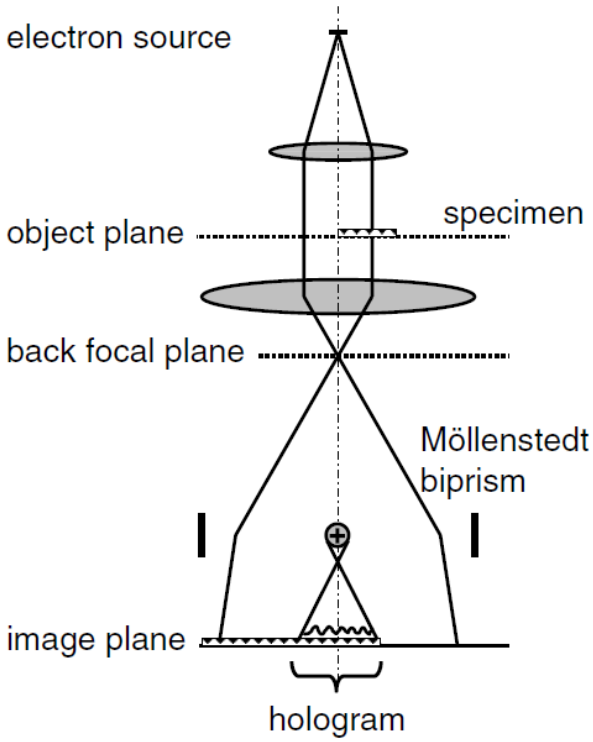


Figure 1.07. Schematic shows the geometry for off-axis electron holography. Copyright (2007) by Institute of Physics publishing group [7].

In another investigation by Tripp et al. [20], single-walled cobalt (Co) nanoparticle rings are examined at nanometer spatial resolution. Off-axis electron holography is used to visualize the magnetic flux. The phase recorded by holography reveals the handedness of the magnetic domains within the Co nanoparticle rings. The method is also useful in imaging the vortex-like spin structure of the skyrmion lattice [22,23]. A study by Li et al. [22] used off-axis electron holography to investigate the formation and characteristics of the skyrmion lattice defects and their relationship to a B20 (type of chiral magnet) crystallographic structure of FeGe (Iron germanide) thin film.

Medium resolution holography is applied to a vast area of research. This is only a glimpse of the powerful technique, which has more to offer.

1.3.3 Off-axis electron holography at atomic resolution

The off-axis electron holography has now reached an atomic resolution limit with aberration correction and biprism design modification [24–26]. At medium resolution, aberrations are almost negligible, and they come into the picture at atomic dimension. Similar to HRTEM, Cs and defocus are the major aberrations that are considered. With the available correctors, the resolution has improved up to the information limit beyond 0.1 nm. An empty hologram is recorded without having the sample in the beam direction [Figure 1.08 (a)]. Figure 1.08 (b) is an atomic resolution hologram of zinc oxide (ZnO) along $[01\bar{1}0]$ zone axis (Z.A.). Fringes are visible on the atomic columns of zinc (Zn) and between the columns.

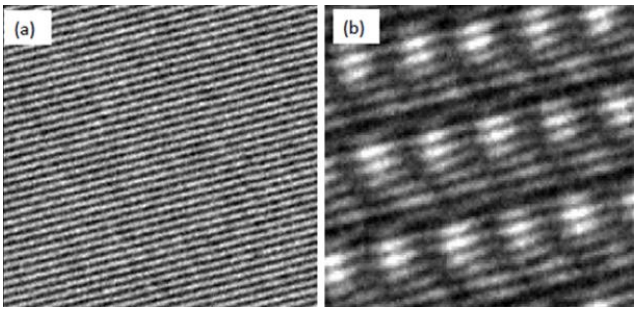


Figure 1.08. Examples of (a) an empty hologram and (b) atomic resolution hologram fringes on the Zn atom.

The intensity of the hologram is given by [7],

$$I(\vec{r}) = 1 + I_{el}(\vec{r}) + I_{inel}(\vec{r}) + 2C.A(\vec{r})\cos(2\pi q_c x + \Phi(\vec{r})) \quad (1.06)$$

Where $q_c = k\beta$ is the carrier spatial frequency, C is the contrast. $I_{el}(\vec{r})$ and $I_{inel}(\vec{r})$ represent all elastic and inelastic interactions. $A(\vec{r})$ and $\Phi(\vec{r})$ are the amplitude and phase, respectively. k is the reciprocal wave vector, and β is the angle between the reference and the modified wave in the *x*-direction.

The Fourier transformation (FT) of this hologram is given by Eq. 1.07 [7] and produces one central band and two sidebands [Figure 1.09]. The sideband contains complete information on the phase of the object.

$$spec(q) = FT[1 + I_{el}(\vec{r}) + I_{inel}(\vec{r})] \qquad \text{central band}$$

$$+C.FT\left[A(\vec{r})\exp\left(i\varphi(\vec{r})\right)\right]\otimes\delta(\vec{q}-\vec{q}c\,) \quad \text{+1 sideband}$$
$$+C.FT\left[A(\vec{r})\exp\left(-i\varphi(\vec{r})\right)\right]\otimes\delta(\vec{q}+\vec{q}c\,) \quad \text{-1 sideband} \qquad (1.07)$$

The Central band contains both elastic and inelastically scattered electrons and both linear and nonlinear terms. It does not include phase $\varphi(\vec{r})$. The ±1 sidebands contain Fourier spectrum of the complete image wave damped by the factor C and convoluted around $\vec{q}=\pm\vec{q}c$. The advantage of sidebands is that they contain only elastically scattered electrons and contain both amplitude and phase, which are linearly related to the object properties. Two sidebands are equivalent and provide all the information about the image wave. This is the advantage of off-axis electron holography over HRTEM, where the twin images or the complex conjugate can be easily separated by masking one of the sidebands.

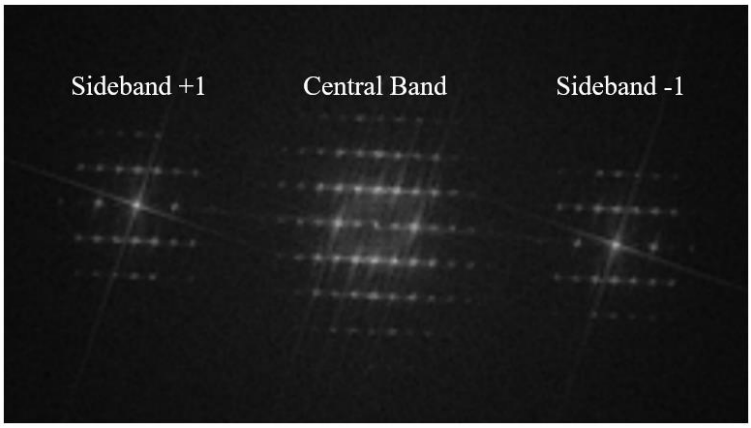


Figure 1.09. Fourier transform of the hologram results in two sidebands: complex conjugate of each other and one central band.

Various studies have been reported on the atomic resolution off-axis electron holography, for example, a quantitative study of 2D materials, potential mapping of graphene sheets, etc. [26,28,29]. Atomic resolution off-axis electron holography offers an exciting opportunity to obtain information at the atomic columns and between the atomic columns. Hence it is one of the most used techniques for quantitative analysis. A few reports try to examine the material properties by in-line and off-axis electron holography. But the results obtained by two different methods do not match even though they are the most widely used techniques [17,27,30–32]. Approaches and formalisms considered in both the methods are different, and thus the deviation is unavoidable. A detailed study about the differences is furnished in chapter 2.

1.4 Quantitative imaging

With the availability of aberration-corrected microscopes and better detectors, active research has begun to extract more information about the material rather than focusing only on the atomic arrangement. For example, counting the number of atoms in the atomic column, detecting point defects and dopants, three-dimensional arrangement of the crystal and its extreme, the electron distributions, and the potential around it. This is the origin of quantitative imaging, which renders more than crystal structure. The goal of the thesis is to contribute to these efforts by developing methods to understand phase-contrast information. Attempts to count, identify the types and defects of atoms in a crystalline solid are reflected in Ref. [33–35] involving developing methods to quantify such information. This section spreads light on such techniques, approaches, and the nature of the information available in the literature.

There are various approaches to quantitative imaging. Here are a few examples determining the three-dimensional arrangements of atoms in bi-layer graphene and MgO crystal [34,36]. A new method based on phase velocity derived from the Big-Bang and expanding universe theory in cosmology is applied for the bi-layer graphene sample. The process is based on first decomposing the spherical wave emanating from the atoms in terms of Fourier components, and the phase of each Fourier component varies linearly with increasing distance from the source and is given by $\pi\lambda g^2 f$ where λ is the wavelength, g is the spatial frequency, and f is the focal distance between the atom and the plane at which the exit wave is reconstructed. The position of each carbon atom has been calculated in this method [34]. In the case of MgO crystal, complete information of the 3D shape of the crystal, surface adsorbed impurity atom is determined from a single image [36]. The method is based on careful interpretation of the 3D nature of electron diffraction. Three different steps are involved in the procedure; choice of sensitive imaging mode for experimental data acquisition, structure refinement procedure, and the statistical confidence check that yields the final 3D structure model's uniqueness. The refinement process makes use of the absolute image intensity levels instead of commonly used relative values. Figure 1.10 shows the reconstructed 3D structure from the 2D projected intensity information.

Electron charge density or the potential of the atom gives information about electronic properties like chemical bonds or the degree of ionization. However, electrons involved in bonding are only valence electrons. As they are small in number compared to total electrons, it is not very

easy to measure. Meyer Jannik C. et al. demonstrated an experimental analysis of charge redistribution due to chemical bonding by HRTEM [33] and matched it with first-principles electronic structure calculation. Charge transfer on the single-atom level for nitrogen-substitution point defects in graphene was studied, and iconicity of the single-layer hexagonal boron nitride was confirmed.

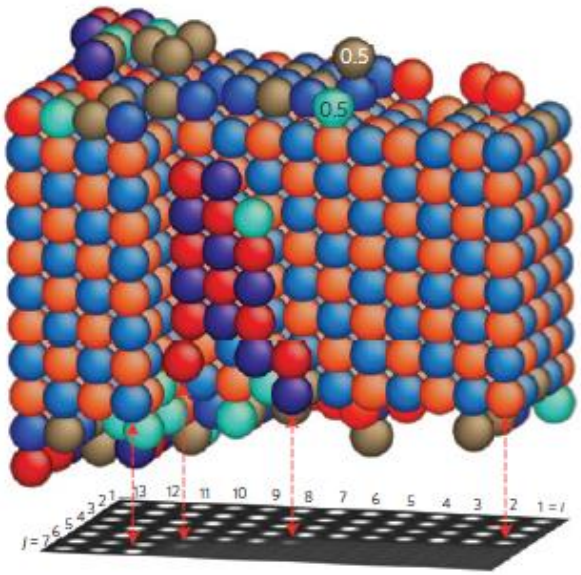


Figure 1.10. Atomically resolved view of the best fitting 3D atomic arrangement for MgO. Red spheres indicate fully occupied Mg sites, and blue spheres fully occupied O sites. Increased color saturation is used to highlight surface atoms. In the surface layers, brown spheres indicate formally half-occupied Mg sites, while cyan spheres indicate formally half-occupied O sites. Copyright (2014) by Nature publishing group [34].

TEM records spatially resolved scattered electron density, which helps to study the material's structure and properties. A new way of exploring the features of the material is based on the probability current of the scattered electrons [37]. 2D lateral quantum mechanical probability current of the scattered electrons are reconstructed in the image plane. The currents are reconstructed from three atomic-resolution TEM images recorded under a slight difference of perpendicular line foci. The probability current is the local measure of the scattering direction of the electrons, therefore providing information on the target properties. Depending on the specimen interaction and energy loss regime considered, one can identify several simple approximate links between the current density and physical properties of the target.

The theoretical approach by computer simulation is made to develop a sensitive method to extract the information of valence electrons [38]. Scattered electron wave is decomposed in a set of primary functions, which are the Eigenfunctions of the Hamiltonian of the projected electrostatic object potential. Each primary function behaves as a communication channel that transfers the information of the object with its transmission characteristic. Information on valence electron distribution is then obtained from a series of exit waves by adequately combining the components of different channels.

These are a few examples of quantitative imaging, and there are alternative methods and techniques available. The area is still open to new and simple approaches. The thesis aims to contribute to this.

Image simulation and reconstructions are at the heart of quantitative imaging. Image simulation is necessary to understand and quantify the number and type of atoms in the acquired images. There are reports which study the match between simulated and experimentally obtained HRTEM images. There is a discrepancy of 3 to 5 times due to what is known as 'Stobb's factor' [39–41]. Researchers are trying to figure out a particular reason rather than the combination of possibilities and obtain the exact match between them. One such example is reflected in Ref. [42]. Simulated and experimental contrasts are compared for different objective apertures. Yet, a difference of 1.2 is obtained in the image contrast for a larger objective aperture size. Various simulation methods and approaches are provided in detail in chapter 3, and also a new method of simulation which matches exactly with the experimental data is proposed.

Similarly, object exit wave (OEW) reconstruction is as important as the simulation to extract the information of the object. There are numerous ways to reconstruct the exit wave from the images, and the particulars are provided in chapter 4. Two distinct methods are put forward based on real and reciprocal space. A single image is enough to reconstruct the OEW in these methods, avoiding the tedious focal series reconstruction process.

1.5 Overview of the thesis

In chapter 2, quantitative counting of Zn and O atoms in zinc oxide (ZnO) epitaxial thin film is carried out by different routes; reconstruction of phase from the side and a central band of atomic resolution off-axis and in-line electron holography. Interesting results are obtained for different methods.

Chapter 3 introduces an alternative approach to the image simulation in HRTEM after a comparative analysis of the existing image simulation methods. The alternative method is based on considering the atom as an electrostatic interferometer. Simulation results are compared with the experimental images of 2D materials of MoS_2 and BN recorded under the optimum combination of third-order spherical aberration and defocus.

In chapter 4, an alternative reconstruction method is proposed for retrieving the OEW from the recorded image intensity pattern in HRTEM. The method is based on applying a modified intensity equation representing the HRTEM image. Additionally, it is shown that the Fourier series expansion of diffraction pattern is effective in retrieving the isolated and periodic image functions of a specific form.

1.6 Bibliography

[1] F. Zernike, Science. **121**, 345 (1955).

[2] R. Sahu, U. Bhat, N. M. Batra, H. Sharona, B. Vishal, S. Sarkar, S. Assa Aravindh, S. C. Peter, I. S. Roqan, P. M. F. J. Costa, and R. Datta, J. Appl. Phys. **121**, 105101 (2017).

[3] R. Sahu, K. Dileep, D. S. Negi, K. K. Nagaraja, S. Shetty, and R. Datta, J. Cryst. Growth **410**, 69 (2015).

[4] K. Dileep, R. Sahu, K. K. Nagaraja, and R. Datta, J. Cryst. Growth **402**, 124 (2014).

[5] O. Scherzer, J. Appl. Phys. **20**, 20 (1949).

[6] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science*, 2nd ed. (Springer, 2009).

[7] H. Lichte and M. Lehmann, Reports Prog. Phys. **71**, 016102 (2007).

[8] M. Haider, H. Rose, S. Uhlemann, E. Schwan, B. Kabius, and K. Urban, Ultramicroscopy **75**, 53 (1998).

[9] H. Rose, Nuclear Instruments and Methods **187**, 187 (1981).

[10] C. Jin, F. Lin, K. Suenaga, and S. Iijima, Phys. Rev. Lett. **102**, 195505 (2009).

[11] C. L. Jia, M. Lentzen, and K. Urban, Science. **299**, 870 (2003).

[12] C. L. Jia, L. Houben, A. Thust, and J. Barthel, Ultramicroscopy **110**, 500 (2010).

[13] D. Gabor, Nature **161**, 777 (1948).

[14] G. Möllenstedt and H. Düker, Zeitschrift Für Phys. **145**, 377 (1956).

[15] E. N. Leith and J. Upatnieks, J. Opt. Soc. Am. **52**, 1123 (1962).

[16] A. C. Twitchett, R. E. Dunin-Borkowski, R. F. Broom, and P. A. Midgley, J. Phys. Condens. Matter **16**, S181 (2003).

[17] A. C. Twitchett, R. E. Dunin-Borkowski, and P. A. Midgley, Philos. Mag. **86**, 5805 (2006).

[18] J. M. Thomas, E. T. Simpson, T. Kasama, and R. E. Dunin-Borkowski, Acc. Chem. Res. **41**, 665 (2008).

[19] D. Cooper, J. Phys. D. Appl. Phys. **49**, 474001 (2016).

[20] S. L. Tripp, R. E. Dunin-Borkowski, and A. Wei, Angew. Chemie **115**, 5749 (2003).

[21] D. Szwarcman, A. Lubk, M. Linck, K. Vogel, Y. Lereah, H. Lichte, and G. Markovich, Phys. Rev. B - Condens. Matter Mater. Phys. **85**, 134112 (2012).

[22] Z. A. Li, F. Zheng, A. H. Tavabi, J. Caron, C. Jin, H. Du, A. Kovács, M. Tian, M. Farle, and R. E. Dunin-Borkowski, Nano Lett. **17**, 1395 (2017).

[23] A. Kovács, Z.-A. Li, K. Shibata, and R. E. Dunin-Borkowski, Resolut. Discov. **1**, 2 (2016).

[24] H. Lichte, Ultramicroscopy **20**, 293 (1986).

[25] M. Linck, B. Freitag, S. Kujawa, M. Lehmann, and T. Niermann, Ultramicroscopy **116**, 13 (2012).

[26] F. Winkler, J. Barthel, A. H. Tavabi, S. Borghardt, B. E. Kardynal, and R. E. Dunin-Borkowski, Phys. Rev. Lett. **120**, 156101 (2018).

[27] U. Bhat and R. Datta, J. Appl. Phys. **125**, 154902 (2019).

[28] V. Boureau, B. Sklenard, R. McLeod, D. Ovchinnikov, D. Dumcenco, A. Kis, and D. Cooper, ACS Nano **14**, 524 (2020).

[29] D. Cooper, C. T. Pan, and S. Haigh, J. Appl. Phys. **116**, 039901 (2014).

[30] T. Niermann and M. Lehmann, J. Phys. D. Appl. Phys. **49**, 194002 (2016).

[31] C. T. Koch and A. Lubk, Ultramicroscopy **110**, 460 (2010).

[32] T. Latychevskaia, P. Formanek, C. T. Koch, and A. Lubk, Ultramicroscopy **110**, 472 (2010).

[33] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Großß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser, Nat. Mater. **10**, 209 (2011).

[34] C. L. Jia, S. B. Mi, J. Barthel, D. W. Wang, R. E. Dunin-Borkowski, K. W. Urban, and A. Thust, Nat. Mater. **13**, 1044 (2014).

[35] S. Borghardt, F. Winkler, Z. Zanolli, M. J. Verstraete, J. Barthel, A. H. Tavabi, R. E. Dunin-Borkowski, and B. E. Kardynal, Phys. Rev. Lett. **118**, 1 (2017).

[36] D. Van Dyck, J. R. Jinschek, and F. R. Chen, Nature **486**, 243 (2012).

[37] A. Lubk, A. Béché, and J. Verbeeck, Phys. Rev. Lett. **115**, 1 (2015).

[38] Q. Xu, H. W. Zandbergen, and D. Van Dyck, Ultramicroscopy **111**, 912 (2011).

[39] M. J. Hÿtch and W. M. Stobbs, Ultramicroscopy **53**, 191 (1994).

[40] A. Howie, Ultramicroscopy **98**, 73 (2004).

[41] A. Thust, Phys. Rev. Lett. **102**, 5 (2009).

[42] F. F. Krause, K. Müller, D. Zillmann, J. Jansen, M. Schowalter, and A. Rosenauer, Ultramicroscopy **134**, 94 (2013).

Chapter 2

Quantitative counting of Zn and O atoms by atomic resolution off-axis and in-line electron holography

In this chapter, quantitative counting of Zn and O atoms in zinc oxide (ZnO) epitaxial thin film is carried out by different routes; reconstruction of phase from the side and a central band of atomic resolution off-axis and in-line electron holography. Results show that the reconstructed phase from both the side and central band and the corresponding number of Zn (Z = 30) and O (Z = 8) atoms are in close agreement with the systematic increase in the number of atoms for the sample area less than the extinction distance. However, complete disagreement is observed for the sample area more than the extinction distance. On the other hand, the reconstructed phase obtained via in-line holography shows no systematic change with thickness for the same sample. Phase detection limits and the atomic model used to count the atoms are discussed.

***U. Bhat** and R. Datta, J. Appl. Phys. 125, 154902 (2019).*

2.1 Introduction

Phase shift is the fundamental information obtained by high-resolution transmission electron interferometry [1]. Phase shift encodes the information on the potential distributions of atomic ensembles, which may be used to deduce the atomic arrangement and properties of the materials [2–4]. Two established approaches, i.e., off-axis and in-line electron holography, can be used to retrieve the phase information experimentally at atomic and sub-atomic length scales. Out of focus HRTEM (high-resolution transmission electron microscopy) images can be considered as an in-line hologram. Holography was first proposed by Gabor and subsequent development of off-axis geometry by Leith and Upatnieks [5,6]. Off-axis geometry eliminates the twin image problem associated with Gabor's original idea of in-line holography [6]. Gabor's proposal was based on using a reference wave to interfere with the object wave, e.g., an electron micrograph to overcome the resolution limit imposed by the geometrical aberrations of the electron lens. Such a hologram contains all the information about the object and the imaging system. The practical off-axis electron holography technique developed by Möllenstedt and Düker implemented an electrostatic biprism for electron interference [7].

Fourier transformation of off-axis electron hologram results in three bands; one central (CB) and two sidebands (SBs), where the sidebands are complex conjugate of each other. Phase information can be directly extracted from one of the two SBs. The CB is equivalent to in-line holography, which contains mixed amplitude and phase signals. Off-axis electron holography is a routine technique for medium resolution imaging of electric and magnetic fields [8–10]. Recently, atomic resolution off-axis electron holography is pursued with the development of a special holography microscope equipped with double electron biprism [11–14]. Double biprism setup eliminates Fresnel fringes and Vignetting effect essential for recording good quality hologram throughout the field of view, usually small, at an atomic resolution [15,16]. In the atomic resolution off-axis electron holography technique, electron interference fringes encode phase information at a sub-atomic length scale.

On the other hand, retrieval of phase shift from in-line holography requires series of images to be recorded at different focus values. Various reconstruction schemes have been developed to retrieve the object exit wave (OEW) function from the experimental through focal image series method [1,17–21]. The development of both the experimental approaches to extract phase

information dates back to the BRITE EURAM program [22]. Few groups have reported comparisons of phase information by two different methods at medium and atomic resolutions. However, quantitative phase information obtained so far through off-axis and in-line holography do not correspond to each other for the same sample area and depend on the frequency range considered for the analysis [14,23–25]. Quantitative imaging is a recent area of active research in the atomic resolution electron microscopy community. Understanding the accuracy of the experimental phase determination and its correlation with the property of materials is crucial for its success and contribution to the materials and microscopy sciences as a whole [12,26–33]. Both aberrations corrected HRTEM and atomic resolution off-axis electron holography provides a unique opportunity to study phase information at the atomic and sub-atomic length scale.

This chapter compares the quantitative atomic-scale phase information by different phase retrieval methods, from SB and CB of off-axis electron holography and in-line electron holography. It is found that the peak phase values and the corresponding number of atoms for both heavy Zn (Z =30) and light O (Z = 8) are in close agreement between the SB and CB of off-axis electron holography for thinner specimen areas with a systematic change with sample thickness. However, for thicker sample areas, the agreement no longer holds. On the other hand, the number of atoms obtained via HRTEM method does not systematically change with the sample thickness variation.

2.2 Experimental details and data analysis

2.2.1 Thin-film crystal growth method

The ZnO epitaxial thin films were grown on 'c' plane ZnO substrates under two different oxygen partial pressure (p_{O_2}) using pulsed laser deposition (PLD) following a procedure described earlier [34,35]. Electron carrier concentrations can be controlled between 10^{19} to 10^{16} cm^{-3} with p_{O_2} 10^{-5} and 10^{-2} Torr, respectively. Though the original aim was to compare the difference in point defect distribution responsible for the change in carrier concentrations in these two samples, we restrict ourselves to the counting of atoms by different phase-contrast imaging routes in this chapter.

2.2.2 TEM sample preparation

Cross-sectional TEM specimens were prepared by mechanical polishing using a combination of coarse grit SiC (< 100 µm) and diamond polishing paper (15 to 1 µm) to a thickness below 20 µm. A tripod polisher was used for the final mechanical thinning procedure. Then the sample was milled to perforation in a Gatan PIPS Ar ion polishing system. An Ar ion energy of 4 and 2 kV was used during the initial and final milling stages, respectively, with a top and bottom gun angle setting of ± 7°. Lower kV at the end helps to remove thin amorphous skin layers. This is essential for the quantitative evaluation of atomic columns.

2.2.3 Off-axis electron holography

A. Principle of off-axis electron holography

The schematic of principles behind in-line and off-axis electron holography image acquisition and phase retrieval are shown in Figure 2.01 and 2.02, respectively. In atomic resolution off-axis electron holography, the electron interference fringes encode phase information at the sub-Å length scale where object wave is an atomic resolution electron micrograph. Double biprism set up at a particular location in the microscope column is essential to avoid Fresnel artifact and Vignetting effect, particularly at the atomic resolution where the field of view is severely restricted [Figure 2.01 (c)] [36]. The first step of off-axis electron holography involves recording electron interference fringes or electron holograms. This hologram is the result of interference between two partial waves; image (ψ_{img}) and vacuum reference waves $\psi_{ref}(x,y)$. The intensity distribution in a hologram is given by [1]:

$$I_{hol}(x,y) = [\psi_{img}(x,y) + \psi_{ref}(x,y)]\,[\psi_{img}(x,y) + \psi_{ref}(x,y)]^{*} \quad (2.01)$$

In the case of off-axis electron holography, a Möllenstedt biprism facilitates interference between image $[\psi_{img}(x,y)\exp(i\pi k\beta x)]$ and reference waves $[\psi_{ref}(x,y)\exp(-i\pi k\beta x)]$ at an angle β [Figure 2.01(b)]. This results in the intensity of the off-axis electron hologram given by [2]

$$I_{hol}(x,y) = 1 + A^2(x,y) + 2C\,A(x,y)\cos(2\pi q_c x + \phi(x,y)) \quad (2.02)$$

Where $q_c = k\beta$ is the carrier spatial frequency of the interference fringes and $C = \frac{I_{max}-I_{min}}{I_{max}+I_{min}}$ is the fringe contrast or fringe visibility (also can be written as V). The instabilities associated with recording the hologram (V_{inst}), the modulation transfer function of the CCD camera (V_{mtf}), and inelastic damping factor (V_{inel}) can be added with the contrast C as $V = C\,.V_{inst}\,.V_{mtf}.V_{inel}$ In the hologram, both amplitude (A) and phase (ϕ) information are recorded in the intensity distribution [2].

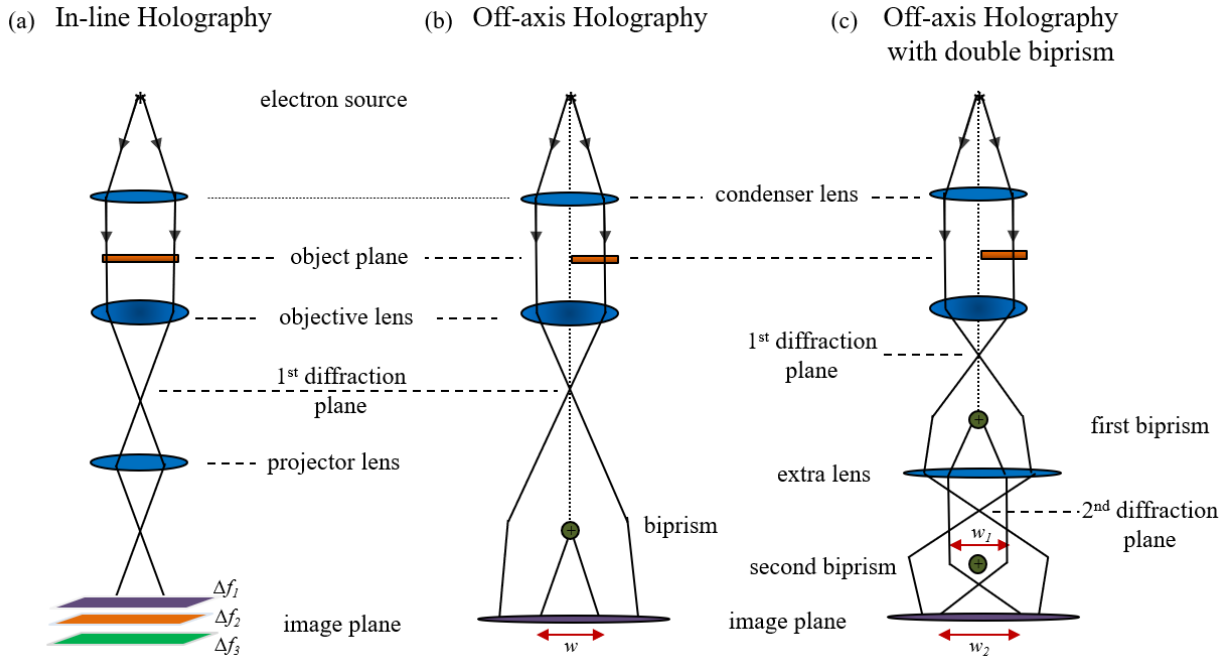


Figure 2.01. The schematics showing the principle of image formation for (a) In-line holography, (b) off-axis electron holography with single biprism, and (c) off-axis electron holography with double biprism set up. Copyright (2019) by American Institute of Physics [37].

B. Principle of OEW reconstruction

The first step involves Fourier transformation of Eq. 2.02, which gives [2]:

$$\begin{aligned} FT\{I_{hol}\} &= \delta(q) + FT\{A^2\} && \text{central band (CB)} \\ &+C.FT\left[\psi_{img}\right] \otimes \delta(\vec{q} - \vec{q}c\,) && \text{+1 sideband (SB1)} \\ &+C.FT\left[\psi^*_{img}\right] \otimes \delta(\vec{q} + \vec{q}c\,) && \text{-1 sideband (SB2)} \end{aligned} \qquad (2.03)$$

The central band (CB) contains only the image amplitude and no phase information. The two sidebands (SBs) include amplitude and phase information and are complex conjugates of each other.

In the second step, one of the SBs, preferably next to the vacuum, is masked out, and inverse Fourier transformed to real space as a complex array of data from which amplitude and phase information are extracted. In the holography reconstruction process, the bending of interference fringes encodes the phase modulation of the electron wave. However, there may be a contribution to the bending of fringes other than object potential: charging of biprism, distortion in the TEM projector lens, etc. These distortions induced phase modulation can be removed at the time of reconstruction using an empty hologram recorded under the same experimental conditions. The residual coherent aberrations are then corrected digitally for any quantitative work.

In present holography reconstruction, series of holograms have been recorded at different defocus settings, i.e., -10 to +10 nm in focus step of 1 nm with Cs =0. There are two reasons for this: it improves the signal to noise ratio in the reconstructed OEW compared to a single hologram image [14]. And the second is that it allows reconstruction of OEW from the CB applying the similar numerical reconstruction method as in the case of in-line holography/HRTEM. Therefore, it gives a unique opportunity to compare the outcome of two different experimental phase retrieval techniques from the same measurement. The numerical reconstruction method implemented for CB reconstruction is a combination of the paraboloid method (PAM) and maximum-likelihood method (MAL).

C. Instrumentation and data analysis

The present data were acquired using an aberration-corrected FEI TITAN 80-300 Berlin holography microscope operated at 300 kV equipped with a double biprism setup. Through focal image series was acquired at focus step (Δf) of =1 nm in the range of -10 to +10 nm. Third order spherical aberration coefficient (C_S) was set to zero. It is already mentioned that the aberration correction improves the phase detection limit by a factor of 4 [11]. Through focal holography method from the series of images provides extraction of phase through CB using standard algorithm used for HRTEM.

In the present case, a combination of PAM and MAL was used for the CB reconstruction. MAL corrects exit wave function iteratively, based on a least-square formalism. Series of images improves the signal to noise ratio significantly for the SB reconstruction [14]. Earlier comparison of phase values for a single image SB reconstruction at the medium resolution reported a low signal to noise ratio [25]. The details of the principle behind the method can be found in Ref. [38] and is shown schematically in Figure 2.02 (a). Figure 2.02 (c) is an example of an atomic resolution hologram, and Figure 2.02 (b) is the corresponding Fast Fourier transform (FFT) image. Reconstruction is carried out for the set of twenty images for both SB and CB. Zn and O atoms are visible in the magnified image of the hologram along $[11\bar{2}0]$ Z.A. [Figure 2.03]. The cut-off frequency for CB and SB reconstruction is 14 and 12 nm^{-1}, respectively. The cut-off frequencies are chosen in a way to avoid overlapping of neighboring bands.

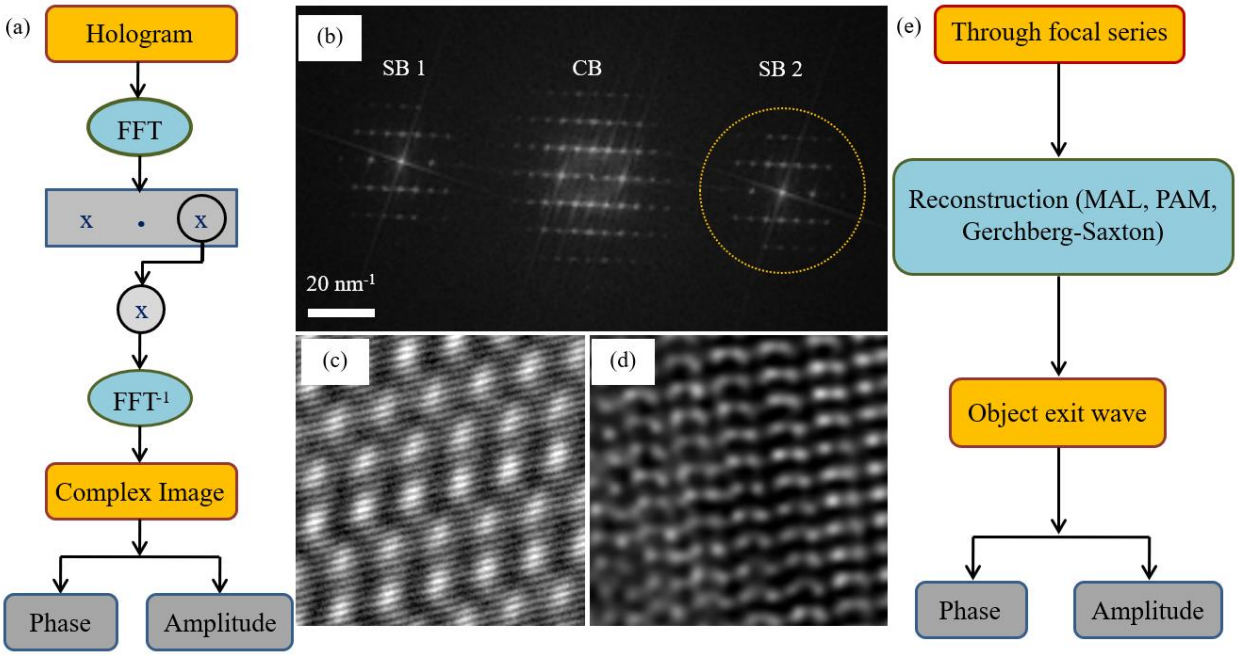


Figure 2.02. (a) and (e) Steps involved in reconstruction methods to extract phase and amplitude from the hologram and HRTEM image series, respectively. (b) Fourier transform of the hologram showing one CB and two SBs. (c) and (d) are the example atomic resolution hologram and HRTEM image of ZnO epitaxial thin film along $[11\bar{2}0]$ Z.A. Copyright (2019) by American Institute of Physics [37].

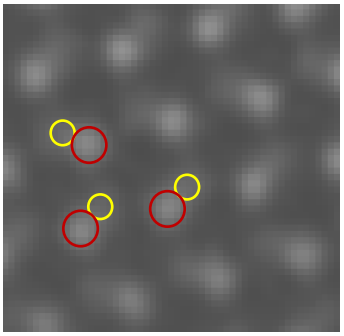


Figure 2.03. Off-axis electron holography image shows the arrangement of Zn and O atoms along $[11\bar{2}0]$ Z.A., where the red circles correspond to Zn atom and yellow circles correspond to O atoms.

2.2.4 In-line electron holography

A. Principle of HRTEM phase retrieval

Unlike holography, phase information cannot be obtained from a single HRTEM image. For phase retrieval, the widely used method is the 'focus variation method'. A series of images are recorded at different focus values with specific focus steps for a given Cs setting. The modification of the unknown phase, with the focus variation, can be linear or nonlinear depending on the sample thickness. There is an analytical solution to the linear reconstruction problem for a very thin sample or within weak phase object approximation (WPOA). However, most of the practical samples modify the phase of the probe wave function nonlinearly. There exists no direct analytical solution to the nonlinear reconstruction problem, and the available schemes are mostly iterative numerical methods. Gerchberg-Saxton algorithm, MAL are examples of such nonlinear reconstruction methods. The intensity recorded in the HRTEM image can be expressed as [1]:

$$I_L(g \neq 0) = \begin{pmatrix} X_0(g) \\ Y_0(g) \end{pmatrix} \left[\begin{pmatrix} \cos(\chi(g)) & -\sin(\chi(g)) \\ \sin(\chi(g)) & \cos(\chi(g)) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] = \begin{pmatrix} X_0(g) \\ Y_0(g) \end{pmatrix} \begin{pmatrix} \cos(\chi(g)) \\ \sin(\chi(g)) \end{pmatrix} \quad (2.04)$$

The expression in square brackets describes a rotation of the projection (1,0) in the image plane to give a new projection axis $(\cos(\chi), \sin(\chi))$, where the object wave function is given by (X_0, Y_0) is left unchanged.

B. Instrumentation and data analysis

HRTEM data was acquired in a double aberration-corrected FEI TITAN 80-300 kV transmission electron microscope at ICMS, JNCASR, Bangalore. An optimized phase contrast transfer function (PCTF) with C_s = -35 µm, f = 8 nm, and a point to point resolution better than 0.8 Å at 300 kV was set as a reference for the experimentation. Image series were recorded with $C_s = -35\ \mu m$, focus range -10 to 10 nm with focus step Δf = 1 nm and exposure time of 1s. However, only ten images around zero focus (±5) are used for the reconstruction. We did not observe any difference in signal to noise ratio in the reconstructed phase between ten and twenty images. The image series was reconstructed using the Gerchberg-Saxton scheme as implemented in the MacTempas package. The reconstruction parameters are; $C_s = -35\ \mu m$, acceleration voltage 300 kV, area of reconstruction 1024×1024 (pixel), and objective aperture size g_{max} =2 Å^{-1}. A strong central beam condition is considered. The phase image obtained was further corrected for the residual aberration using the digital aberration correction scheme available within the package. Example HRTEM image and the reconstruction steps are shown in Figure 2.02 (d) and (e), respectively. Example images before and after aberration correction are shown in Figure 2.04.

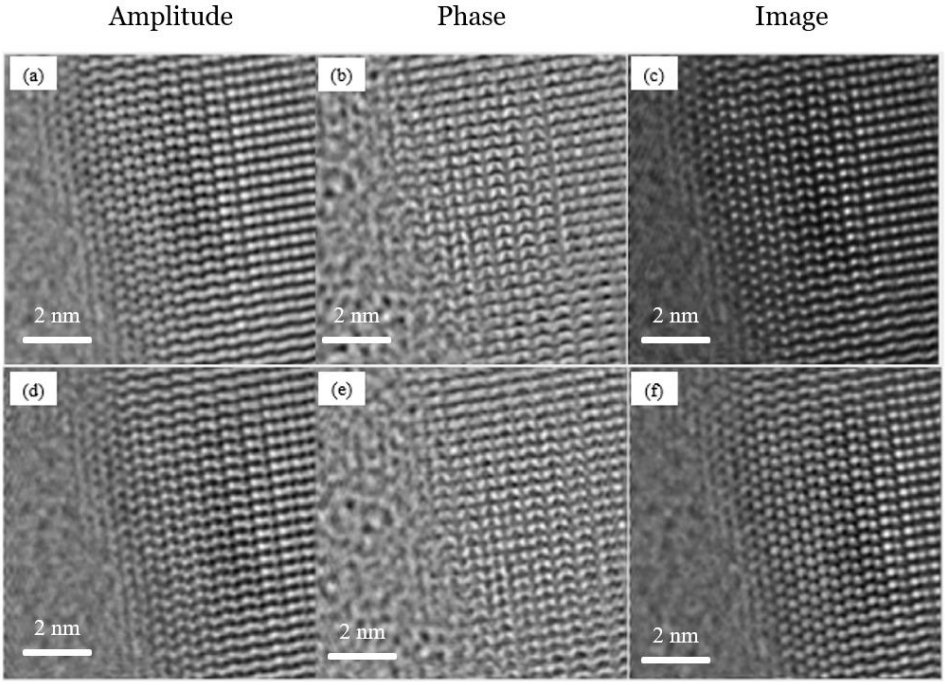


Figure 2.04. Example of HRTEM reconstructed amplitude, phase, and image (a)-(c) before and (d)-(f) after digital aberration correction. Cs = -35µm, f = 2nm are used for the aberration correction in this example.

2.3 Results and discussion

2.3.1 Phase detection limit

Resolution is the most crucial parameter in high-resolution transmission electron microscopy. In the presence of aberration, the point resolution g_S is defined by the first zero crossings of the phase contrast transfer function (PCTF) on the frequency axis under optimum C_s and defocus Δf [Figure 2.05]. The information limit g_i of a microscope is the maximum information which can be transferred and is defined by the last point of the PCTF function just above the noise level and usually damped by various incoherent aberrations. The information encoded between the point resolution and information limit is not directly interpretable. For example, in an aberration-corrected microscope, one can obtain a resolution, g_S better than 0.8Å, which is enough to resolve any atomic columns in the crystalline material along high symmetry orientation. This reveals the structure of the material directly in terms of the periodic arrangement of atoms.

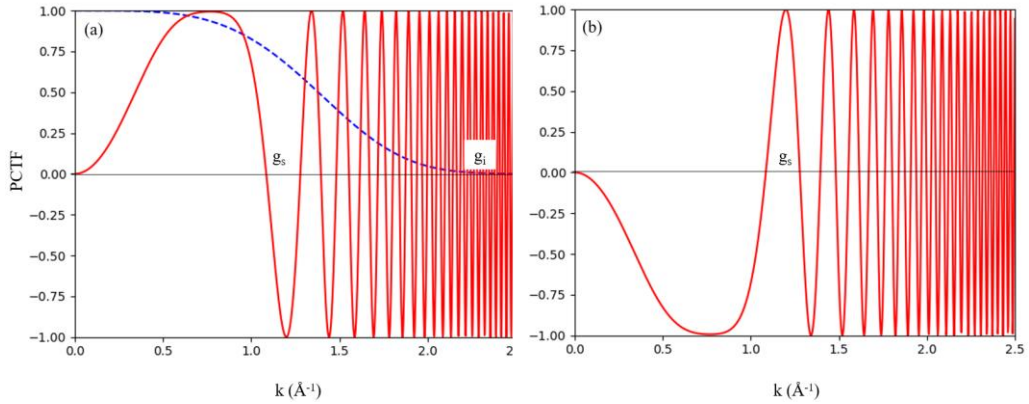


Figure 2.05. (a) PCTF function at 300 kV under optimum lens parameters, Cs = -35µm f = 8nm with the envelope function (blue dotted line) corresponding to spread in defocus 1nm. The point resolution and information limits are marked as g_s and $g_{i,}$ respectively. The positive PCTF gives negative phase contrast or white atom contrast. (b) PCTF function corresponding to positive phase contrast or dark atom contrast with Cs = 35µm, f = -8nm.

Like resolution, a minimum detectable change in amplitude and the phase signal of an electron wave after interacting with the specimen potential is equally important to evaluate the smallest gradient of electric and magnetic fields, distinguishing between the columns and counting

atoms along the columns. Below is a brief discussion on the phase detection limit in both off-axis and in-line electron holography in the context of the present analysis.

In electron holography, following the procedure as described by Lichte [38], the phase detection limit in a medium resolution hologram is given by

$$\sigma_\varphi = \frac{\sqrt{2e}}{p\sqrt{V^2 j_0 \tau STE(u_c)}} \tag{2.05}$$

Where $N = \frac{1}{e} j_0 p^2 \tau$, e is the charge of the electron, V is the fringe visibility, $STE(u_c)$ is the signal transfer efficiency of the CCD camera, and j_0 is the current density during the exposure time τ over the area p^2. The three essential parameters in Eq. 2.05 are N, p, and V. The fundamental limit in phase detection is governed by the shot noise or stochastic impacts of single electrons due to the probabilistic nature of the electron wave. This is given by

$$\sigma_\varphi = \sqrt{\frac{2c}{V^2 N}} \tag{2.06}$$

The fundamental phase detection limit improves with the increasing electron dose N. Lichte has shown that for V=40%=0.4, STE=0.8, and N=9000/nm^2 σ_φ is 0.0314rad for p^2=1nm^2. c is 1 for the images which contain a vacuum. However, almost no changes in the phase detection limit by improving contrast up to 0.8, but decreases with a further increase in contrast.

In our experimental holograms, acquired in Berlin, the average electron dose is 16 ×10^6 /nm^2, V=15%, and fringe spacing (s) is 0.0469 nm. Therefore, for a reconstructed area p^2=100nm^2 (512 X 512), $\sigma_\varphi = 0.00023$ rad. The minimum reconstruction area for the present data is approximately $(4 \times 0.8)^2$ =(0.32 nm)^2, where 0.8 Å is the point resolution of the microscope. Thus, the phase detection limit for the smallest reconstruction area, corresponding to the present data, is 0.007365 rad. The theoretical model suggests, the change in peak (mean) phase due to incremental change in Zn and O atoms in the atomic column are 0.284 (0.138) and 0.1098 (0.094) rad, respectively. The reference values given in the parenthesis are the peak phase values obtained from the calculation considering the isolated atom model. The mean phase has been calculated for an inner and outer cut-off potential of 1-50 pm for the Zn atom. Therefore, it is possible to count the incremental Zn and O atoms in the atomic columns of the ZnO from the present atomic resolution holography data irrespective of the area of reconstruction.

Lichte [38] has shown that the phase detection limit improves with increasing electron dose N nm^{-2} and lateral resolution p of the reconstructed wave. At atomic resolution, the width of the hologram (w_{hol}) is related to the resolution q_{max} ($w_{hol} \geq 4\,psf$), where *psf* is the point spread function. Therefore, the above equation is modified for a C_s corrected microscope to

$$\delta\varphi_{lim} = \frac{4\sqrt{\pi}\, snr\, C_s}{|\mu|.V_{inel}.V_{inst}.V_{MTF}\sqrt{-\ln(|\mu|)\frac{B_{ax}}{ek^2}\varepsilon t DQE(q_c)}} \times \frac{q_{max}^4}{k^3} \tag{2.07}$$

Where *snr* is the signal to noise ratio, q_{max} is the resolution in reciprocal space, μ is the degree of spatial coherence, V_{inel}, V_{inst}, V_{MTF} are the hologram contrast arising due to inelastic scattering, instabilities, and Modulation Transfer Function (MTF) of the CCD, respectively. $DQE(q_c)$ is the Detection Quantum Efficiency of the CCD camera, B_{ax} is the brightness of the electron source, ε is the illumination ellipticity, k and e are the wavenumber and charge of the electron, respectively [11].

The phase detection limit changes within the same order of magnitude with small variations in *V*, *N*, and *p*, thus not affecting the counting of both Zn and O atoms. The dependency of the theoretical phase detection limit on *V*, *N*, and *p* is shown in Figure 2.06. The dashed vertical lines are marked corresponding to the current experimental parameters. Geiger et al. [11] reported a phase detection limit of 2π/80 (0.0785) rad for an aberration-corrected holography microscope. Cooper and Voelkl improved the phase detection limit to 0.001 and 2π/1000 (0.00628) rad by long exposure and multiplicity of holograms together with bi-prism and sample drift correction, respectively [39,40]. However, a double biprism setup was not used, eliminating Fresnel fringe and significantly improving the phase detection limit. However, there is another limit posed by reconstruction methods where standard deviation in vacuum phase value poses experimental phase detection limit.

On the other hand, in HRTEM, the phase detection limit has not been discussed. Experimentally, distinguishing B and N atoms in monolayer BN has been reported with corresponding peak phase values of 0.022 and 0.026 rad, respectively, with a difference of 0.004 rad between the two atoms [Figure 2.07] [41]. Both the shape and contrasts are responsible for the detection of atoms in this case. The peak phase value on the atom position depends on the atomic scattering factor, structure factors, microscope transfer function, resolution, damping envelopes,

and modulation transfer function (MTF) of the camera. This will be reflected in the recorded image intensity as well. The changes in peak values for both phase and intensity images can be calculated theoretically. However, there is another factor, the standard deviation in the vacuum phase value from the reconstruction method eventually determines the experimental phase detection limit. Experimentally, the standard deviation of the reconstructed phase image in the vacuum will limit the interpretable phase change, which is 0.023 rad from the present result. In the case of off-axis holography reconstruction, the number is better, i.e., 0.00488 rad.

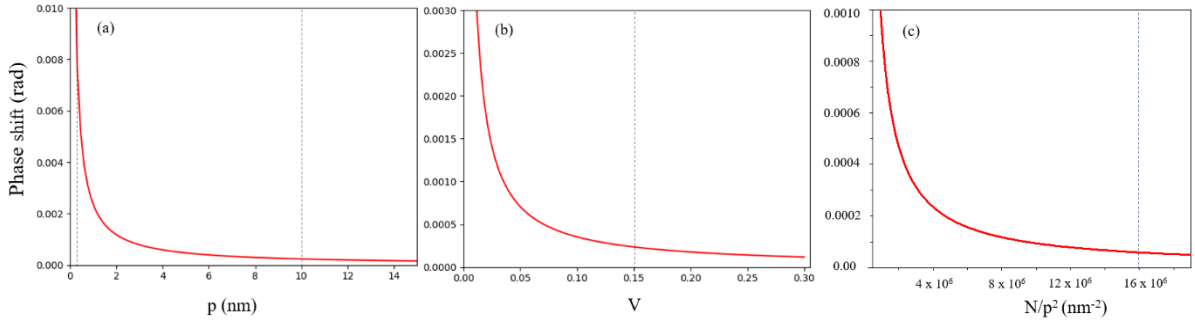


Figure 2.06. Dependence of theoretical phase detection limit on area of reconstruction (p^2), visibility and electron dose are given. (a) Phase shift v/s pixel size ($N/p^2=16\times10^6$ /nm^2 V=15%), (b) phase shift v/s visibility ($N/p^2=16\times10^6$ /nm^2 p=10nm), (c) phase shift v/s electron dose (p=10nm, V=15%).

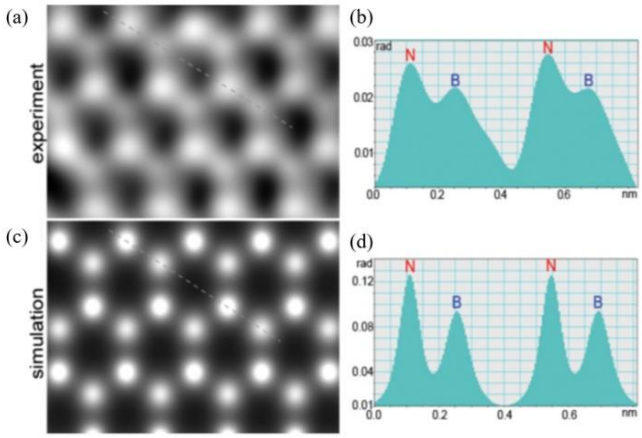


Figure 2.07. (a) Experimental image of monolayer BN with (b) line profile. (c) Simulated image with (d) line profile. For a BN monolayer, the peak phase shift value of B and N are 0.022/0.09 and 0.026/0.13, with a difference of 0.004/0.04 rad by experiment and simulation. The difference between the simulation and experimental values is because of Stobb's factor. Copyright (2009) by American Physical Society [41].

2.3.2 Atomic potential model

It is necessary to compare the reconstructed phase shift results with the theoretical reference values to quantify the number of atoms. This involves modeling the atomic potential as an imaging electron directly interacts with it, resulting in the object exit wave (OEW) function. Moreover, the lens contrast transfer function (CTF) and aperture size (*k*) modify the phase of the OEW further on the way to the recording device. The nucleus size (1.6 to 15 *fm*) is tiny compared to the atom's size. An atom consists of a nucleus and surrounding clouds of electrons which may extend to 0.1 to 0.5 nm. For a stationary atom, the Coulomb potential is $\propto \frac{1}{r}$ and there is a singularity at the center of the atom. The imaging electrons mostly see the nuclear potential, and the surrounding electrons shield the effect [42]. Inelastic events are negligible compared to elastic events (imaging electrons) for a thin sample. Various theoretical atomic potential models are available in the literature [42–44]. In the present investigation, Hartree-Fock isolated atomic potential projected along the *z*-direction is considered [17]. The atomic potential in 3D is given by,

$$V_a(x,y,z) = 2\pi^2 a_0 e \sum_{i=1}^{3} \frac{a_i}{r} \exp\left(-2\pi r\sqrt{b_i}\right) + 2\pi^{5/2} a_0 e \sum_{i=1}^{3} c_i {d_i}^{-3/2} \exp\left(-\pi^2 r^2/d_i\right) \quad (2.08)$$

with $r^2 = x^2 + y^2 + z^2$. Where, a_0 is the Bohr radius, a_i, b_i, c_i, d_i are the parameterized coefficients. Then the mean phase shift in the absence of dynamical scattering is calculated by the equation,

$$\Phi(x,y) = \sigma \int V(x,y,z)dz \quad (2.09)$$

Where $\sigma = \frac{\pi}{\lambda E}$ is the interaction parameter with wavelength λ and accelerating voltage *E* [1]. Projected potential along the *z*-direction, based on the Hartree-Fock isolated atom model, is given by

$$v_z(x,y) = \int_{-\infty}^{+\infty} V_a(x,y,z)dz$$

$$= 4\pi^2 a_0 e \sum_{i=1}^{3} a_i K_0\left(2\pi r\sqrt{b_i}\right) + 2\pi^2 a_0 e \sum_{i=1}^{3} \frac{c_i}{d_i} \exp\left(-\pi^2 r^2/d_i\right) \quad (2.10)$$

with $r^2 = x^2 + y^2$. $K_0(x)$ is the modified Bessel function of order zero [17].

Various resolution limiting factors such as diffraction limit, thermal vibration, and aberration of the microscope result in measurable peak phase value in the phase image of the atom.

The phase image of the atom can be approximated to a Gaussian function. A Gaussian function is parameterized by the peak height and the full-width half maxima (FWHM). Therefore, the reference phase shift can be considered either based on the peak value of the phase or the mean value of the phase. The mean value will depend on both the peak value and FWHM of the phase distribution function. To calculate the mean value of the phase, the mean inner potential is required. The mean value of potential is calculated by integrating either two-dimensional projected potential or three-dimensional atomic potential between the two limits and dividing with the surface area/volume, respectively. The mean phase is then calculated by multiplying this with the interaction parameter σ. The mean phase values calculated for different inner and outer cut-off values are given in Table 2.01, and the profile for Zn and O atomic potential is shown in Figure 2.08 (a). The theoretical mean phase values corresponding to Zn and O atoms are considered for an outer bound of 50 and 25 pm, respectively, for Zn and O atoms keeping the same values with the experimental outer cut-off.

Inner-Outer cut off (pm)	Mean phase shift (rad)	
	Zn	O
0.1-100	0.056	0.022
1-100	0.054	0.022
5-100	0.054	0.019
10-100	0.051	0.021
5-50	0.133	0.055
5-60	0.107	0.045
5-80	0.073	0.031
10-50	0.122	0.052
10-60	0.099	0.043
10-80	0.069	0.030
1-50	0.138	0.056
1-25	0.269	0.094

Table 2.01. The mean phase shift for the same outer cut-off and varying inner cut-off, same inner cut-off with a varying outer cut-off for Zn and O atoms.

The above calculation is based on the weak phase object approximation of isolated atoms. Any phase change due to more than one atom is calculated via the linear projection of potential. We have plotted the theoretical peak reference value obtained through the MacTempas package

for comparison purposes [Figure 2.08 (b) and (c)]. The multislice method employed in the MacTempas code considers the electron propagation between the successive slices, which account for the dynamical scattering by the crystal. The slicing method for fewer atoms may be responsible for nonlinear phase increment than the linear projection of the potential method.

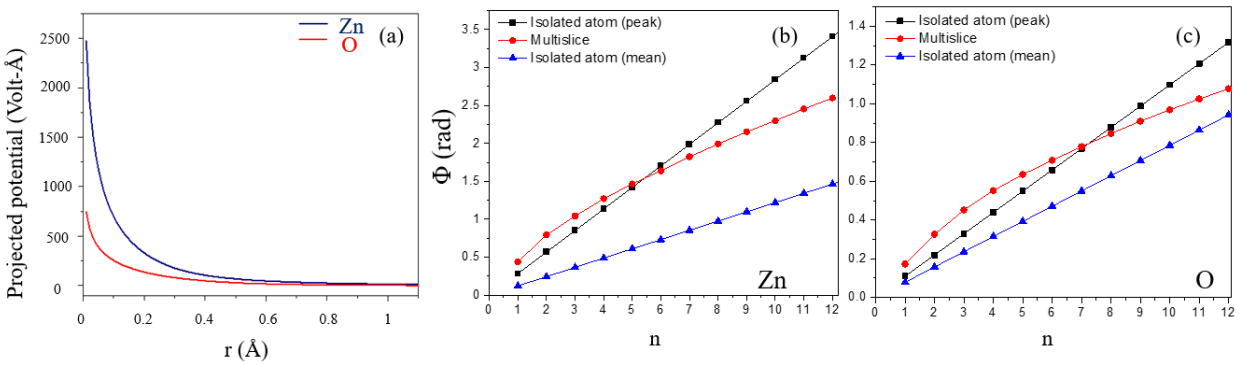


Figure 2.08. (a) The projected potential of Zn and O atoms calculated using Eq. 2.10. The peak phase shift of (b) Zn and (c) O atoms with the increase in the number of atoms in the column calculated using the isolated atom model and multislice method considering dynamical scattering. The resolution was set to 0.5 Å. The mean phase shift is plotted with the inner and outer bound of potential 1 to 50/25 pm (Zn/O), respectively. Copyright (2019) by American Institute of Physics [37].

The potential function has a singularity at $r = 0$ due to $\frac{1}{r}$ dependence. Therefore, it is necessary to consider an inner and outer bound of the potential while calculating the phase shift and image of the atom. The atomic scattering factor $f_e(k)$ (according to Moliere) and image of the atoms depends on the inner and outer bound of the potential [Figure 2.09. However, it is observed that there is a limit in both inner and outer bounds, beyond which the $f_e(k)$ and peak values of the atom image do not change significantly. In the present report, the limits corresponding to inner and outer bound are 0.01 and 1 Å, respectively, considered for calculating images of Zn and O atoms.

According to Moliere, the atomic scattering factor is given by Eq. 5.18 in reference [17],

$$f_e(k) = \frac{2\pi i}{\lambda}\int_0^{\infty} J_0\,(2\pi kr)\,\{1 - \exp\,[i\sigma \int V(x,y,z)dz]\}rdr \qquad (2.11)$$

Where $r^2 = x^2 + y^2$, $J_0(x)$ is the Bessel function of order zero.

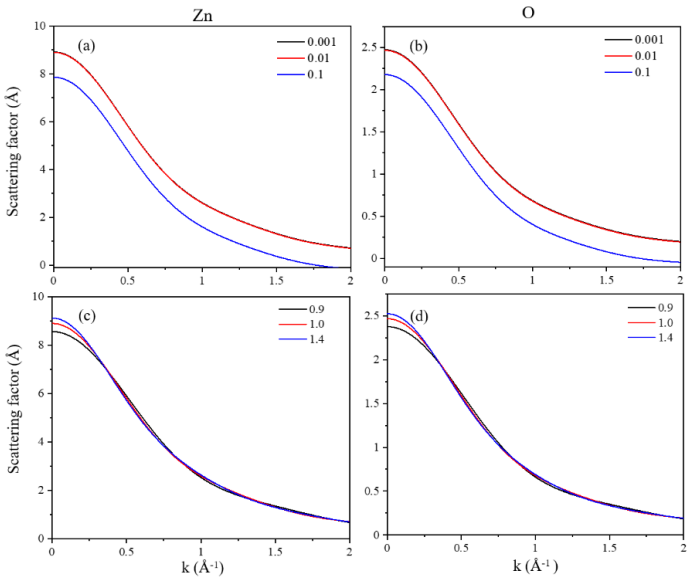


Figure 2.09. The scattering factor variation with the inner and outer bound of the potential for both Zn and O atoms. (a) and (b) show no change in scattering factor by changing the inner cut-off from 0.001 to 0.01 Å (for the same outer cut-off 1 Å), but changes to the inner cut-off of 0.1 Å for both Zn and O, respectively. On the other hand, keeping the inner cut-off fixed (0.01 Å), there is only a change in amplitude at a small scattering angle (<0.25 Å $^{-1}$) by changing the outer cut-off for Zn (c) and O (d) atoms.

The image of an isolated atom based on the above mentioned model potential can be calculated directly using electron scattering amplitude from the following equation [17];

$$g(x) = \left|1 + 2\pi i \int_0^{k_{max}} f_e(k) \exp[-i\chi(k)] J_0(2\pi kr) k dk\right|^2 \quad (2.12)$$

Where, $f_e(k)$ is the electron scattering factor in the Moliere approximation using the projected atomic potential. $\chi(k)$ is the aberration function, $k_{max} = \alpha_{max}/\lambda$ is the maximum spatial frequency transferred by the objective aperture, and $J_0(x)$ is the Bessel function of order zero [17].

The effect of the inner and outer bound of the potential and the objective aperture diameter on the shape and peak values of phase shift and image intensity for a single atom are given in Figure 2.10. The real part of the wave transfer function $\cos(\chi)$ is neglected (i.e., set to zero) and the imaginary part $sin(\chi)$ is set to 1 to mimic the Scherzer-like transfer function for the atoms in a periodic lattice within a weak phase object approximation.

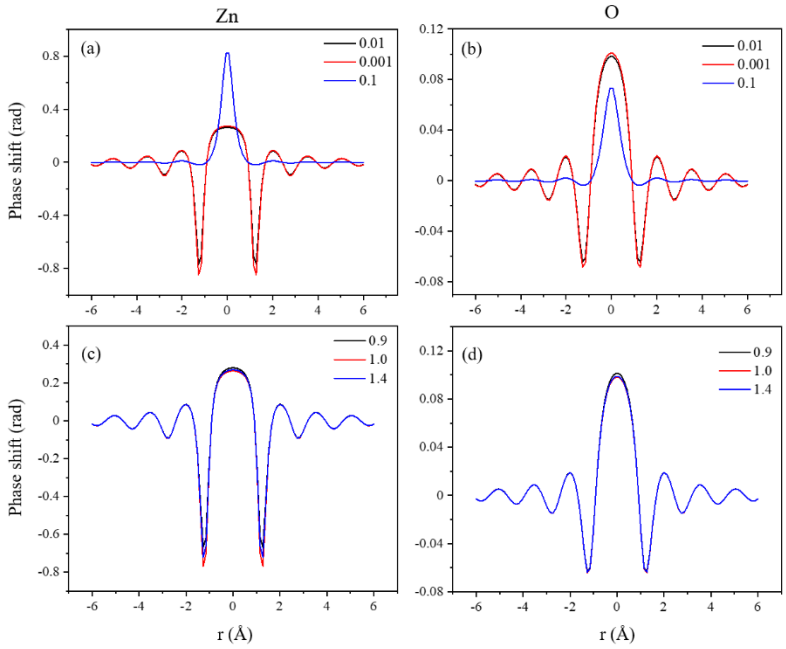


Figure 2.10. (a) and (b) The peak phase of the Zn and O atom as a function of inner cut-off (fixed outer cut-off 1 Å) and (c) and (d) as a function of outer cut-off (fixed inner cut-off, 0.01Å). The peak phase value does not change with the outer cut-off potential from 0.9 to 1.4 Å. Peak phase values also do not change for the inner cut-off of 0.001 and 0.01 Å but change significantly for 0.1Å.

Like the scattering factor, the peak value of phase and intensity does not change significantly below an inner cut-off value of 0.01Å. No significant change is observed with the outer bound. This is true for both Zn and O atoms. The peak phase and intensity values also depend on the aperture size [Figure 2.11]. In the present case, an aperture size of 2 $Å^{-1}$ is used.

Two different theoretical phase values, peak and mean for a given atom, can be considered to interpret the reconstructed phase for counting the number of atoms. However, atoms are never stationary in the crystal, and due to finite temperature, they oscillate (0.0073 and 0.0072 Å, for Zn and O atoms in ZnO at 293K [45] about their mean position in the lattice. Aberrations in the microscope cause further blurring on the atom positions. However, the amplitude of thermal vibration at room temperature and resulting blurring is smaller than the blurring due to aberration and is not considered in the present investigation (frozen phonon). By numerical evaluation, one

can find that the peak phase shift values have a coarse dependence of $Z^{0.6\text{-}0.7}$, and deviation can be observed due to valence electron filling with the atomic number [46].

On the other hand, the mean phase shift value is sensitive to the inner and outer bound of the potential (Table. 2.01). The mean phase shift value does not change enormously with the inner and outer cut off below 0.01 Å and above 1 Å of the potential, respectively, but varies significantly in between. An inner cut at 1 pm and an outer cut-off of 50/25 pm corresponding to the size of the Zn and O atoms are considered for extraction of mean phase shift from the experimental data. The theoretical mean phase shift value is calculated by three-dimensional integration of the potential and dividing with the volume bounded by the limits. The peak phase shift values for a microscope resolution of 0.8Å obtained from literature and multislice calculation as implemented in MacTempas for the atoms in a crystal along with mean phase shift values are given in Figure 2.08 (b) and (c). The two curves corresponding to peak phase values match well for fewer atoms in a column but deviate from linearity due to dynamical scattering for a higher number of atoms. The mean phase is smaller (~ factor of 0.5) than the peak phase value, which has implications on the number of atoms by two different reference parameters and is discussed next.

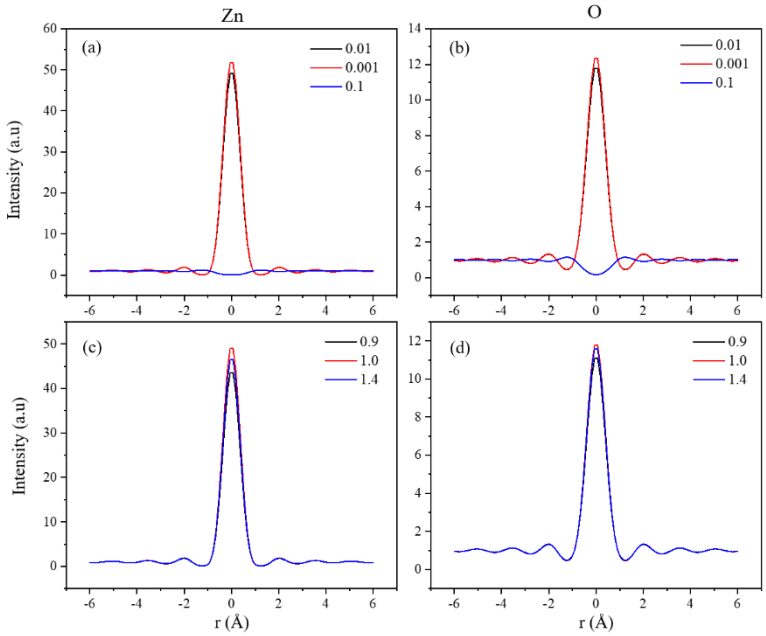


Figure 2.11. (a) and (b) Intensity plots of Zn and O atom as a function of inner cut-off (fixed outer cut-off 1 Å) and (c) and (d) as a function of outer cut-off (fixed inner cut-off, 0.01Å). The intensity value does not change with the outer cut-off potential from 0.9 to 1.4 Å.

2.3.3 Atomic resolution off-axis electron holography

This section analyzes the experimental phase information retrieved from both SB and CB off-axis electron holograms of ZnO film with varying thicknesses. ZnO films with two different thicknesses along $[11\bar{2}0]$ and $[01\bar{1}0]$ Z.A. are considered. The extinction distances (ξ_g) are 89.89 and 90.59 nm for $[11\bar{2}0]$ and $[01\bar{1}0]$ Z.A. respectively for the 0002 reflections. Figure 2.12 and Figure 2.15 are the amplitude and phase images corresponding to CB and SB obtained for the two different regions marked as P and Q. The peak phase values on top of the Zn and O columns have been evaluated. Selected columns at three different distances corresponding to different thickness levels from the edge of the specimen for area P are indicated in the figures. The same columns are considered for the comparisons of two different OEW reconstructed using the CB and SB. The number of atoms corresponding to Zn, evaluated from the peak and mean phase values, are plotted in Figure 2.13 for area P. Difference in the number of Zn atoms between CB and SB is within ±1 and ±4, corresponding to reference peak and mean phase, respectively. Another noticeable point is that a similar amount of Zn and O atoms are obtained for different areas of region P. This suggests that peak phase values used from the theoretical model fit well for both light and heavy atoms adjacent to each other. The reconstructed phase values and the corresponding number of Zn and O atoms in the neighboring columns for area P are given in Figure 2.14. The number of atoms is in close agreement (with the difference of ±1 atom) for Zn and O.

Figure 2.16 compares phase shift and the number of Zn atoms reconstructed by CB and SB for region Q. There is a gradual increase in the number of atoms for SB, but an almost constant (but different than the SB) number of atoms are obtained from the CB. Thus, for region Q, the match is poor between CB and SB because of the relatively thicker sample inducing a stronger dynamical effect. Reconstructed OEW for regions P and Q from both SB and CB are given in Figure 2.17 for comparison.

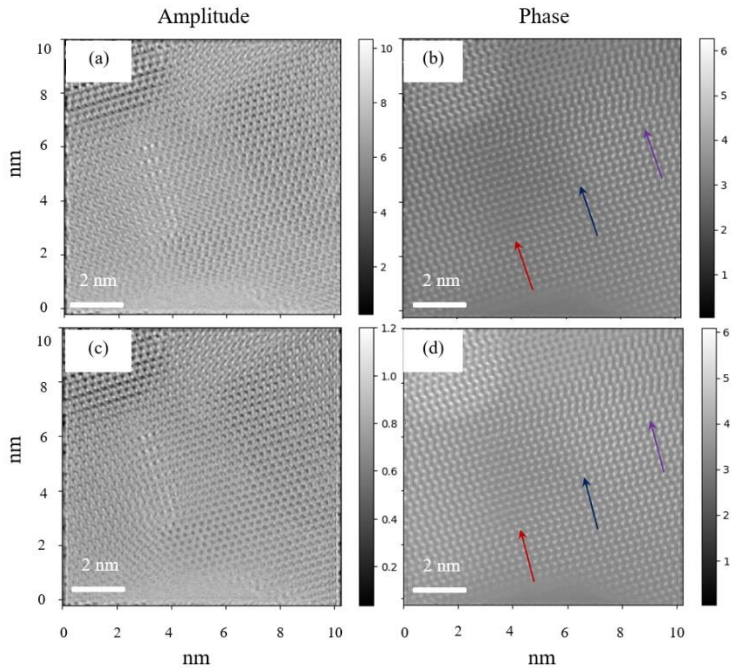


Figure 2.12. (a) and (c) Amplitude and (b) and (d) phase image of ZnO along [11$\bar{2}$0] Z. A. for area P obtained through reconstruction CB and SB of off-axis electron hologram, respectively. Three different arrows are indicated in the phase image, along which the peak and mean phase values are extracted. Copyright (2019) by American Institute of Physics [37].

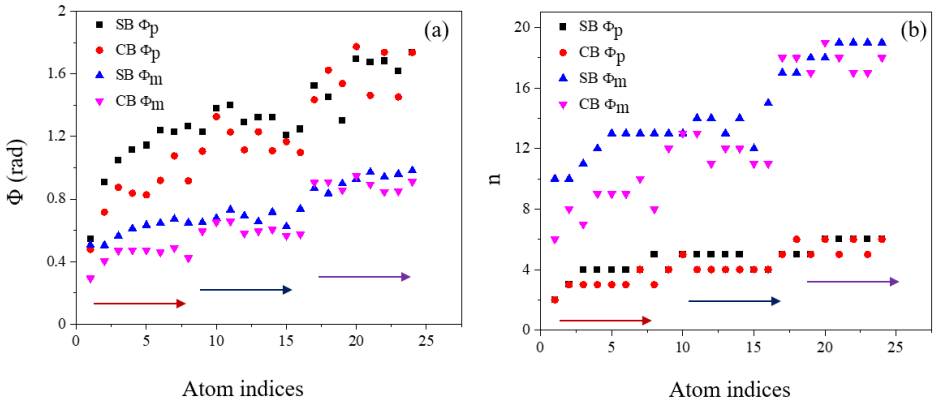


Figure 2.13. (a) Peak (Φ_p) and mean (Φ_m) phase shift and (b) the corresponding number of Zn atoms along three different arrows from area P reconstructed from SB and CB. The number of Zn atoms coincides between SB and CB with ±1 atom error. However, the number of atoms derived from the mean phase value is three times higher than the peak phase value. Copyright (2019) by American Institute of Physics [37].

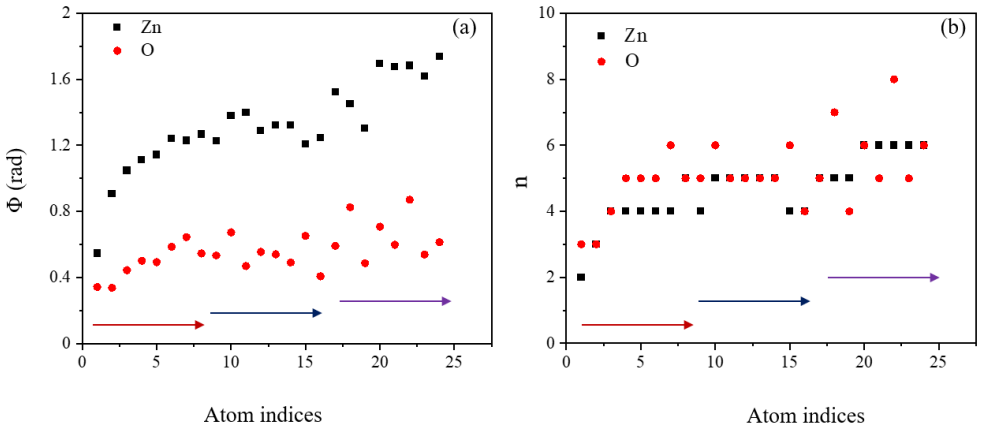


Figure 2.14. Comparison of (a) peak phase shift and (b) the corresponding number of atoms with the variation of thickness in Zn and O columns for area P. Almost a similar number of atoms are obtained for Zn and O atoms at the neighboring sites. Copyright (2019) by American Institute of Physics [37].

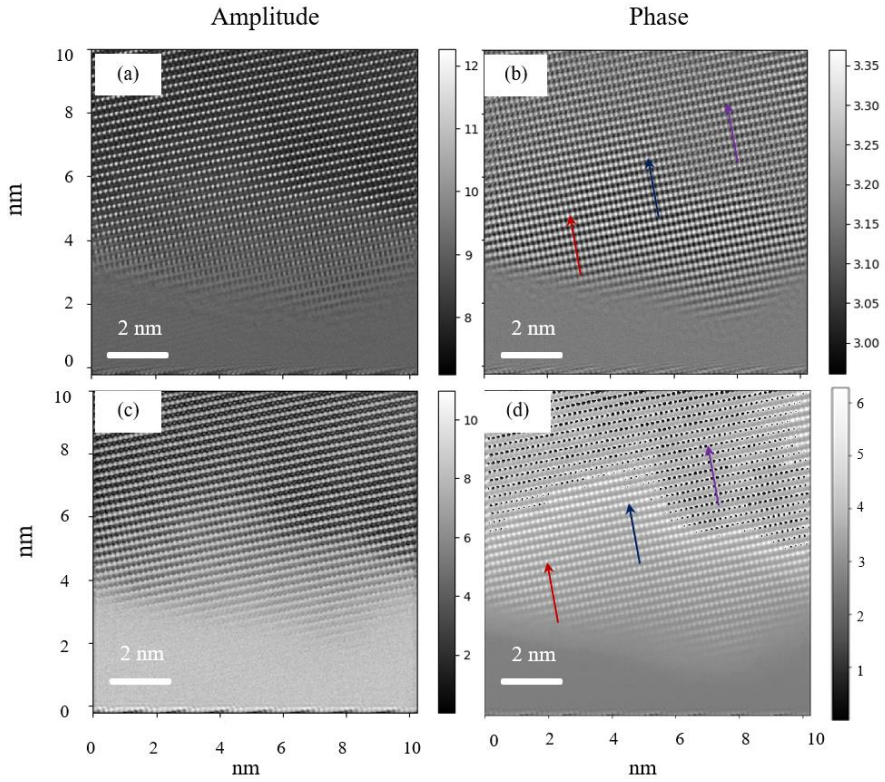


Figure 2.15. (a) and (c) Amplitude and (b) and (d) phase image of ZnO along [11$\bar{2}$0] Z. A. for area Q obtained through reconstruction CB and SB of off-axis electron hologram. Three different arrows are indicated in the phase image, along which the peak and mean phase values are extracted. Copyright (2019) by American Institute of Physics [37].

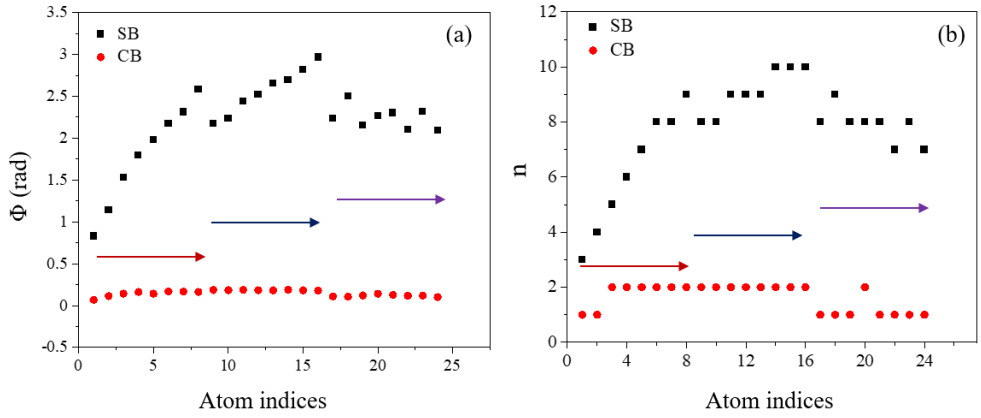


Figure 2.16. (a) Peak (Φ_p) phase shift and (b) the corresponding number of Zn atoms along three different arrows from area Q reconstructed from SB and CB. No agreement is found in the number of Zn atoms between SB and CB. Copyright (2019) by American Institute of Physics [37].

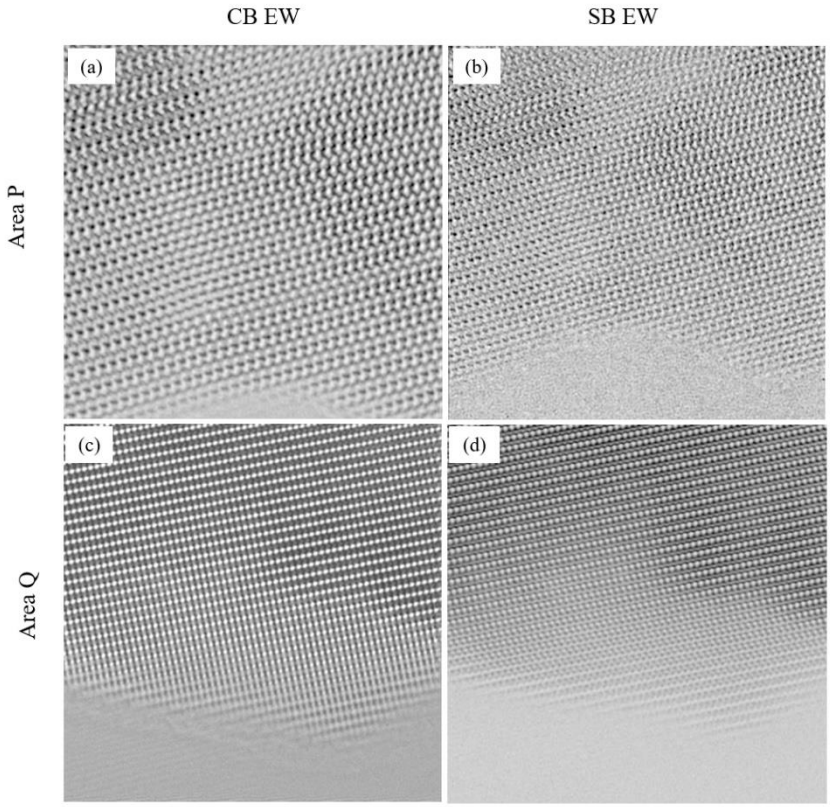


Figure 2.17. Reconstructed wave function of area P (a) and (b), and area Q (c) and (d) from central band (a), (c) and (b), (d) sideband.

2.3.4 In-line electron holography/HRTEM

Figure 2.18 shows the reconstructed phase image of ZnO film along $[11\bar{2}0]$ orientation from different thickness areas of the sample. The peak phase values from Zn columns are given in the line scan for the columns indicated in the image. The peak phase values remain almost the same between thinner and thicker regions of the sample at around 0.138 rad, which corresponds to approximately one Zn atom. In the case of O, the peak phase value remains the same at about 0.09 rad, which corresponds to approximately one atom.

Niermann et al. [14] first described the difference between in-line and off-axis electron holography at atomic resolution in GaAs crystal along $[1\bar{1}0]$ Z.A. The phase and amplitude reconstructed from the SB and CB agree well up to a thickness of 3/2 times the extinction length. Still, significant deviations were observed at the lower frequency range and thicker specimen area. The agreement between the two methods for the thinner area was due to the similar wave function reconstructed in the limit of linear imaging with negligible inelastic scattering. However, due to significant inelastic scattering, reconstruction methods corresponding to CB and SB yielded two different wave functions for the thicker area. SB's mathematical formulation contains an average OEW function, while CB contains the sum of the squared OEW function. It was mentioned that the deviation observed in the thicker area between CB and SB reconstructed wave function may be either due to fundamental quantum mechanical differences or numerically challenging inversion of the imaging process. Our results agree with Niermann et al. [14] observation for the CB and SB reconstruction of the off-axis experiment in the present investigation. However, an observation made in the HRTEM experiment is not comparable with the CB reconstruction. This could be because of the reconstruction scheme employed in the reconstruction package.

The counting of atoms depends on the theoretical reference phase values, i.e., mean or peak values of phase. We obtain ~ 3 times higher number of atoms for Zn and O using reference mean phase than peak phase value. The experimentally observed higher mean phase values than theory could be because of incoherent aberrations, vibrations present, any shortfall in modeling electron materials interaction and electron-optical systems, or any other reasons.

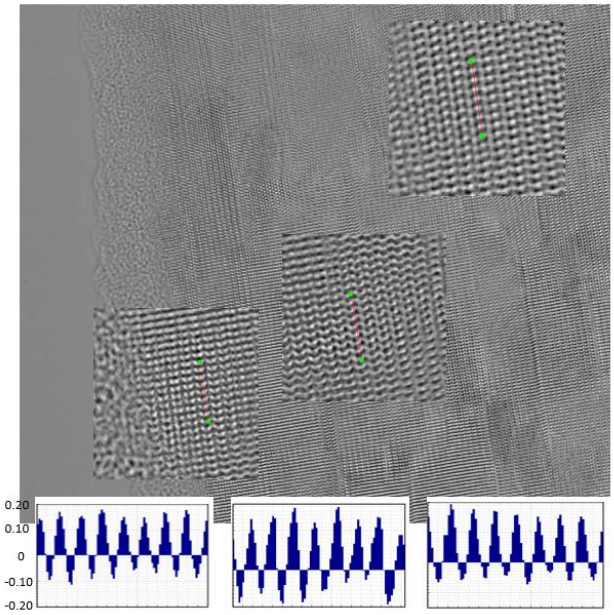


Figure 2.18. Reconstructed phase image of HRTEM image series for different thickness regions. Throughout the sample area, almost constant phase shift and atom number corresponding to one are obtained. Copyright (2019) by American Institute of Physics [37].

2.4 Conclusion

In conclusion, the atomic resolution reconstructed phase of Zn and O atoms in ZnO epitaxial thin film is compared between off-axis and in-line holography techniques. While the off-axis holography method shows an excellent agreement in the number of atoms for both Zn and O atoms extracted from SB and CB for thin sample area, however, for thicker samples, the number of atoms does not reconcile. In HRTEM reconstruction, the number of atoms does not change systematically with increasing sample thickness. A constant number of one atom is obtained throughout the reconstructed area.

2.5 Bibliography

[1] D. B. W. Barry C. Carter, *Transmission Electron Microscopy Diffraction, Imaging, and Spectrometry* (2017).

[2] H. Lichte and M. Lehmann, Reports Prog. Phys. **71**, 016102 (2008).

[3] W. Coene, G. Jannsen, M. O. d. Beeck, and D. V Dyck, Phys. Rev. Lett. **69**, 3743 (1992).

[4] D. Van Dyck and M. Op de Beeck, Proceedings of the XIIth International Congress on Electron Microscopy, pp. 26–27 (1990).

[5] D. Gabor, Nature **161,** 777 (1948).

[6] E. N. Leith and J. Upatnieks, J. Opt. Soc. Am. **52**, 1123 (1962).

[7] G. Möllenstedt and H. Düker, Zeitschrift Für Phys. **145**, 377 (1956).

[8] S. Frabboni, G. Matteucci, G. Pozzi, and M. Vanzi, Phys. Rev. Lett. **55**, 2196 (1985).

[9] M. R. McCartney, D. J. Smith, R. Hull, J. C. Bean, E. Voelkl, and B. Frost, Appl. Phys. Lett. **65**, 2603 (1994).

[10] D. Cooper, R. Truche, a. C. Twitchett-Harrison, R. E. Dunin-Borkowski, and P. a. Midgley, J. Microsc. **233**, 102 (2009).

[11] D. Geiger, H. Lichte, M. Linck, and M. Lehmann, Microsc. Microanal. **14**, 68 (2008).

[12] M. Linck, B. Freitag, S. Kujawa, M. Lehmann, and T. Niermann, Ultramicroscopy **116**, 13 (2012).

[13] H. Lichte, Ultramicroscopy **20**, 293 (1986).

[14] T. Niermann and M. Lehmann, J. Phys. D. Appl. Phys. **49**, 194002 (2016).

[15] G. Schaal, C. Jönsson, and EF Krimmel, Optik (Stuttgart) **24**, 529 (1966).

[16] K. Harada, A. Tonomura, Y. Togawa, T. Akashi, and T. Matsuda, Appl. Phys. Lett. **84**, 3229 (2004).

[17] E. J. Kirkland, *Advanced Computing in Electron Microscopy: Second Edition*, 2nd ed. (Springer, 2010).

[18] W. O. Saxton, Ultramicroscopy. **55,** 171 (1994).

[19] W. M. J. Coene, A. Thust, M. Op De Beeck, and D. Van Dyck, Ultramicroscopy **64**, 109

(1996).

[20] R. KILAAS and R. GRONSKY, Ultramicroscopy **11**, 289 (1983).

[21] P. W. Hawkes, Comput. Des. **5**, 200 (1973).

[22] D. Van Dyck, H. Lichte, and K. D. Van Der Mast, Ultramicroscopy **64**, 1 (1996).

[23] C. T. Koch and A. Lubk, Ultramicroscopy **110**, 460 (2010).

[24] T. Latychevskaia, P. Formanek, C. T. Koch, and A. Lubk, Ultramicroscopy **110**, 472 (2010).

[25] C. Ozsoy-Keskinbora, C. B. Boothroyd, R. E. Dunin-Borkowski, P. A. Van Aken, and C. T. Koch, Sci. Rep. **4**, 7020 (2014).

[26] D. Van Dyck, J. R. Jinschek and F.R. Chen, Nature **486**, 243 (2012).

[27] K. Ishizuka and B. Allman, J. Electron Microsc. (Tokyo). **54**, 191 (2005).

[28] C. L. Jia, S. B. Mi, J. Barthel, D. W. Wang, R. E. Dunin-Borkowski, K. W. Urban, and A. Thust, Nat. Mater. **13**, 1044 (2014).

[29] S. Borghardt, F. Winkler, Z. Zanolli, M. J. Verstraete, J. Barthel, A. H. Tavabi, R. E. Dunin-Borkowski, and B. E. Kardynal, Phys. Rev. Lett. **118**, 086101 (2017).

[30] A. Wang, F. R. Chen, S. Van Aert, and D. Van Dyck, Ultramicroscopy **116**, 77 (2012).

[31] A. Lubk, A. Béché, and J. Verbeeck, Phys. Rev. Lett. **115**, 176101 (2015).

[32] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Großß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser, Nat. Mater. **10**, 209 (2011).

[33] F. Winkler, J. Barthel, A. H. Tavabi, S. Borghardt, B. E. Kardynal, and R. E. Dunin-Borkowski, Phys. Rev. Lett. **120**, 156101 (2018).

[34] B. Loukya, P. Sowjanya, K. Dileep, R. Shipra, S. Kanuri, L. S. Panchakarla, and R. Datta, J. Cryst. Growth **329**, 20 (2011).

[35] R. Sahu, K. Dileep, B. Loukya, and R. Datta, Appl. Phys. Lett. **104**, (2014).

[36] H. Lichte, Ultramicroscopy **64**, 79 (1996).

[37] U. Bhat and R. Datta, J. Appl. Phys. **125**, 154902 (2019).

[38] A. Harscher and H. Lichte, Ultramicroscopy **64**, 57 (1996).

[39] D. Cooper, R. Truche, P. Rivallin, J. M. Hartmann, F. Laugier, F. Bertin, A. Chabli, and J. L. Rouviere, Appl. Phys. Lett. **91**, 143501 (2007).

[40] E. Voelkl and D. Tang, Ultramicroscopy **110**, 447 (2010).

[41] C. Jin, F. Lin, K. Suenaga, and S. Iijima, Phys. Rev. Lett. **102**, 195505 (2009).

[42] T. G. Strand and T. Tietz, Nuovo Cim. B Ser. 10 **41**, 89 (1966).

[43] P. A. Childs and D. L. Misell, J. Phys. C Solid State Phys. **4**, (1971).

[44] WILI.IAM J. HYATT, Phys. Rev. **104**, 1298 (1956).

[45] K. Kihara and G. Donnay, Can. Mineral. **23**, 647 (1985).

[46] M. Linck, H. Lichte, M. Lehmann, Int. J. Mat. Res. (formerly Z. Metallkd.), **97,** 890 (2006).

Chapter 3

Image simulation in HRTEM considering atom as an electrostatic charge center

This chapter introduces an alternative approach to image simulation in high-resolution transmission electron microscopy after a comparative analysis of the existing image simulation methods. The alternative method is based on considering the atom center as an electrostatic interferometer akin to the conventional off-axis electron biprism within few nanometers of focus variation. Simulation results are compared with the experimental images of 2D materials of MoS_2 and BN recorded under the optimum combination of third-order spherical aberration $C_s = -35$ μm and defocus $\Delta f = 1$, 4, and 8 nm and are in good agreement.

Manuscript accepted by Journal of Physics Communication.

***U. Bhat** and R. Datta, 2021 J. Phys. Commun. https://doi.org/10.1088/2399-6528/ac1839*

3.1 Introduction

Understanding the high-resolution transmission electron microscope (HRTEM) images is not straightforward. What we see is only the two-dimensional projection of the three-dimensional specimen. It is necessary to know the structure of the object thoroughly to interpret the information contained in the micrograph. In all ways, visualization is proved to be the most effective way to realize things. Hence it requisite to visualize the specimen. Simulation makes it easy to know the object from different perspectives.

Imaging any object in transmission and reflection geometry is generally carried out by detecting the scattered (incoherent) and diffracted (coherent) radiation on a recording device, e.g., a camera placed at different reference planes away from the object on the optic axis. Image is the replica of the object and not the object itself, and the information about the object is carried to the detector through the complex wave function [1–5]. Maximum spatial details that can be obtained are limited by the diffraction and the microscope performance. The entire topic of quantitative HRTEM falls into two broad categories; (*i*) object exit wave (OEW) reconstruction to retrieve the phase information from the recorded image, and (*ii*) image simulation to interpret the OEW with the object structure. Various schemes are available to reconstruct the OEW function to recover the phase related to the object potential and the crystallography [6–8]. There are several aspects in HRTEM image simulation that need to be considered, e.g., probe electron, the interaction between the fast electron and the specimen potential, lens action, and characteristics of the recording device [5,9]. The probe illumination is typically a plane wave of an electron with relativistic energy in the range of 100-300 kV ($\lambda = 1.97\ pm$ at 300 kV). The amplitude $A(x, y)$ and the phase $\phi(x, y)$ of the OEW function of the form $\psi = Ae^{i\phi}$ extracted from the recorded intensity pattern are used to interpret the potential information of atoms in specimen. The information on object potential can be used to extract wide range of information such as number of atoms or thickness along the beam propagation direction, identification of atoms, valence electron sharing between the atoms etc. [7,10–14]. However, the change in phase (ϕ) of the probe electron wave after interaction with the specimen potential and resulting modulation in intensity pattern has been treated in fundamentally distinctive ways, e.g., (*i*) transmission function based on weak phase object approximation (WPOA) along with Zernike type $\pi/2$ or $\lambda/4$ phase plate equivalent to phase contrast transfer function (PCTF) to account for the lens aberration, where the phase change is incorporated

in terms of change in the magnitude of momentum vector k of the probe electron due to specimen potential [4,5], (*ii*) phase change according to scattering amplitude in terms of atom scattering $f(k)$ and structure factor $F(g)$ along with PCTF [9,15,16], and (*iii*) self-interference in HRTEM and holographic fringe shift in off-axis electron holography where a phase term (±φ) is added inside the trigonometric functions with respect to the reference phase [17,18]. Kindly note that the phase change due to aberration through PCTF is not added with the object wave rather it is applied as frequency filter and point spread function (*psf*) in the diffraction and image planes, respectively.

In this chapter, a comparative analysis is provided on the various existing methods for simulating the image of the atom. It is shown that the different ways of considering phase change in the probe electron wave function due to atomic potential result in different magnitude and intensity patterns for the same atom. Subsequently, an alternative method is introduced where the geometry of interference based on the momentum vector direction is emphasized. The process is based on an atomic potential center as an interferometer akin to the electron biprism within a short range of focus variation (<10 nm) from the reference Gaussian image plane and resembles Abbe's diffraction picture. In this alternative method, phase change has been treated like off-axis electron holography considering the wave interference at an angle. Its analogy with other approaches can be understood with interference geometry and associated momentum vector direction. Simulation results are compared with the experimental image of 2D materials of MoS_2 and BN recorded under specific settings of third-order spherical aberration $C_s = -35$ μm and defocus $\Delta f = 1$, 4, and 8 nm and are found to be in good agreement.

3.2 Coherent image formation at near and far field

A brief discussion is provided on the analogy of slit diffraction patterns in light optics, both near and far-field regimes. A typical Fresnel and Fraunhofer diffraction regimes and corresponding patterns are shown in Figure 3.01. An analytical formula is derived from the Fraunhofer integral and Fourier transformation-based method to evaluate the far-field diffraction pattern in light optics. The analytical formula embodies various parameters, e.g., wavelength (λ) of the illumination, scattering angle (θ), dimension, and periodicity of the slits. The approach is based on the physical picture of path difference and associated constructive and destructive interference between waves with the same momentum vector.

Fourier transformation is an essential mathematical concept to go back and forth between the intensity patterns formed in two reference planes. The results should be consistent independent of the coordinate space and methods chosen for the image calculation. We describe the basic procedures involved with various approaches to calculate the image of atoms and the different results they yield.

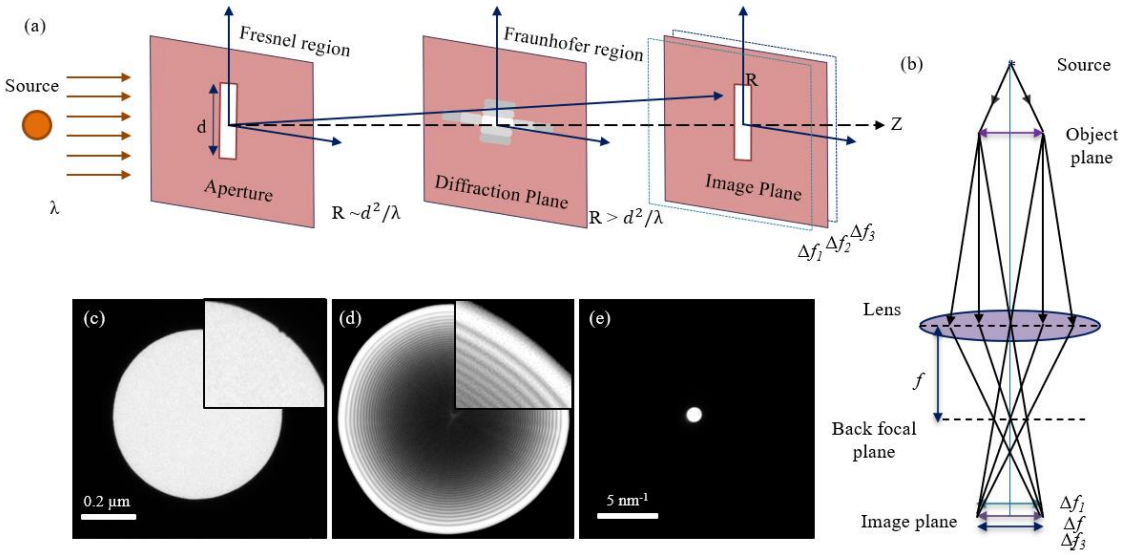


Figure 3.01. (a) Near field (Fresnel) and far field (Fraunhofer) regions of image formation. Also shown the exact Gaussian image plane and Fresnel regimes away from this plane. (b) Schematic showing the role of the lens on the information transfer. (c)-(e) Examples of typical intensity distribution based on electron diffraction as observed at three different regions of interest, (c) at exact focus, (d) slightly away from the focus, and (e) at far-field, respectively. In the illustration, no specimen is used on the electron beam path around the optic axis.

3.2.1 Fresnel diffraction pattern

The diffraction geometry involved in Fresnel zone construction for image formation at the near field is shown in Figure 3.01 (c). Angular correlation between the wave vectors pointing at different directions lying on the surface of a sphere and the phase difference between various wave vectors is acquired due to path difference of waves while converging to a point with respect to outward curvature of the spherical wavefront (more precisely parabolic wavefront) [Figure 3.02 and 3.03]. This is the basis of Fresnel zone construction. Kindly note that there is no path difference

between various wave vectors regarding the emitting point. The radius is the same for a sphere but not so for a parabola and plane wavefront geometry.

The Fresnel diffraction is a near field pattern perpendicular to an interface due to discontinuity in the scattering potential near an edge or interface [1,19]. Fresnel-Huygens principle is based on the following two postulates :

(i) Every point of a wavefront is a source of secondary disturbance giving rise to spherical wavelets, and the propagation of the wavefront is regarded as the envelope of these wavelets.

(ii) The secondary wavelets mutually interfere.

Now applying the above principle, which is based on the geometry as shown in Figure 3.02, each surface element dS of the incident wavefront ψ_{in} at a radial distance R, generates a spherical wavelet contribute an amplitude $\mathrm{d}\psi_{sc}(P)$ at a point P on the optic axis beyond the wavefront.

$$\mathrm{d}\psi_{sc}(P) = -iA(2\theta)\psi_{in}\frac{e^{ikR}}{R}\mathrm{d}S \tag{3.01}$$

after integration over the surface Eq. 3.01 becomes

$$\psi_{sc}(P) = -i\int_{wavefront} A(2\theta)\psi_{in}\frac{e^{ikR}}{R}\mathrm{d}S \tag{3.02}$$

After choosing the appropriate limit over the radial extent, Eq. 3.02 becomes

$$\psi_{sc}(P) = i\frac{\lambda\psi_{in}^{0}}{r_0+R_0}e^{ik(R_0+r)} \tag{3.03}$$

Eq. 3.03 above describes the propagation of the spherical wave. In HRTEM image simulation similar spherical wave is used in superposition with the incident plane wavefront in the formulation of transmission function based on a solution of Schrödinger equation in integral form (sec. 3.2.2) [9,19].

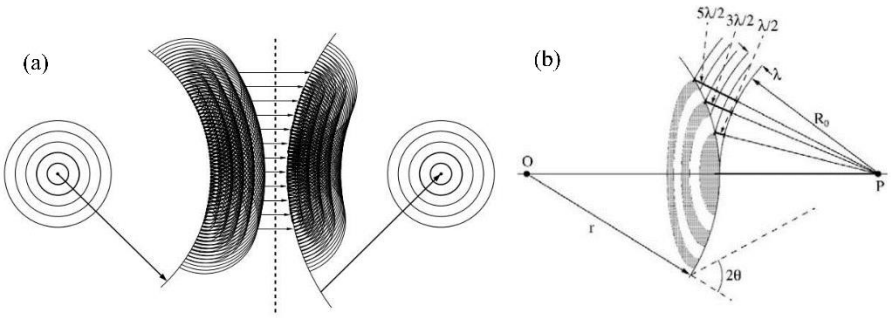


Figure 3.02. (a) The geometry of Huygens principle for diverging (left) and converging (right) wavefronts. The action of the lens is along the dashed line. (b) Construction of Fresnel zones considers the self-interference of the spherical wave possessing a range of momentum vectors while converging to a point P. Copyright (2012) by Springer Science and Business Media [19].

The Fresnel diffraction pattern for an edge can be calculated using the following equation.

$$\psi_{SC}(P) = \frac{i\psi_{in}^{0} e^{ik(r_0+R_0)}}{2(r_0+R_0)} [C(X) + iS(X)]_{X_0}^{\infty} [C(Y) + iS(Y)]_{-\infty}^{\infty} \quad (3.04)$$

Where *C(X)* and *S(X)* are the Fresnel cosine and sine integrals and the plot of $C(X) + iS(X)$ is called a 'Cornu spiral'.

The phase change in the propagating spherical wave is due to a range of angular momentum vectors associated with a spherical wave (for plane wave, only one momentum direction) and the self-interference while converging. Various angular momentum vectors will acquire path differences between themselves due to the outside curvature of the spherical surface with respect to the converging point at ***P,*** and the problem of interference is solved by well-known Fresnel zone construction. More precisely, the wavefronts geometries are spherical, parabolic, and plane surface for Rayleigh-Sommerfeld, Fresnel, and Fraunhofer regimes. The rate of change of phase between wavevectors is governed by the outward curvature of the wavefront, which is different for three different regimes. A similar picture is also captured in self-interference between the probe and scattered waved, describing the HRTEM image intensity pattern [6,20]. The momentum vector direction of ψ_0 is along the k_0 direction, and ψ_i along various k directions. Kindly note that the interference geometry is different for Fraunhofer pattern and off-axis electron holography (see sec. 3.2.2 and 3.4).

3.2.2 Fraunhofer diffraction pattern

The plane wavefront from a pair of spatial points at the slit opening ensures the intensity distribution along different scattering angles. The Fraunhofer integral can be employed for various geometry of slits to calculate intensity patterns in the diffraction plane as a function of scattering angle based on analytical expressions. As an example, the Fraunhofer integral for a rectangular aperture of sides 2a and 2b with the origin at the center O of the rectangle and with Qx and Qy axis parallel to the sides is given by,

$$U(P) = C\int_{-a}^{a}\int_{-b}^{b} e^{-ik(p\xi+q\eta)}d\xi d\eta = C\int_{-a}^{a} e^{-ikp\xi}d\xi \int_{-b}^{b} e^{-ikq\eta}\,d\eta \qquad (3.05)$$

The intensity is represented as

$$I(P) = |U(P)|^2 = \left(\frac{\sin kpa}{kpa}\right)^2\left(\frac{\sin kpb}{kpb}\right)^2 I_0 \qquad (3.06)$$

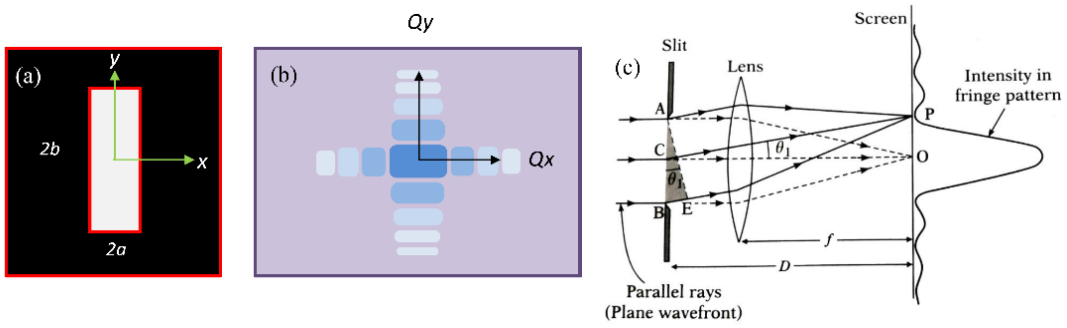


Figure 3.03. (a) The geometry of rectangular aperture and (b) corresponding far-field or Fraunhofer intensity pattern. (c) The geometry of Fraunhofer interference showing the origin of path difference.

Experimental observation of example Fresnel pattern and far-field Fraunhofer pattern from TEM is given in Figure 3.04. As already mentioned, the interference geometry for the far-field pattern is different than the Fresnel zone construction. The correlation of path difference between emerging waves from various spatial points at the aperture plane is considered in this case. The path difference is not between different momentum vectors but rather a spatial separation between two points with the same momentum or scattering. In off-axis electron holography, the interfering waves have momentum direction mirror symmetry to each other (sec. 3.4.3).

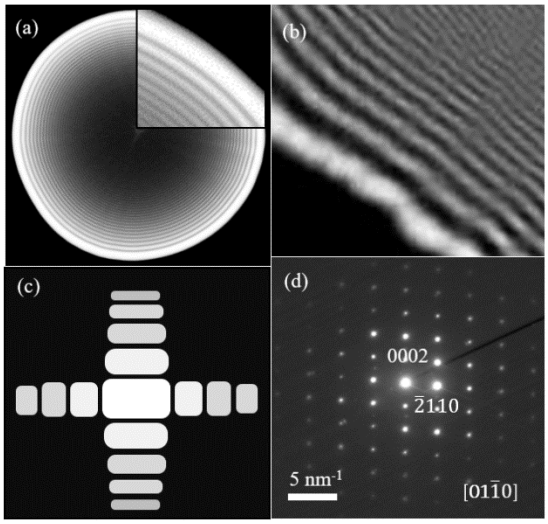


Figure 3.04. Fresnel diffraction pattern as observed under slight defocus (~ 10s of nm) condition from the (a) edge of an aperture and (b) thin specimen edge. (c) Far-field Fraunhofer pattern from the single slit and (d) electron diffraction pattern of ZnO Crystal along $[01\bar{1}0]$ Z.A.

3.2.3 Fourier transformation-based method

In the preceding two sections, the propagation of phase information from the aperture plane to the near field regime (Fresnel) diffraction plane (Fraunhofer) is briefly described, emphasizing specific diffraction geometry of waves depending on the momentum vectors. The change in momentum vector directions of the scattered/emergent waves and concomitant interference phenomena are at the heart of calculating the intensity pattern.

On the other hand, the Fourier transformation (FT)-method is based on the principle that the intensity pattern corresponding to image and diffraction planes is related by FT without considering diffraction geometry. For example, a periodic crystal-oriented diffraction pattern along high symmetry direction is the Fourier transform of the periodic crystal potential. However, the independent variables are different. In the case of FT, the independent variable is the frequency of wave vector *k*, and for Fraunhofer's analytical formula, it is scattering angle θ. The abs-FT pattern needs to be calibrated with respect to θ or reciprocal lattice vector g. The Fourier method's mathematic operation is similar to Abbe's imaging theory. More details on Abbe's imaging theory and its connection with off-axis electron holography formalism are given in sec. 3.2.4.

Now, according to the Fourier method, the far-field Fraunhofer pattern is obtained by Fourier transformation of the object function $f(x)$ (or $f(x,y)$ in 2D) as

$$\mathcal{F}(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x.k} dx \tag{3.07}$$

Strictly speaking, it is the magnitude of $\mathcal{F}(k)$ or absolute FT, which results in the Fraunhofer pattern at far-field and not the individual real $(\boldsymbol{Re}(\boldsymbol{x},\boldsymbol{k}))$ and imaginary $(\boldsymbol{Im}(\boldsymbol{x},\boldsymbol{k}))$ components.

$$\boldsymbol{Abs\ (or\ modulus\ or\ magnitude)\ \mathcal{F}(k) = Fraunhofer\ pattern} \tag{3.08}$$

However, it is indispensable to know the $\boldsymbol{Re}(\boldsymbol{x},\boldsymbol{k})$ and $\boldsymbol{Im}(\boldsymbol{x},\boldsymbol{k})$ components to go back to the original function $f(x)$ through inverse Fourier transformation. The established phase retrieval procedures involve retrieving the object wave function (OEW) through the appropriate filter function applied directly on $I(k,z)$ derived from the image intensity $I(x,y,z)$ recorded at sufficient resolution [6,21,22]. Experimentally, only if the magnitude of $\mathcal{F}(k)$ is recorded in the diffraction plane, then $\boldsymbol{Re}(\boldsymbol{x},\boldsymbol{k})$ and $\boldsymbol{Im}(\boldsymbol{x},\boldsymbol{k})$ components are lost, and this is the well-known phase problem. Though it is not a problem if one derives $\mathcal{F}(k)$ or $I(k,z)$ from the recorded intensity pattern in the image plane with sufficient spatial resolution an. The phase related to object potential and crystallography is preserved and can be retrieved by following specific procedures.

To elaborate more, what happens in the FT process is that the object function at first is expanded into its continuous cosine and sine series with frequency ranging from $k = -n\ to\ +n$ as written below

$$Re(k) = \int_{-\infty}^{\infty} f(x) cos(2\pi x.k) dx \tag{3.09}$$

and

$$Im(k) = \int_{-\infty}^{\infty} f(x) sin(2\pi x.k) dx \tag{3.10}$$

$$Abs\ square\ of\ \mathcal{F}(k) = \left|\sqrt{Re^2 + Im^2}\right|^2 \tag{3.11}$$

And the phase, $\theta(k) = \tan^{-1}\frac{Im}{Re}$ (3.12)

The expansion of the object function (discrete or periodic) into various cosine and sine harmonics of different frequencies resulting in Fourier waves with various frequencies and are similar to Abbe waves (sec. 3.2.4). The absolute FT is then plotted for each frequency, and the result is the well-known Fraunhofer pattern. Moreover, each frequency of the Fourier wave is

associated with the corresponding phase. This phase as a function of *k* can be calculated [Figure 3.05(d)].

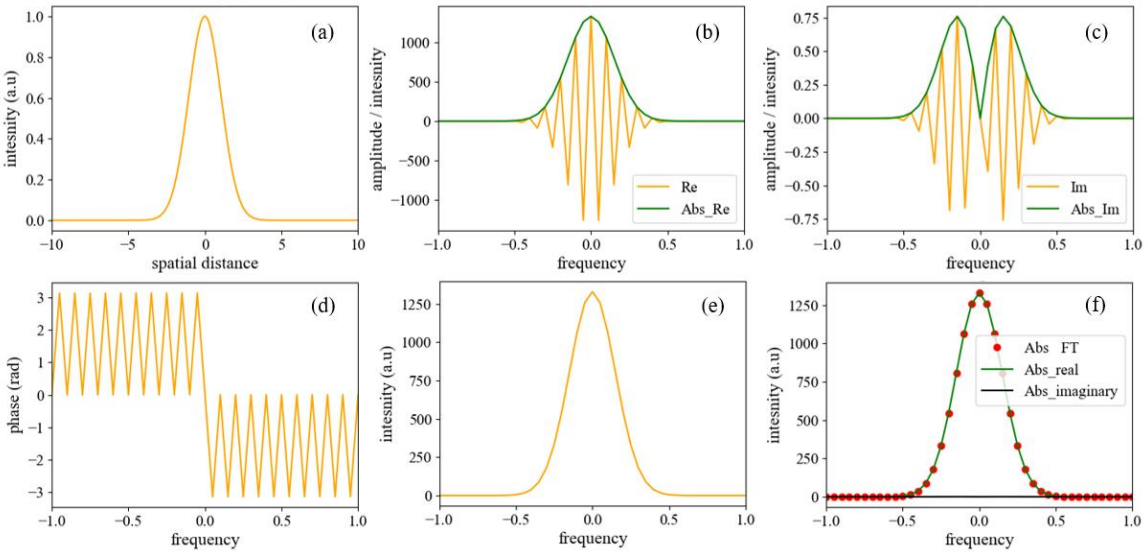


Figure 3.05. (a) Example of a Gaussian function $y = \exp(-\frac{x}{1.5})^2$ (b) and (c) corresponding **Re** and **Im** parts along with respective absolutes after Fourier transformation, respectively. (d) The phase of Fourier waves as a function of frequency, (e) Absolute of FT, and (f) total abs-FT, abs-**Re**, and abs-**Im** plotted together, showing that abs-**Im** is negligible abs-**Re** is almost equal to abs-FT.

The implication is that for a discrete function, e.g., if the function is a narrow slit, it will require a higher magnitude of frequency to get a first zero in the frequency pattern compared to a broader slit. The subsidiary effect is due to the non-perfect cancellation of integrand depending on frequency. For a periodic function or slit, a similar explanation holds. In the Fourier-based method, one adjustment is required to calibrate or relate various frequencies with the scattering angle, which does not arise naturally. This can be achieved by associating the frequencies at which Abs FT magnitudes are zeroes with the angle at which destructive interference occurs. The latter aspect is the physical picture required to derive the analytical formula based on the Fraunhofer integral (sec 3.2.2). The results calculated independently by FT and the analytical approach are shown in Figure 3.06. for seven periodic slits and have a few differences.

The analytical formula for various slit geometry is derived based on the path difference of waves between two spatial points at the slit opening emanating along different directions based on constructive and destructive interference. The method has the advantage over the Fourier-based

method because the formula contains useful quantities, e.g., wavelength (λ) of the illumination, scattering angle (θ), dimension, and periodicity of the slit.

For diffraction from N number of periodic slits, the intensity expression is given by

$$I(\theta) = I_0 \left[\frac{\sin\left(\frac{\pi a}{\lambda} sin\theta\right)}{\frac{\pi a}{\lambda} sin\theta}\right]^2 \left[\frac{\sin\left(\frac{N\pi d}{\lambda} sin\theta\right)}{\sin\left(\frac{\pi d}{\lambda} sin\theta\right)}\right]^2 \quad (3.13)$$

Where a is the slit width, d is the inter-slit distance or periodicity. λ is the wavelength of the monochromatic wave.

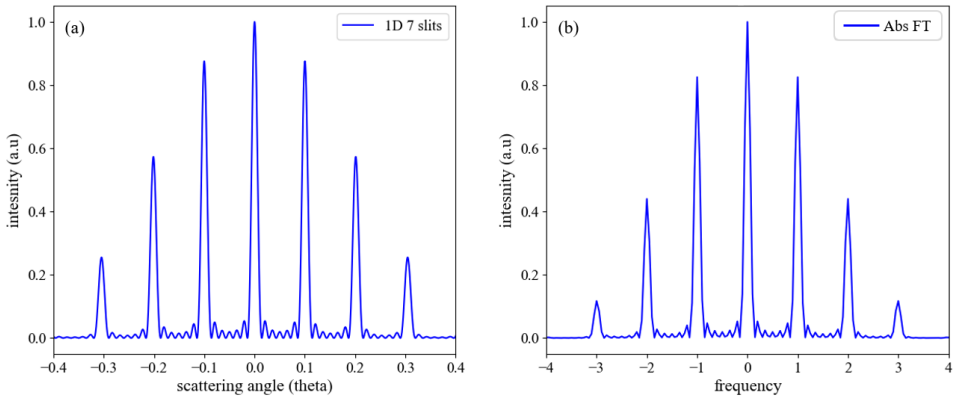


Figure 3.06. (a) Comparison of Fraunhofer pattern calculated by Fraunhofer analytical method for periodic seven slits with slit size a = 2λ and periodicity d =10 λ and (b) Fourier transformation of seven periodic slits followed by abs-FT. (b) the appearance of zero magnitude in abs FT is due to zero values of integration of $\mathrm{Re}(x,k)$ and $\mathrm{Im}(x,k)$ parts with respect to x corresponding to respective frequency values.

Fourier transform physically represents that each object will give off $\boldsymbol{Re}$ and $\boldsymbol{Im}$ parts of the Fourier wave with frequencies $k = -n \ to \ +n$. For each frequency, one needs to evaluate the integrand corresponding to $\boldsymbol{Re}$ and $\boldsymbol{Im}$ part and sum up to obtain the abs-FT for that frequency. Therefore, in assessing the $\boldsymbol{Re}$ and $\boldsymbol{Im}$ part, the amplitude and phase of each Fourier wave, i.e., the numbers outside and inside the cosine and sine part, mixed up. Or, in other words, they collapse to a single number. Moreover, the amplitude of each cosine and sine part will depend on the values of corresponding cosine and sine multiplied with $f(x)$. This is how mathematical procedures mix up the information on the function and frequency of waves. In HRTEM and off-axis holography image reconstruction, the starting data is an image. Therefore, every step while performing FT is preserved, and then only it is possible to go back to the image again.

In the Fourier transformation (FT)-based approach, the far-field image of the object is the modulus of the FT (abs-FT) of the object function. However, the phase angle calculated for Fourier waves or equivalently Abbe waves corresponding to each frequency does not contain the scattering angle. Thus, the frequency of Fourier waves needs to be calibrated either with respect to the scattering angle obtained through an analytical method or an experiment using a standard sample with known lattice parameters [Figure 3.06].

3.2.4 Abbe's theory of imaging

According to Abbe's theory, it is postulated that the point of interaction between illuminating monochromatic wave and object generate waves with a continuous spectrum of frequencies. The higher the frequency of the wave higher is the scattering angle. The waves of equal frequency and converging angle meet (by ideal microscope lens) to produce standing waves corresponding to different discrete frequencies. These standing waves superpose to form an image wave function. The largest frequency component allowed by the aperture defines the resolution according to Abbe's famous resolution formula. These standing waves are precisely similar to the Fourier transformed waves discussed above, except object function being considered and form the basis of Fourier optics. The diffraction geometry is given in Figure 3.07 and shows interference between pair of wave vectors having the same inclination angle. There is a continuous presence of such pairs of partial waves with a continuous range of inclination angles. This is similar to off-axis electron holography only for a pair of momentum vectors across a mirror plane.

Now, the resolution criterion is derived as follows. The objective lens acts as a low-pass filter. The cut-off frequency in 1D is given by

$$k_{max} = \frac{2\pi}{\lambda f_{obj}} x_{max} = \frac{2\pi}{\lambda f_{obj}} \theta_{max} \tag{3.14}$$

The effect of limiting frequencies higher than the maximum allowed by the aperture is to determine the resolution of the system.

The Fourier transformation of objective aperture gives the impulse response or *psf* due to limiting objective aperture to the image plane in 2D.

$$g(r) = \frac{J_1(k_{max} r)}{k_{max} r} \tag{3.15}$$

Where, $r = x^2 + y^2$. Rayleigh's criteria for resolution are postulated in terms of a minimum distance between two such blurred points due to aperture. This is defined by the first root of two such functions, i.e., the maximum of one function overlaps with the root of the second function. This occurs at $x_o = 1.22\pi$. The resolution is given by

$$\rho_0 = 0.61\frac{\lambda}{NA} \tag{3.16}$$

Where, $NA = \frac{r_{max}}{f_{obj}}$ is the numerical aperture, which represents the maximum half-angle subtended by the entrance pupil.

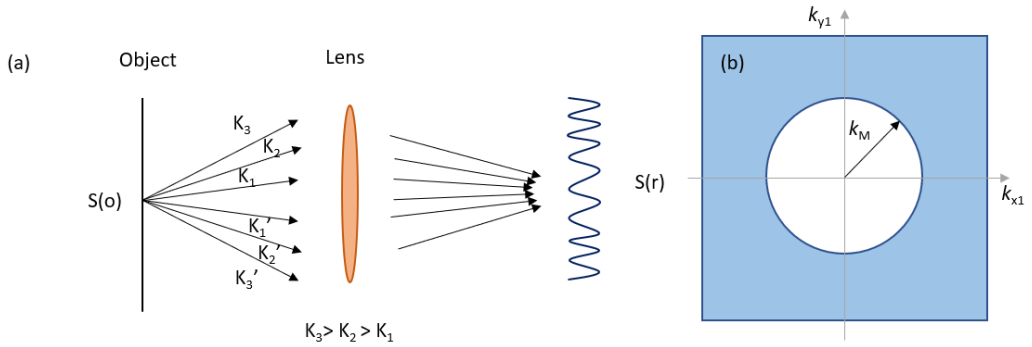


Figure 3.07. (a) Abbe's picture of image formation shows a range of frequencies generated at the point of interaction between illumination wave and object point, (b) Maximum frequency allowed by aperture will define the resolution of the image.

Now in the context of diffraction from atomic potential, the Fraunhofer pattern is the atom scattering amplitude $f(k)$ and the structure factor $F_g(k)$ for isolated and periodic atoms, respectively. Diffraction depends on the strength and crystallography of the scattering potential. More details about the atom scattering factor are provided in sec. 3.3.2.

3.3 Existing methods of image simulation

There are different methods of image simulation, mainly based on two approaches,

1. Zernike phase object and WPOA
2. Atom scattering factor

These methods of image simulations are discussed in detail and compared for the results they yield. A comparative study is carried out for theoretical simulation methods and the experimental data.

3.3.1 Image simulation based on Zernike phase object and WPOA

The weak phase object approximation (WPOA) originates in the Zernike phase contrast theory, which considers the specimen as a pure phase or weak scattering object. The object transmission function is represented by a complex function for small phase change ϕ, where ϕ is a real function corresponding to the discrete or periodic transparent object. This is known as weak phase object approximation (WPOA) [1].

The definition of a transmission function can be found in Ref. [23] (pp 446-447). Transmission function can be characterized as any diffraction grating that introduces variation in amplitude and phase of the incident wave. It is given by

$$F(\xi,\eta) = \frac{V(\xi,\eta)}{V_0(\xi,\eta)} \tag{3.17}$$

Where, $V_0(\xi,\eta)$ is the disturbance on (ξ,η) plane in the absence of object and $V(\xi,\eta)$ is the disturbance on the same plane when the object is present. This is consistent with the formalism of transmission function for electron imaging discussed before.

Zernike's phase contrast method section of Ref. [1] (page 472), the 'phase object' is defined by a complex amplitude function (for light) as follows

$$F(x) = e^{i\phi(x)} \tag{3.18}$$

Where $\phi(x)$ is a real periodic function and whose periodicity is equal to the grating period (in the case of periodic grating). $\phi(x)$ is small compared to unity, and the above equation can be approximated to

$$F(x) \sim 1 + i\phi(x) \tag{3.19}$$

Now one can notice the origin of observing $\phi(x)$ in the intensity information. If this approximation is not made, $\phi(x)$ would be lost.

WPOA is a straightforward and widely applied approach to simulate the HRTEM images of thin samples. In HRTEM, WPOA describes the phase shift of probe electron wave due to the electrostatic potential $V_t(x,y)$ of the specimen projected along the beam propagation direction, and the transmission function has the following expression.

$$t(x) = \exp\left[i\sigma V_t(x,y)\right] \tag{3.20}$$

After invoking WPOA and ignoring the terms with σ^2 and higher-order,

$$t(x) = 1 - i\sigma V_t(x, y) \tag{3.21}$$

Where, $V_t(x, y)$ is the projected specimen potential and $\sigma = \frac{2\pi me\lambda}{h^2}$ is the interaction constant. m and e are the mass and charge of the electron, and h is the Plank's constant. The approximation in Eq. 3.21 is essential to retain the object information in the image plane. The accompanied transmitted wave function is derived considering the change in the magnitude of electron wavelength from λ to λ' due to attractive positive potential (V_s) of the specimen. Refraction of the electron through atomic potential is associated with the change in momentum vector direction and is addressed in the description of the alternative method (sec. 3.4). λ' an average change corresponds to a mean inner potential (MIP) for a given spatial extent or a function of spatial position from the atom's center at medium and atomic resolution. At medium resolution, for average projected potential V_t, the transmitted wave function of the electron within kinematical scattering in 1D is given by

$$\psi_t(x) \sim t(x) \exp(2\pi i k_z z) \tag{3.22}$$

Incorporating the transmission function in the above equation, the transmitted wave function is given by,

$$\psi_t(x) \sim \exp(2\pi i k_z z) \exp(i\sigma V_t) \tag{3.23}$$

The above equation can be modified to

$$\psi_t(x) \sim \exp\big(i(2\pi k_z z + \sigma v_z(x)\big) \tag{3.24}$$

The plane wave component in Eq. 3.22 contributes to the background as the direct (DC) component and poses difficulty in in-line holography along with twin image components. This similarity is the basis for terming HRTEM images as in-line holograms. However, the above description of WPOA does not draw an analogy between Gabor's in-line holography, Fresnel diffraction geometry, and defocus HRTEM image in terms of interference geometry. Considering the change in momentum vector direction due to interaction with the object potential draws analogies and differences between all the three pictures regarding interference geometry.

Now considering lens response, the image wave function becomes

$$\psi_i(x,y) = 1 - i\sigma\phi_p(-x,y) * \mathcal{F}\{P(u,v)\exp\big(i\chi(u,v)\big)\} \tag{3.25}$$

And the image intensity is given by [4,5]

$$I(x,y) = \psi_i(x,y)\psi_i^*(x,y) \approx 1 + 2\sigma\phi_p(-x,y) * \mathcal{F}\{sin\chi(u,v)P(u,v)\} \tag{3.26}$$

Where $$\sigma\,\mathcal{F}\{sin\chi(u,v)P(u,v)\} = \frac{2\pi}{\lambda^2}\int_0^{\theta_{ap}} sin\chi(\theta)J_0(\frac{2\pi\theta r}{\lambda})\theta d\theta \tag{3.27}$$

is the lens response function. Suppose we don't make any such series approximation. In that case, phase information will be inside the cosine and sine trigonometric functions. The intensity of transmitted radiation derived by multiplying the transmitted wave function with its complex conjugate will result in a constant value, and thus phase information is lost. Therefore, by expanding the transmission function in a series (weak phase object approximation), the phase is retained, even after multiplying with the complex conjugate. Thus, we can say that by mere mathematical manipulation, information on the object phase is preserved.

If we don't consider the series expansion and approximation, then the image wave function can be written as

$$\psi_i(x,y) = \{t(x)\exp(2\pi ik_z z)\} * \mathcal{F}\{P(u,v)\exp\big(i\chi(u,v)\big)\} \tag{3.28}$$

Image intensity

$$I(x,y) = 1 \tag{3.29}$$

Now, let $$p = 2\pi k_z z + \sigma v_z(x) \tag{3.30}$$

$$\psi_i(x,y) = \{\cos(p) + isin(p)\} * \{\cos(x) + isin(x)\} \tag{3.31}$$

$$I(x,y) = \psi_i(x,y)\psi_i^*(x,y) \approx 1 \tag{3.32}$$

However, there is a way out of the series approximation or WPOA. To read this additional phase shift, we need to have a reference wave so that the fringe shift will be visible and phase change can be measured, as done in Ref. [6]. This is based on the physical picture of self-interference. The intensity expression looks similar to the expression corresponding to off-axis electron holography except for the additional carrier frequency and writing component of Gabor's holography [20]. A hologram is typically an interference pattern formed between incident plane wave illumination and scattered or diffracted waves due to object and is a two-step process, writing

and reading [Figure 3.08 (a)]. The intensity distribution at the near field of the object at some propagation distance, typically in the Fresnel regime, can be captured near the Gaussian image plane through a CCD camera. In other ways, the experiment hologram contains the information on how the object has diffracted the electron beam and the information at a propagation distance. As the hologram records the complete wavefield information, including the effect of coherent aberration, it can remove aberration from the image to represent the object through subsequent digital ex-situ procedures.

The amplitude of interference pattern which constitutes the ***writing component*** of the hologram is given by

$$A = \sqrt{UU^*} = \sqrt{A^{(i)2} + A^{(s)2} + 2A^{(i)}A^{(s)}\cos(\psi_s - \psi_i)} \tag{3.33}$$

Where, $U = Ae^{i\psi}$ is the complex disturbance at a point in the screen *H* at some distance behind the object, *A* is the amplitude, ψ is the phase, superscript, and subscript corresponding to *(i)* and *(s)* denote incident and scattered waves, respectively.

The ***reading component*** or reconstruction of the hologram is written in the multiplication of object transmission function with the plane wave illumination. Together, they form the transmitted wave function. Thus, the reconstructed wave or substituted wave is written as follows [Figure 3.08 (b)]

$$U' = \alpha_p U^{(i)} = KA^{(i)2}e^{i\psi_i}[A^{(i)} + \frac{A^{(s)2}}{A^{(i)}} + A^{(s)}e^{i(\psi_s - \psi_i)} + A^{(s)}e^{-i(\psi_s - \psi_i)}] \tag{3.34}$$

Where α_p is the amplitude transmission factor of the positive hologram. $U^{(i)} = A^{(i)}e^{i\psi_i}$ denotes the incident wave or the coherent background, and *K* is the proportionality constant [1].

The transmitted wave function in HRTEM within weak phase object approximation (WPOA) has the form [9] similar to the ***reading component*** of the optical hologram. However, the description of transmission function in terms of various parameters is different.

HRTEM intensity can also be written as [6],

$$I_{in\ line} = |\psi_0 + \psi_i|^2 = A_0^2 + A_i^2 + 2A_0A_i cos(\phi_i - \phi_0) \tag{3.35}$$

This is similar to the ***writing component*** of the hologram. Compare between Eq. 3.33 and 3.35, here ψ_0 and ψ_i are equivalent to ψ_i and ψ_s of Eq. 3.35

The information about the object is carried by the transmission function component of the transmitted wavefunction. The formation of an image from the object using an appropriate lens is equivalent to the recorded hologram with a supplementary illumination with the lens [Figure 3.08 (c)]. However, the mathematical expressions corresponding to ***reading*** and ***writing*** are not the same, and retaining the phase information in terms of intensity modulation is different in the context of HRTEM.

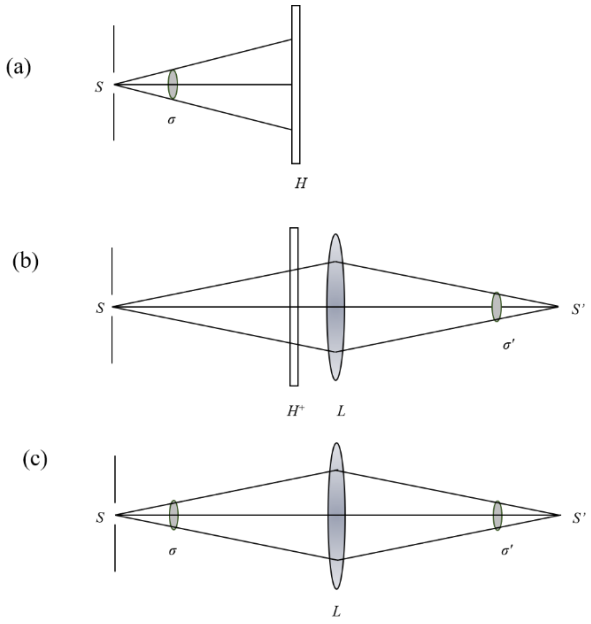


Figure 3.08. (a) and (b) show the writing of the hologram and object image reconstruction from the recorded hologram, respectively. (c) Equivalent image formation without a hologram. A hologram is utilized to remove the coherent lens aberrations by ex-situ digital processes to render the object's image faithfully. Copyright (1999) by Cambridge University Press [1].

Now, Figure 3.09 shows the image intensity calculated using Eq. 32.26 (based on WPOA) with and without considering the lens response for isolated Mo, S, N, and B atoms [7]. Not considering lens response is similar to Zernike like phase transfer, and as the potential function is asymptotic, peak value will remain undefined with a background value of 1 [Figure 3.09 (b)]. Considering aberration through optimum PCTF ($C_S = -35\,\mu m$ and $\Delta f = 8\,nm$), peak values

(and FWHM) of ~ 22000 (0.25 Å) and 3500 (0.25 Å) are obtained for Mo and B atoms, respectively [Figure 3.09 (a)]. This gives a peak intensity ratio of ~ 6.2. Peak intensity increases linearly with the atomic number. The trend contrasts with experimental observation, where changes in peak intensity in the first decimal place are observed with the atomic number [Table 3.01]. According to Eq. 3.26, the high peak value is due to the convolution procedure and cannot be normalized individually as the image without PCTF is not known [Figure 3.10].

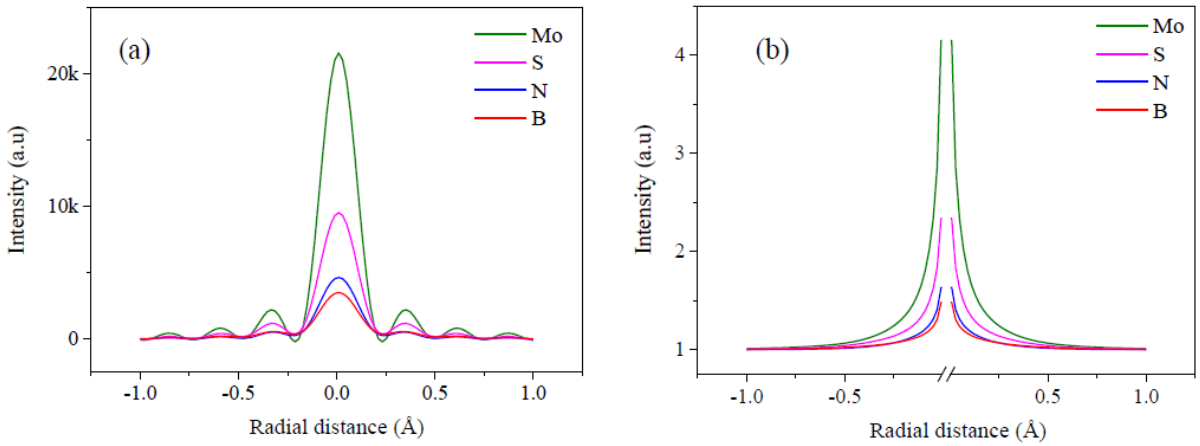


Figure 3.09. Intensity profile of Mo, S, N, and B atoms (a) using Eq. 3.26 with Cs =-35µm, Δf =8 nm, (b) with Zernike like phase transfer, i.e., $I(x,y) \approx 1 + 2\sigma\phi_p(-x,y)$.

Defocus (nm)	**Peak Intensity Mo (exp±0.04)**	**Peak Intensity B (exp±0.02)**	**Peak Intensity Mo (sim)**	**Peak Intensity B (sim)**
+1	1.08	1.04	1.03	1.02
+4	1.23	1.05	1.19	1.09
+8	1.24	1.05	1.25	1.11

Table 3.01. Peak intensity after considering aberration blurring for $C_s = -35$ µm and $\Delta f = +1, +4,$ and $+8$ nm compared to the experimental image recorded under similar conditions.

The convolution operation of the transmitted wave function with the point spread function results in very unusual behavior of atom image with slight variation in lens parameter and aperture angle. This convolution operation changes the peak height of the resultant curve significantly. Therefore, one needs to normalize the resultant based on flux balance, i.e., area under the object function. After considering flux balance, the resultant curve drops in peak value and delocalizes (blurring) compared to the object function. Experimentally acquired images need to be deconvoluted with the known *psf* and compare the image with the simulated image. However, any balanced flux approach will be relative between atoms, and for absolute contrast, the present alternative method may be used.

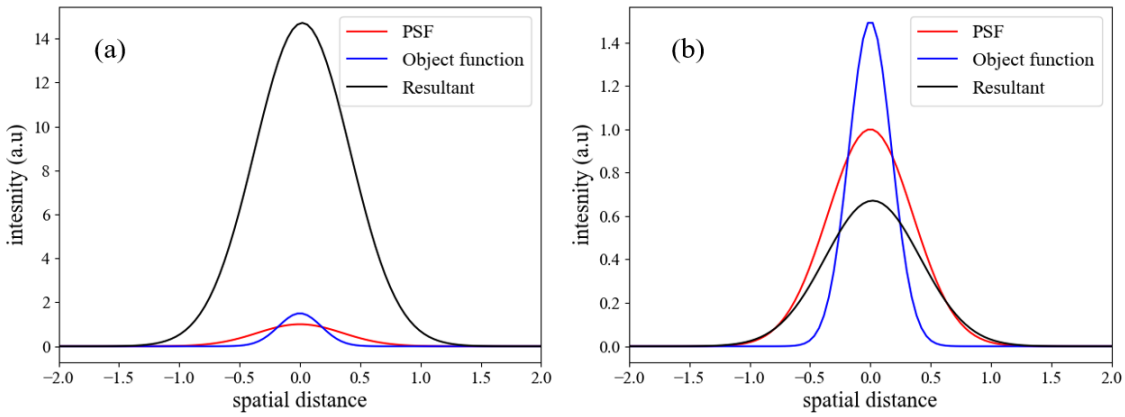


Figure 3.10. (a) Examples showing the area under the graph (resultant black) is not preserved after convolution between a model object function (Gaussian blue, with peak value 1.5 and FWHM 0.5) and convoluting function (Gaussian red, with peak value 1 and FWHM 1). (b) Resultant black is normalized based on the total area under the curve of the model object function (flux balance), which then shows the broadening at the expense of reduced peak intensity.

As already mentioned, the electron phase shift due to aberration cannot be added to the trigonometric function in the diffraction plane as that shifts the wave amplitude. Rather, it is used as a coherent envelope function or frequency filter in the diffraction plane and a point spread function (*psf*) in real space. According to Eq. 33.27, the PCTF gives weight to the magnitude of *psf* and aperture function in the Bessel function equivalent to Abbe's theory that sets the resolution in full width at half maximum (FWHM) in the final image [5]. Scherzer phase transfer will have maximum weight for the optimum value of spherical aberration and defocus that depends on the integration value of PCTF over the bandpass limits of spatial frequency. The weight of specimen potential in Eq. 3.27 has a similar effect as the atom scattering factor discussed in sec. 3.4. As the

convolution procedure changes the magnitude of the resultant function significantly [Figure 3.09 (a)], a flux balance approach is helpful to observe the qualitative decrease in intensity and increase in FWHM due to aberration compared to the ideal image free from any aberration [Figure 3.10]. The method described in sec. 3.4 can be used as a reference atom image without any lens response to employ quantitative flux balance after application of PCTF to compare the effects with a change in C_S, Δf and resolution.

The mean inner potential (MIP) extracted from the mean phase shift has been utilized at medium resolution. However, in a single atom at the sub-atomic resolution, the potential variation around the atom is essential and can be used for various other studies [7]. Moreover, the square amplitude of the transmitted wave function will be unity unless series approximation/WPOA is conjured. Also, no information on the phase can be obtained (Eq. 3.29 and 3.32) unless a reference wave is used in the mathematical expression according to in-line or off-axis holographic geometry [6,17]. However, the above description based on WPOA does not have any information on the geometry of interference in momentum vector directions. Instead, considering the change in momentum vector direction due to interaction with the object potential and ensuing interference effect, draw an exact comparison between the various pictures (sec. 3.2.2 and 3.4.3).

3.3.2 Image simulation based on scattering factor

Another approach of image simulation incorporates atom scattering factor directly instead of specimen potential. The transmitted wave function, in this case, can be derived from the Schrödinger integral equation and has the following form [9,19]

$$\psi_t(x) = \exp(2\pi i k_z z) + f_e(q)\frac{\exp(2\pi i q.r)}{r} \tag{3.36}$$

Where, $q = k - k_0$, and $f_e(q)$ is the atom scattering factor and is defined by,

$$f(q) = -\frac{m}{2\pi\hbar^2}\int V(r')e^{-2\pi i q.r}\, d^3r \tag{3.37}$$

which is the FT of the scattering potential. The above equation can be modified after implementing the projected potential and written as

$$f_e(k) = \frac{2\pi i}{\lambda}\int_0^\infty J_0(2\pi k r)\{1 - \exp[i\sigma\int V(x,y,z)dz]\}rdr \tag{3.38}$$

$f_e(k)$ is the electron scattering factor in the Moliere approximation, which has the advantage over the Born scattering factor due to an imaginary component. $\chi(k)$ is the aberration function, $k_{max} = \alpha_{max}/\lambda$ (rad Å^{-1}) is the maximum spatial frequency allowed by the objective aperture and $J_0(x)$ is the Bessel function of order zero.

The wave function solution based on the differential form of the Schrödinger equation is a plane wave. On the other hand, the integral form gives spherical waves and amplitude factor as the atom scattering factor. The equivalence between the two solutions can be perceived by enveloping all the spherical waves from many adjacent scattering centers, resulting in a plane wavefront. The picture is akin to Huygens's construction that a plane wavefront is the envelope of many forward scattered spherical wavelets and equivalent to First Born approximation. This result is used along with Moliere's scattering factor to calculate the image of isolated atoms by using the following equation. [9,16]

$$g(x) = \left|1 + 2\pi i \int_0^{k_{max}} f_e(k) \exp[-i\chi(k)] J_0(2\pi kr) k dk\right|^2 \tag{3.39}$$

The scattering factor based on Moliere approximation for isolated B and Mo atoms is plotted in Figure 3.11. (a) and (c). For a crystal with periodic lattice, the scattering factor is replaced by the structure factor. Example for BN and MoS_2 lattice along $[0001]$ Z.A. is shown in Figure 3.11 (b) and (d), respectively.

Eq.3.39 has some precursor work [15,16]. For example, in Scherzer's theory of phase contrast, the phase shift is introduced by the CTF function in the back focal plane and the object's wave function in the image plane at the Fraunhofer approximation. It is given by

$$\psi_i(r_i, v_i) = \frac{1}{\lambda_i f_i M} \int_0^{r_m} r dr \int_0^{2\pi} dv \times \exp\left(2\pi i \frac{r_i \lambda \cos(v - v_i)}{\lambda_i f_i M}\right) \psi_a(r, v) \tag{3.40}$$

$$\psi_i(r_i) = \frac{2\pi}{M\lambda} \int_0^{\alpha} \exp\left[i(s^2\Theta^4 - \tau s\Theta^2\right] \times S(\Theta) J_0\left(\frac{2\pi r_i}{M\lambda}\Theta\right)\Theta d\Theta \tag{3.41}$$

In the publication by Eisenhandler and Siegel [15], similar to the above formalism, the image wave function and corresponding intensity are given by

$$\psi(x^i, y^i) = \frac{2}{M\lambda} \int_0^{\pi} \int_0^{\alpha_{max}} \psi(\alpha, \theta) e^{i\chi(\alpha,\theta)} \times \cos\left(\frac{2\pi x^i cos\theta}{\lambda}\right) \cos\left(\frac{2\pi y^i sin\theta}{\lambda}\right) \alpha d\alpha d\theta \tag{3.42}$$

$$\left|\psi_{total}(x^i, y^i)\right|^2 \approx M^{-2} + 2M^{-1} Re[\psi(x^i, y^i)] \tag{3.43}$$

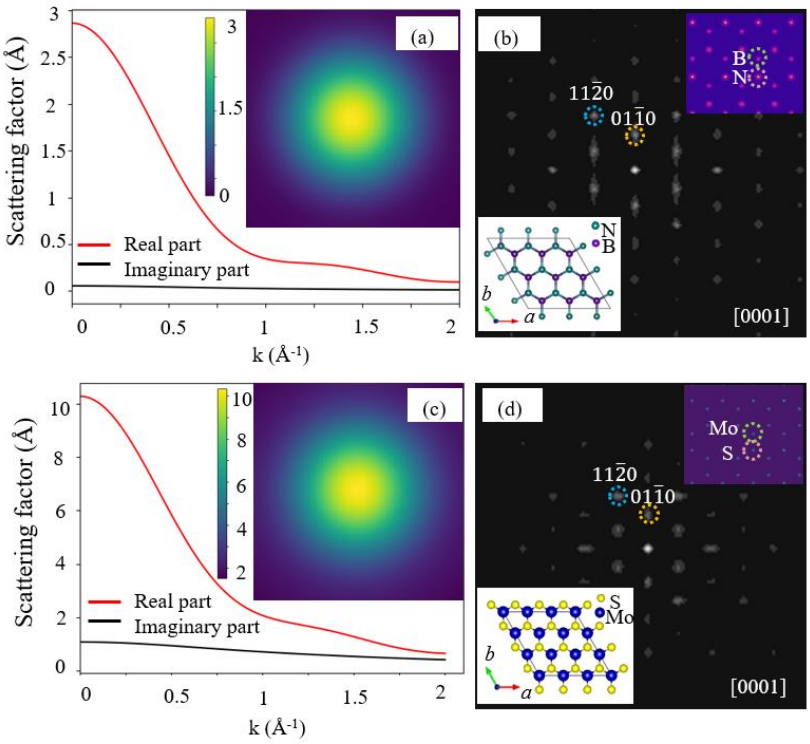


Figure 3.11. (a) Atom scattering factor of an isolated B atom in 1D (inset 2D distribution) and (b) structure factor of monolayer BN along [0001] direction. (c) and (d) scattering factor and structure factor for Mo and MoS_2 lattice along [0001] direction, respectively.

In the above formulations of Eq. 3.41 and 3.42 following considerations are made; Some of the illuminating monochromatic electron waves after passing through the sample scatter elastically. The scattered wave undergoes a phase change, and the amplitude of the scattered wave is a function of the scattering angle. The diffraction plane's intensity distribution is the Fraunhofer pattern obtained by the Fourier transform of object plane distribution. The amplitude distribution and phase relationship are essential to go back to the image plane by inverse Fourier transformation. It is considered that each point of diffraction pattern emits Huygens wavelets, and the wavelets which can escape the aperture at the back focal plane will interfere/recombine at the image plane to form the image. This can be obtained by inverse Fourier transformation, and any information lost is due to wavelets propagated outside the aperture. Eq. 3.41 and 3.42 are the final image wave function and intensity calculated for an aberrated lens. Figure 3.12. shows the example intensity pattern for four different isolated Mo, S, B, and N atoms considering PCTF and complete CTF with $Cs = 35$ µm, $\Delta f = 8$ nm using Eq. 3.39.

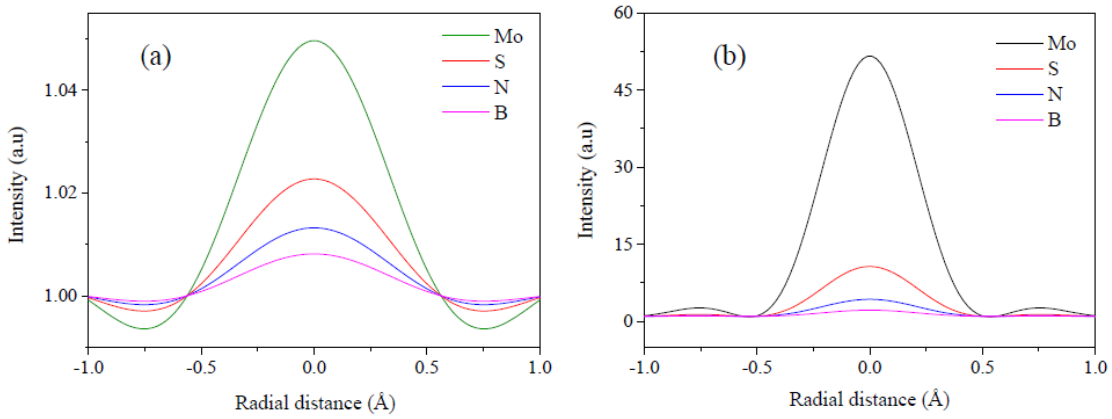


Figure 3.12. The intensity profile of isolated Mo, S, N, and B (a) calculated using Eq. 3.39 with $\mathrm{Cs} = -35$ μm and $\Delta\mathrm{f} = 8$ nm. Only a sine part of the PCTF is considered, which is valid for weak phase object approximation. (b) The effect of complete CTF on the intensity of Mo, S, N, and B.

The phase-contrast image intensity calculated using Eq. 3.39 varies weakly with atomic number and the peak phase shift $\varphi_{max}(rad)$ follows ~ $Z^{0.6} - Z^{0.7}$ where Z is the atomic number [9,23]. Though the trend can be complicated depending on the valence electron filling, specific atoms with higher *Z* can have smaller contrast than atoms with lower Z next to each other in the periodic table [9,17,23]. The peak intensity is almost the same irrespective of the atomic number and changes only in the second decimal place. Figure 3.13 and Table 3.02 summarized the peak intensity and FWHM values calculated using Eq. 3.26 (WPOA) and Eq. 3.39 (atom scattering factor) for Mo, S, B, N, Zn, and O atoms. One can notice that the difference in peak intensity and FWHM maximum calculated by two different methods are markedly different. The intensity values calculated based on Eq. 3.39 show frivolous dependence considering only PCTF irrespective of atom number. However, the peak values are much smaller, and FWHM are higher by a factor of two, respectively calculated by Eq. 3.39 and Eq. 3.26, and the difference will remain even after the flux balance.

Now, Eq. 34.27 describes the *psf* in the image plane. A similar expression derived by Scherzer, [16] Kirkland, [9] incorporates the atom scattering factor, and the transmitted wave function was derived based on Fraunhofer approximation and Schrödinger integral equation, respectively. To see how the *psf* acts on the image contrast, one needs to look at Eq. 3.26, 3.39, 3.14, and 3.18. The Bessel function term has the origin in an aperture function that sets the resolution criteria according to Abbe's theory. This term will control the image pattern width in

terms of FWHM. The sine function, atom scattering factor, and interaction constant σ will contribute to the weight to the image intensity, as the magnitude of phase contrast, atomic number, and characteristics of probe electron, respectively. The effect of atomic number enters into Eq. 3.20 (based on WPOA) as atomic potential. However, all these expressions are equivalent but give different image patterns depending on how the transfer function is considered. Kindly note that the aberration phase shift is never added as an additional phase inside the trigonometric operator of wave functions either in the image or diffraction plane. Instead, it acts as a frequency filter in the diffraction plane and *psf* through convolution in the image plane. If this phase is added, then it will cause a change in the pattern periodicity and is never observed.

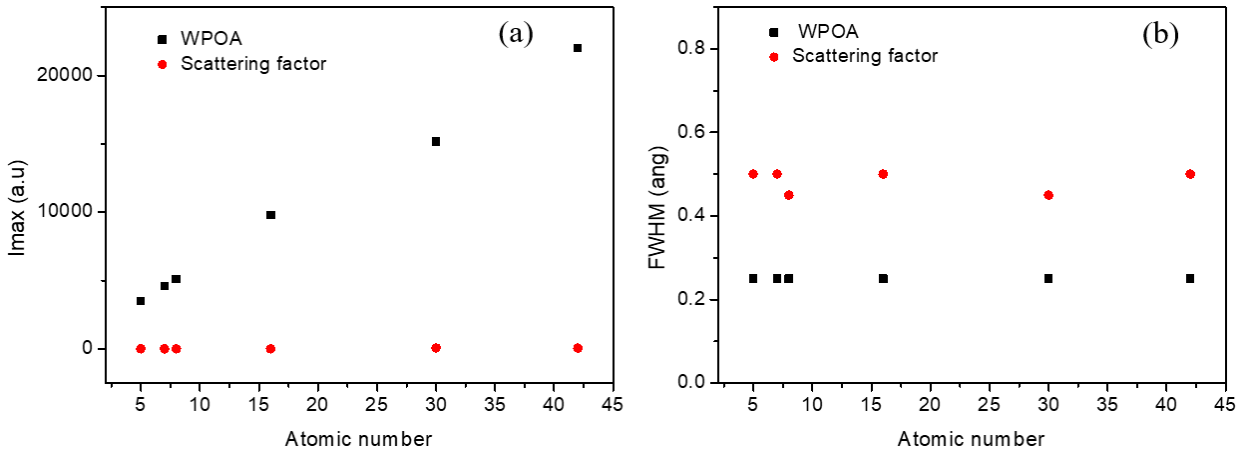


Figure 3.13. (a) The intensity and (b) FWHM values calculated using Eq. 3.26 (WPOA) and Eq.3.39 (Atom scattering factor) for Mo, S, B, N, Zn, and O atoms. Values are plotted for only sine part or PCTF function (Cs = -35 μm, Δf = 8 nm) considered for the intensity calculation. For complete ACTF and PCTF consideration based on Eq.3.39, see Table 3.02. The difference in peak intensity and FWHM maximum calculated by two different methods are markedly different.

Atom	Method		I_{max}	FWHM	Atom	Method		I_{max}	FWHM
Mo (42)	WPOA	LR*	22000	0.25	S (16)	WPOA	LR	9800	0.25
		WLR*	3.5	0.06			WLR	2.3	0.1
	Scattering factor	both sine &cosine	52	0.5		Scattering factor	both sine &cosine	11	0.5
		Only sine	1.05	0.75			Only sine	1.02	0.75
Zn (30)	WPOA	LR	15150	0.25	O (8)	WPOA	LR	5100	0.25
		WLR	3.23	0.1			WLR	1.7	0.1
	Scattering factor	both sine &cosine	70	0.45		Scattering factor	both sine &cosine	4	0.45
		Only sine	1.06	0.63			Only sine	1.01	0.63
B (5)	WPOA	LR	3500	0.25	N (7)	WPOA	LR	4600	0.25
		WLR	1.06	0.06			WLR	1.65	0.1
	Scattering factor	both sine &cosine	2.2	0.5		Scattering factor	both sine &cosine	4.4	0.5
		Only sine	1.008	0.75			Only sine	1.013	0.75

Table 3.02. Intensity and FWHM values calculated using Eq.3.26 (WPOA) and Eq.3.39 (Atom scattering factor) for Mo, S, B, N, Zn, and O atoms.

*LR: considering lens response and

*WLR: without lens response.

3.4 New method: Atom as an electrostatic charge center

An alternative method based on the atom as an electrostatic charge center and its action as an interferometer is described in this section. The atomic nucleus size is extremely small (~1.6-15 fm) compared to the overall size of the atom with electron clouds (~ 0.1-0.5 nm). Therefore, the nucleus can safely be considered as a source of a positive point charge (+Z*e*) with associated Coulomb potential that decays inversely away from the charge center [Figure 3.14 (a)]. The electron clouds surrounding the nucleus only screens the radially symmetric positive Coulomb potential.

Moreover, the surrounding electrons clouds scatter the probe electron inelastically [24]. The probability of inelastic events for the fast probe electron is negligible compared to the dominant elastic events for samples having a thickness less than the extinction length and at a short exposure time, typically 1-2 sec. This screened positive potential is attractive to the negatively charged probe electron wave while passing by the nucleus. The force experienced by the traversing electron will depend on the distance from the nucleus. The calculated screened projected potential of Mo atom following the Hartree-Fock model [9] and associated attractive force in terms of bending angle is shown in Figure 3.14 (a) and (b). Within this picture, the projected potential of an atom can be considered as an electrostatic circular prism similar to the experimental off-axis electron holography biprism except attractive force acting from all directions with varying strength as a function of the radial distance around it. This implies that the atom bends electron trajectory along $\theta = 0 - 2\pi$ azimuthal direction around it where momentum vectors lying on a conical surface representing scattering angles within the same order of magnitude. In contrast, a cylindrical biprism does the same but along single pair of momentum directions across a mirror plane [Figure 3.15].

Therefore, it is necessary to describe the interference pattern from a different geometrical perspective than typical unidirectional electrostatic biprism for the atomic case. The picture is on the classical concept of wave optics but differs from channeling quantum mechanical Bloch waves of electrons through the crystal lattice. In the following sub-sections, the method based on off-axis electron holography is presented briefly to introduce image contrast and then extend the principle to ‘the atom as electrostatic charge center’ to simulate the image of the atom.

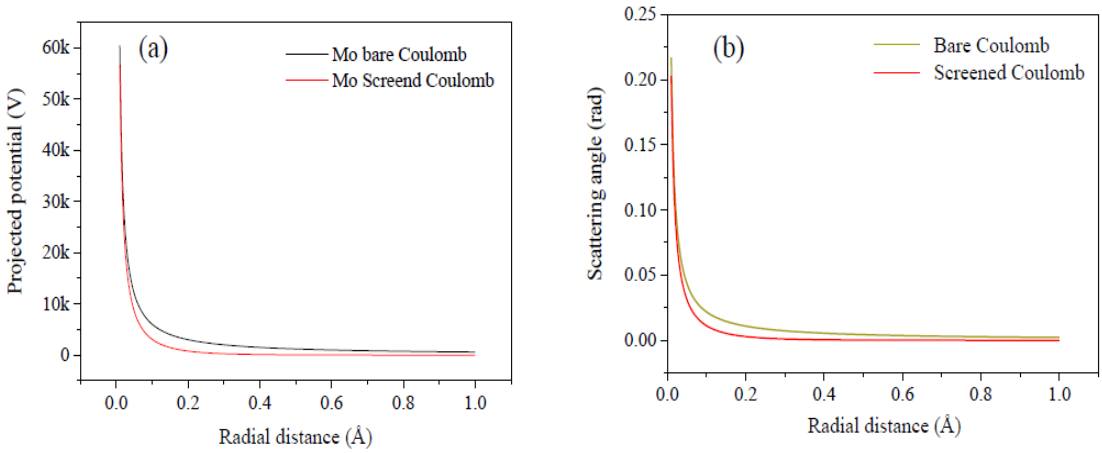


Figure 3.14. (a) The projected and bare Coulomb potential of Mo atom and corresponding (b) bending angle as a function of distance from the nucleus for both screened and bare Coulomb potential.

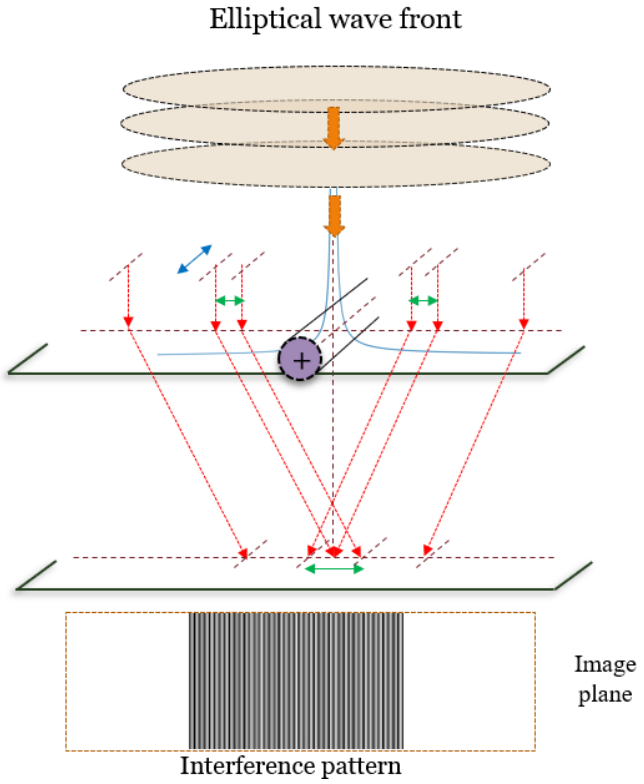


Figure 3.15. Image formation in off-axis electron holography with a single frequency wave with momentum vector direction equal and mirror symmetry with respect to the biprism axis.

3.4.1 Image intensity pattern in off-axis electron holography

Now there is another way to calculate patterns or images based on converging waves. This is based on the geometry of interference associated with off-axis optical/electron holography. The equation for each frequency in 2D is given by [17,18]

$$I_{hol}(\vec{r}) = I_0 + I_{ima}(\vec{r}) + 2|\mu|A_0 A(\vec{r})\cos\left(2\pi\overrightarrow{q_c}.\vec{r} + \phi(\vec{r})\right) \quad (3.44)$$

Where, $I_0 = A_0^2$ is the intensity of the reference wave. $I_{ima}(\vec{r})$ is the image intensity and $q_c = k\beta$ is the spatial carrier frequency of the hologram. μ is the degree of coherence.

There is a connection between Abbe's model and the carrier frequency associated with the off-axis electron holography. According to Abbe, waves intercepted by the aperture at a higher scattering angle are associated with higher frequency. In contrast, a higher angle of interference in off-axis geometry is associated with a higher carrier frequency. Both give finer details and are associated with better spatial resolution.

The new simulation approach proposed in the present chapter is similar to the off-axis electron holography concept by considering a single atom as an electrostatic interferometer. The principle is the extension of conventional off-axis electron holography. It considers an atom as an electrostatic charge center that induces interference of waves along complete azimuthal orientations. The interference geometry is different in this case than Fresnel and Fraunhofer's cases (see Figure 3.02 and 3.03 for comparison). In the off-axis electron holography, the pair of momentum vectors across the mirror plane produces an interference pattern.

3.4.2 Image contrast in off-axis electron holography with electron biprism

The formation of the electron interference pattern and resulting contrast in off-axis electron holography is emphasized here. The basic principle is based on single-electron wave interference, and the expression corresponding to HRTEM is given in Eq. 3.35 (sec. 3.3.1). The intensity pattern of the hologram is given by Eq. 3.45. Details on the off-axis electron holography methods and practices can be found in Ref. [17,18].

$$I(x,y) = I(x) = a_1^2 + a_2^2 + 2a_1a_2\cos\left(2\pi q_c x + \Delta\phi\right) \quad (3.45)$$

Where, $q_c = 2k_x$ is the carrier spatial frequency of the hologram. a_1 and a_2 are the amplitudes of

waves undergoing interference at an inclination angle due to the action of biprism. x is the spatial variable of the interference pattern, and $\Delta\phi$ is the difference in phase between the two interfering partial waves, typically acquired by one of the two waves due to object potential. $\Delta\phi = 0$ for the vacuum wave. Information on $\Delta\phi$ appears as a small deviation in hologram fringe. The maxima and minima of the intensity pattern can be determined from Eq. 3.45 and remain the same throughout the interference field. The carrier frequency q_c can be considered as carrying the information about the field strength and associated potential of the biprism. Field strength, potential, and carrier frequency can be empirically put together in a functional form. It is important to note that the total flux or energy of the interfering waves must be preserved on the resulting interference field. This is essential while calculating the image intensity and distinguishing the contrast between the various atoms based on the atom as a charge center.

The q_c of the hologram that represents the fringe distance depends on the angle of superposition for a given wavelength. The higher the angle of superposition, the larger will be the carrier frequency. Resulting in finer fringe spacing due to the horizontal component of the wave vector is larger at a higher inclination angle. This is equivalent to Abbe's picture, where the wave vector is larger or higher frequency at a higher scattering angle (sec. 3.2.4). Thus, the two concepts may be unified and said the higher the inclination angle, the smaller the fringe spacing or inter feature distance resulting in better spatial resolution. This was Abbe's hypothesis describing the diffraction-limited imaging, where restricting the higher frequency by numerical aperture coming at a higher scattering angle limits the spatial resolution.

3.4.3 Image contrast due to atom as a charge center equivalent to biprism

The image formation by the interference of waves due to atom charge center similar to 1D electrostatic biprism is described here. The intensity pattern can be calculated following Eq. 3.45, considering the radially symmetric atomic potential after incorporating wave interference effect from a given radial zone of extent Δr. The zones described here are similar to binary type Fresnel zone plates with multiple foci, depending on the scattering angle with a different order of magnitude. Calculating interference patterns along all azimuthal inclination angles for the peripheral zone area requires two additional considerations than the unidirectional interference pattern. The first consideration is that the wave flux will depend linearly on the perimeter $2\pi r$

which is a function of radial distance r from the center of the atom (Eq. 3.46). Larger the perimeter of the zone area [$\pi(r_2^2 - r_1^2)$], higher will be the flux of the wave approaching for the interference. The relative intensity contribution at the center of the pattern from different rims belongs to the same spatial coherent zone is scaled with $2\pi r$, where *r* is the radial distance from the center of the atom [Figure 3.16, Eq. 3.46].

$$I_{rad\ int}(r) = a_1^2 + a_2^2 + 2a_1a_2\cos(2\pi q_c r) * 2\pi(r_{max} - r) \tag{3.46}$$

The second consideration is that of flux balancing between the flux of wave at the plane of the atom as given by the coherent rim area $\pi(r_2^2 - r_1^2) * I$, where I is the intensity at a given pixel point within this zone, and resulting interference field over a circular area around the optic axis is given by $\pi r^2 \times I_{rad\ int}$ (1st law of thermodynamics).

$$\int_{-drr}^{drr} I(r)\, dr = \pi(r_2^2 - r_1^2) \times 1 \tag{3.47}$$

Where, $drr = (r_2 - r_1)/2$ and $r = \sqrt{x^2 + y^2}$. $I = 1$ is the minimum intensity count on a pixel of the size 1 pm considered for the present calculation.

The resulting interference pattern is different compared to the unidirectional interference geometry. The pattern is radially symmetric with peak intensity at the origin. Intensity decreases gradually, unlike the unidirectional pattern, but the average intensity remains the same throughout the field of superposition [Figure 3.16]. This is due to superior spatial coherency associated with the elliptical shape compared to the round shape [17]. However, the wave's spatial coherency can be correlated with electric field magnitude for a given spatial extent and equivalent order of bending angle due to Coulomb attraction.

For the elliptical spread of the beam, most of the spatial extent of the wave will approach from electric field regions away from the biprism center. It will have a bending angle almost within the same order of magnitude [Figure 3.17]. This is the region of high spatial coherency resulting in high contrast to regions close to the biprism where field strength and associated inclination angle varies strongly with the spatial distance. The analog with high energy electron diffraction at a small angle vs. large-angle scattering and their relationship with the spatial coherency is described. At a small angle, coherency in elastic scattering is well preserved. It gives rise to crystal Bragg diffraction peaks, whereas at larger angles, scattering becomes more and more incoherent, i.e., Rutherford type scattering. However, for the atom case, there is no advantage for shaping the probe

elliptical as potential falls much more rapidly due to neighboring atoms in a crystal and the short range of potential compared to biprism.

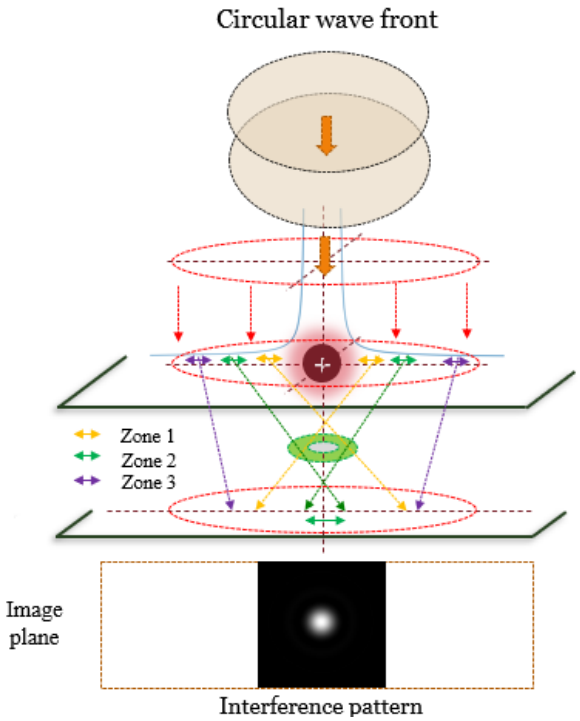


Figure 3.16. The interference pattern due to the atom as a charge center. Momentum vectors propagate along the radial direction in the case of an atom.

In the case of an atom, potential varies strongly close to the nucleus and slowly beyond a certain distance. Therefore, there will be a different contribution from various radial zones with specific rim widths. These zones are defined by the angle of superposition within the same order of magnitude. The overall pattern can be controlled by the lens focus [Figure 3.17]. The lens focus will effectively increase the width of the interference field for the coherent zone. Therefore, we have divided the entire spatial range surrounding the atom starting from 1 pm to a suitable outer range in various zones. For example, 1, 4, and 8 nm focus step (experimental through focus condition) and calculated intensity pattern for different zones separately and summing up the contributions [Figure 3.18 and 3.19]. The contributions to the overall intensity pattern from different zones are incoherent. They can be thought of as either binary Fresnel zone plate or annular Airy apertures giving rise to Airy pattern at the specific focal length. Each zone acts as a filter to specific momentum vector components similar to the PCTF as a spatial frequency filter.

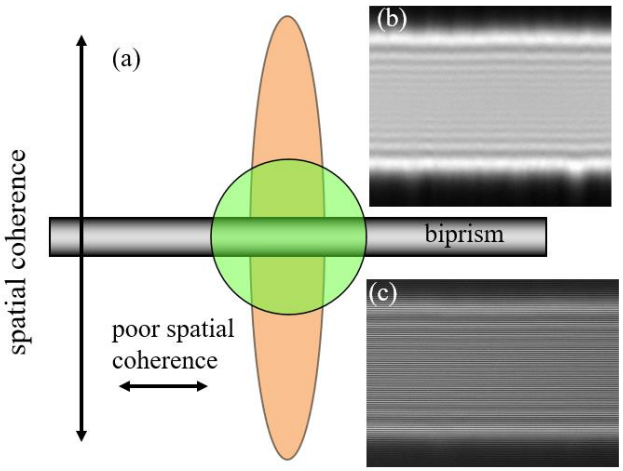


Figure 3.17. (a) Schematic showing relative spread of probe for conventional round and elliptical illumination and their connection with spatial coherency. (b) Fringe contrast corresponding to round and (c) highly astigmatic (elliptical) probe in off-axis electron holography.

The zone's spatial extent increases away from the atom due to decreasing potential and associated scattering angle. For example, the outer spatial extent is 10.31 and 4.19 pm corresponding to 1 nm focal length on the optic axis for Mo and B atoms, respectively. The extent of the outer zone contributing to the coherent interference field increases with increasing focus (Table 3.03). Various zone schemes and intensity values for different atoms are given in Tables 3.03 to 3.08. Depending on the focus settings, various zones will contribute to intensity pattern (focus) and background (away from focus) differently. For a given intermediate focus setting, the focus length smaller and larger than that will contribute to the background differently. The zones closer to the atom center would have formed interference patterns and propagated the information along the original direction and diffracted direction. Diffracted interference patterns will appear as a displaced image with intensity contribution minimal compared to the direct pattern. An intensity difference of almost an order of magnitude is observed. This is similar to the relative intensity between CB and SBs in the off-axis and direct and diffracted beams in HRTEM. Figure 3.19 shows various contributions from different zones of Mo atom for 1 nm.

Moreover, the flux contributions will be different from different zones and proportional to the annular zone area. As already mentioned, the area and corresponding flux will be higher for the zones away from the nucleus. The image of lighter atom B for similar focus settings is given

in Figure 3.20. Kindly note that the zones with large radial distances will contribute to and modify the image's peak value with increasing focus.

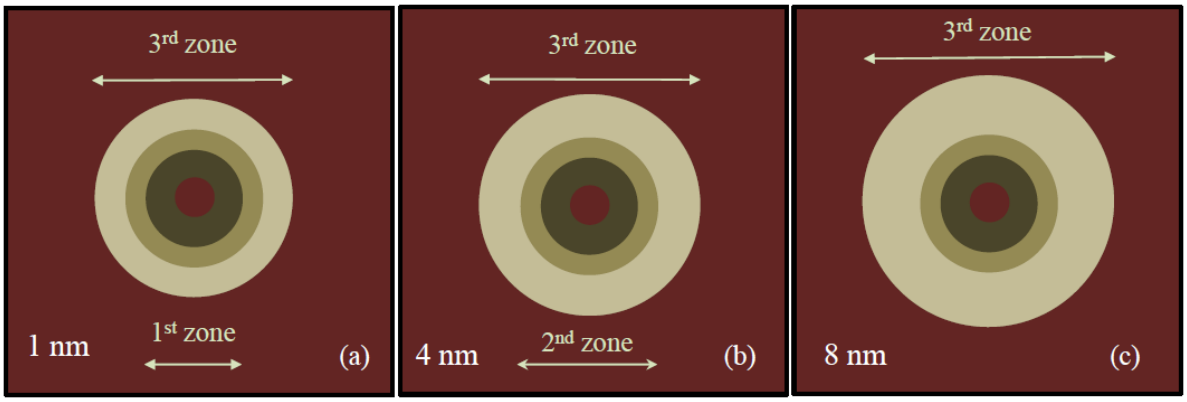


Figure 3.18. (a)-(c) Three different zones and their extent corresponding to 1, 4, and 8 nm focus step with respect to the optic axis for Mo atom, respectively. The extent of the third zone increases with defocus from 10.70 pm to 21.18 pm, having the same mean bending angle.

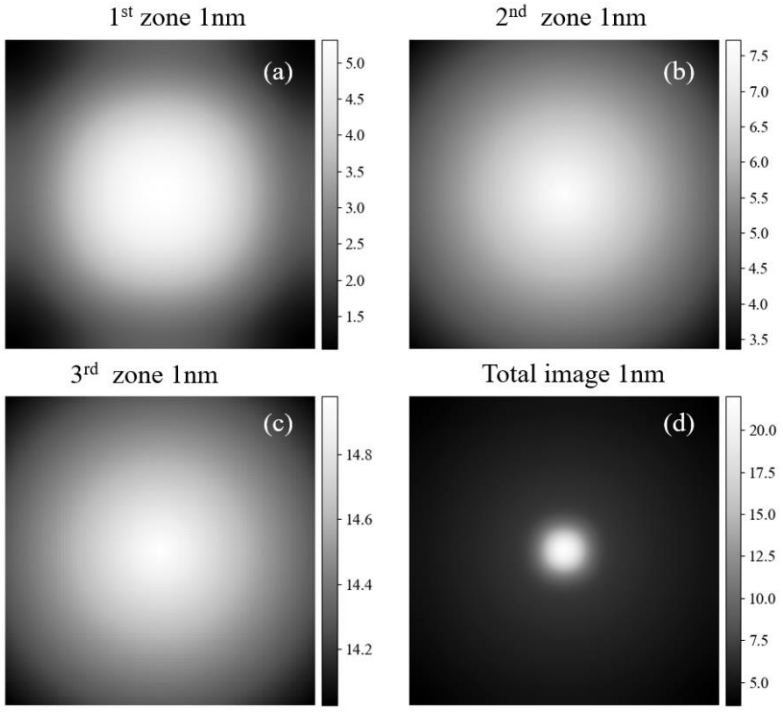


Figure 3.19. (a)-(c) Image pattern formation due to various zones at 1 nm focus and (d) overall image of the Mo atom at 1 nm focus.

Radial distance (pm) & zones	Rim width (pm)	Mean scattering angle (rad)	Length on optic axis (nm)	Intensity (with flux balance)
1-1.77	0.77	0.1558	1	5.82
1.77-10.31	8.54	0.0488	1	8.45
10.31-10.70	0.39	0.0099	1	16.89
10.31-16.91	6.60	0.0070	4	11.33
10.31-21.18	10.87	0.0062	8	9.56

Table 3.03. Various radial zones and corresponding focal length on the optic axis for Mo (Z=42) atom for 1, 4, and 8 nm length on the optic axis. Zones are divided based on scattering angles falling in the same order of magnitude. The corresponding schematic is given in Figure 3.18.

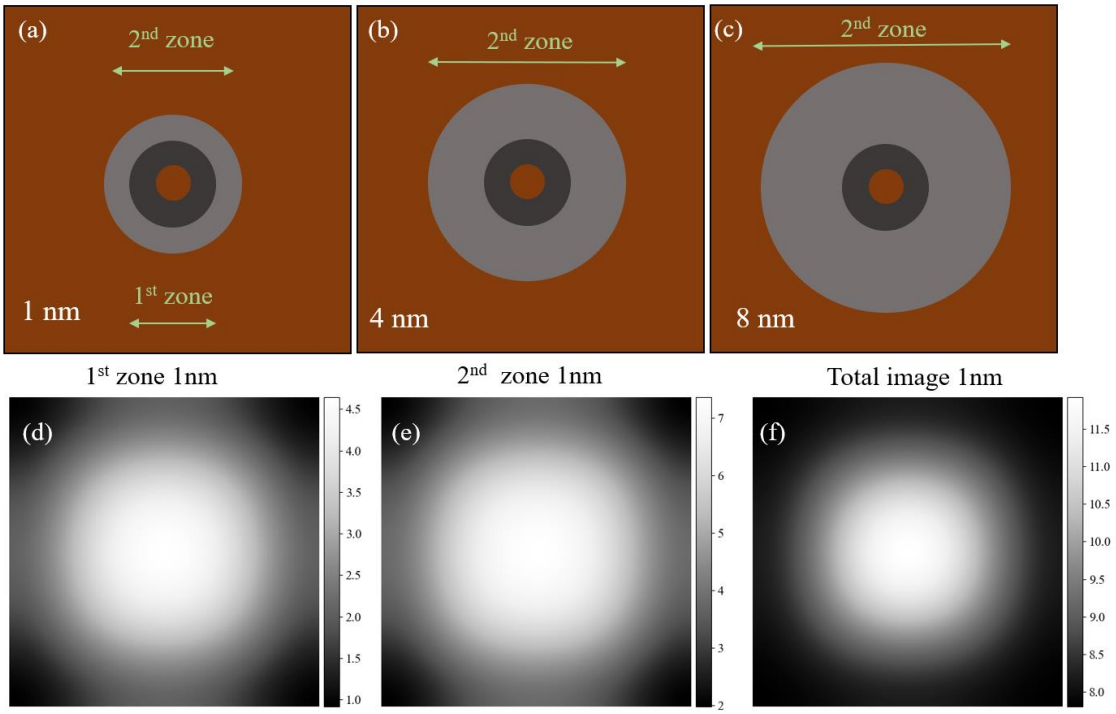


Figure 3.20. (a)-(c) Two different zones and their extent corresponding to 1, 4, and 8 nm focus step with respect to the optic axis for B atom, respectively. The extent of the second zone increases with defocus from 4.49 pm to 10.31 pm, having the same mean bending angle. (d) and (e) Image pattern formation due to various zones at 1 nm focus and (f) overall image of the B atom at 1 nm focus.

Radial distance (pm) & zones	Rim width (pm)	Mean scattering angle (rad)	Length on optic axis (nm)	Intensity (with flux balance)
1-1.77	0.77	0.1092	1	6.82
1.77-8.76	6.99	0.0275	1	7.61
8.76-9.54	0.78	0.0092	1	13.73
8.76-15.36	6.60	0.0067	4	9.93
8.76-18.85	10.09	0.0061	8	11.03

Table 3.04. Various radial zones and corresponding focal length on the optic axis for Zn (Z=30) atom for 1, 4, and 8 nm length on the optic axis. Zones are divided based on scattering angles falling in the same order of magnitude.

Radial distance (pm) & zones	Rim width (pm)	Mean scattering angle (rad)	Length on optic axis (nm)	Intensity (with flux balance)
1-6.04	5.04	0.043	1	9.60
6.04-7.21	1.17	0.0081	1	10.58
6.04-12.25	6.21	0.0060	4	11.63
6.04-15.36	9.71	0.0055	8	11.64

Table 3.05. Various radial zones and corresponding focal length on the optic axis for S (Z=16) atom for 1, 4, and 8 nm length on the optic axis. Zones are divided based on scattering angles falling in the same order of magnitude.

The relationship between the phase of the OEW and the resulting magnitude of intensity can be understood as follows. The meaning of phase change is the change in momentum vector direction, which is similar to Fresnel, Fraunhofer, off-axis electron holography type of interference geometry where wave interference is involved at an angle. This is different from the WPOA type phase shift, where no information on interference geometry is available. Radial interference geometry modifies the unidirectional straight electron interference fringe to a radially symmetric pattern with a peak intensity at the center of the pattern. In the off-axis electron holography, the

intensity pattern oscillates periodically with the same magnitude in unidirectional interference geometry across a mirror plane. Now this phase term appears as carrier frequency Q in off-axis electron holography and, in the present case, alters the peak intensity of the radial pattern depending on field strength and extent of a given radial zone. The field strength and corresponding potential information are the object information that can be interpreted and used to identify atoms.

Radial distance (pm) & zones	Rim width (pm)	Mean scattering angle (rad)	Length on optic axis (nm)	Intensity (with flux balance)
1-3.32	2.32	0.0252	1	8.87
3.32-5.65	2.33	0.0075	1	7.09
3.32-9.15	6.22	0.0061	4	8.40
3.32-12.25	8.93	0.0057	8	8.87

Table 3.06. Various radial zones and corresponding focal length on the optic axis for O (Z=8) atom for 1, 4, and 8 nm length on the optic axis. Zones are divided based on scattering angles falling in the same order of magnitude.

Radial distance (pm) & zones	Rim width (pm)	Mean scattering angle (rad)	Length on optic axis (nm)	Intensity (with flux balance)
1-3.33	2.33	0.022	1	8.42
3.33-5.27	1.94	0.0067	1	7.42
3.33-9.15	5.82	0.0054	4	8.41
3.33-11.87	8.54	0.0050	8	8.87

Table 3.07. Various radial zones and corresponding focal length on the optic axis for N (Z=7) atom for 1, 4, and 8 nm length on the optic axis. Zones are divided based on scattering angles falling in the same order of magnitude.

Radial distance (pm) & zones	Rim width (pm)	Mean scattering angle (rad)	Length on optic axis (nm)	Intensity (with flux balance)
1-2.16	1.55	0.014	1	5.09
2.16-4.49	1.64	0.0069	1	7.77
2.16-7.99	6.21	0.0052	4	7.31
2.16-10.31	8.93	0.0049	8	6.76

Table 3.08. Various radial zones and corresponding focal length on the optic axis for B (Z=5) atom for 1, 4, and 8 nm length on the optic axis. Zones are divided based on scattering angles falling in the same order of magnitude. The corresponding schematic is given in Figure 3.20.

In WPOA, series approximation and PCTF contrast are required to observe intensity patterns due to object potential, but only defocus is sufficient in the present case. In practical HRTEM, the atom contrast never goes to zero even if PCTF is zero. Traditionally, this is explained based on amplitude contrast due to ACTF, which is the higher-order contribution from interaction constant σ according to Eq. 3.26 and is attributed as a non-linear imaging condition. However, the presence of contrast, even if defocus and the corresponding PCTF is zero, can only be explained by the present method in terms of radial interference originating from sectors of potential very close to the nucleus. Therefore, in the current process, as the meaning of phase is different, incoherent zones will have non-identical phase change in terms of Q. Thus, more than one wave function is superimposed on top of each other incoherently, and it will be appropriate to interpret intensity directly to the object information. An alternative proposal on image reconstruction based on intensity directly instead of wavefunctions is discussed in the next chapter.

3.4.4 Effect of spherical aberration and defocus on the image

The effect of third-order spherical aberration (C_s) and defocus (Δf) are now considered to modify the image contrast further. The image contrast calculated based on the alternative method without considering lens transfer does not match the experimental images as a function of the atomic number recorded under the particular settings of imaging conditions. Generally, a suitable combination of C_s and Δf are used to transfer maximum phase contrast for a given bandwidth of

spatial frequency with the best possible point resolution provided by the first zero crossings of PCTF [Figure 3.21]. A point resolution better than 1 Å can be achieved in an aberration-corrected microscope for a combination of $C_s = -35\ \mu m$ and $\Delta f = 8\ nm$ [25]. As already mentioned, the effect of aberration phase shift can be considered either in the diffraction plane or image plane, depending on the requirement. In the image plane, the aberration figure due to spherical aberration is given by a disk of radius r_s at the Gaussian image plane [Figure 3.22]. The effect of C_s and Δf act in opposite directions, which depends on the third power and linearly with the scattering angle, respectively. The effect is either increasing or decreasing the full width at half maxima (FWHM) of the *psf* and impairs or improving the resolution of the imaging process. A typical angular deviation from the ideal ray path and corresponding aberration figure can be correlated for a given combination of C_s and Δf [Figure 3.23 and Table 3.09].

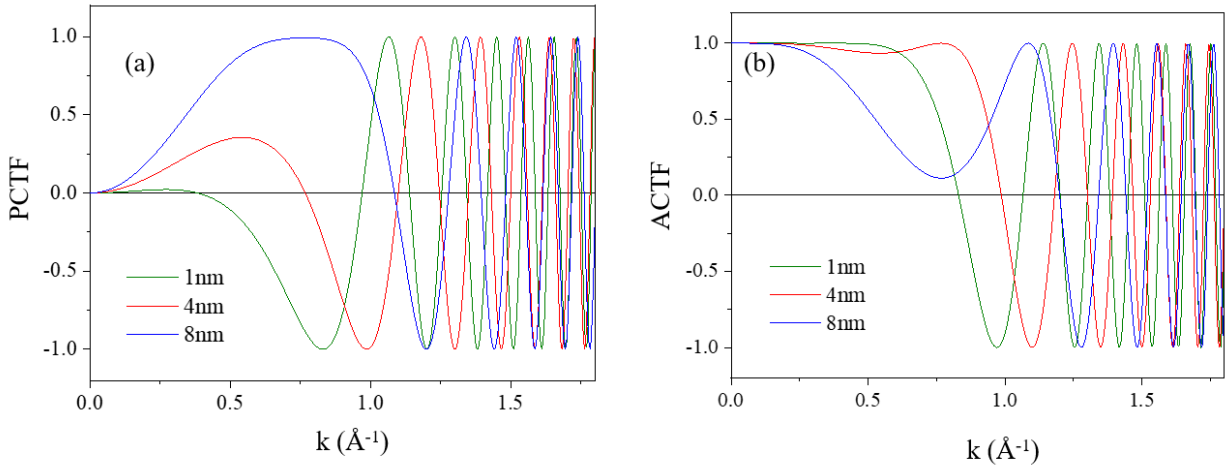


Figure 3.21. (a) PCTF and (b) ACTF function for $C_3 = -35$ µm and positive focus setting of $\Delta f = +1$, +4, and +8 nm.

The effect of defocus (C_1) and third-order spherical aberration coefficient (C_3) on the image aberration is given by

$$r_s = MC_3\theta^3 + MC_1\theta \tag{3.48}$$

Where r_s is the radius of the disk at the Gaussian image point.

Opposite values of C_3 and C_1 reduce the aberration, which is a function of scattering angle [Figure 3.23 (a)]. The effect is to decrease the r_s and improves the resolution of the imaging process. Under

best conditions (e.g., Scherzer criteria), the optimum balance gives the best minimum value of r_s, also called the disk of least confusion.

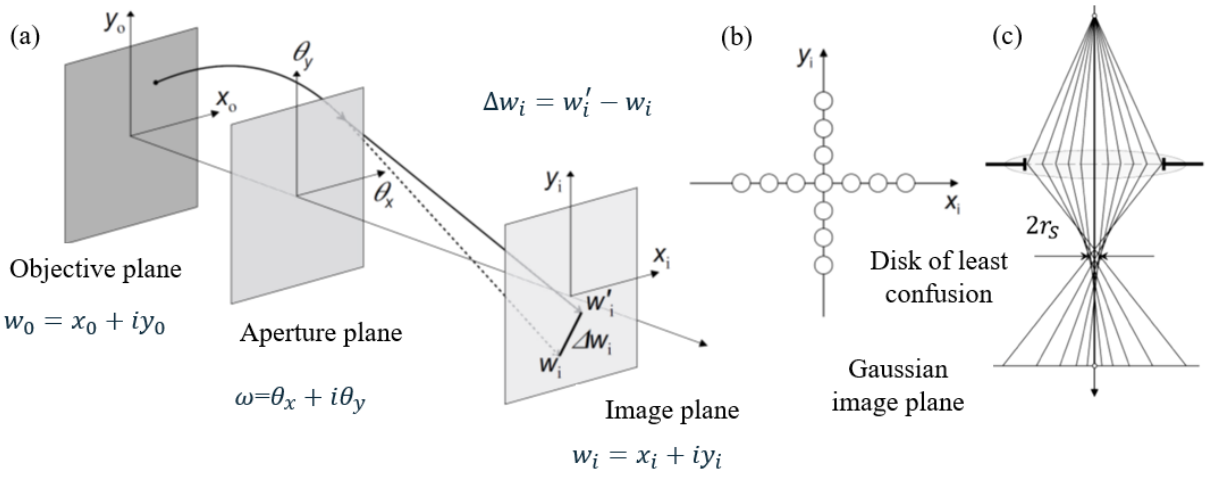


Figure 3.22. (a) Definition of aberration figure in the image plane, (b) effect of C_s on object point on the image plane, and (c) schematic representation of the effect of positive Δf for a given C_s on the image blurring in terms of aberration figure. Copyright (2015) by World Scientific Publishing Company [3].

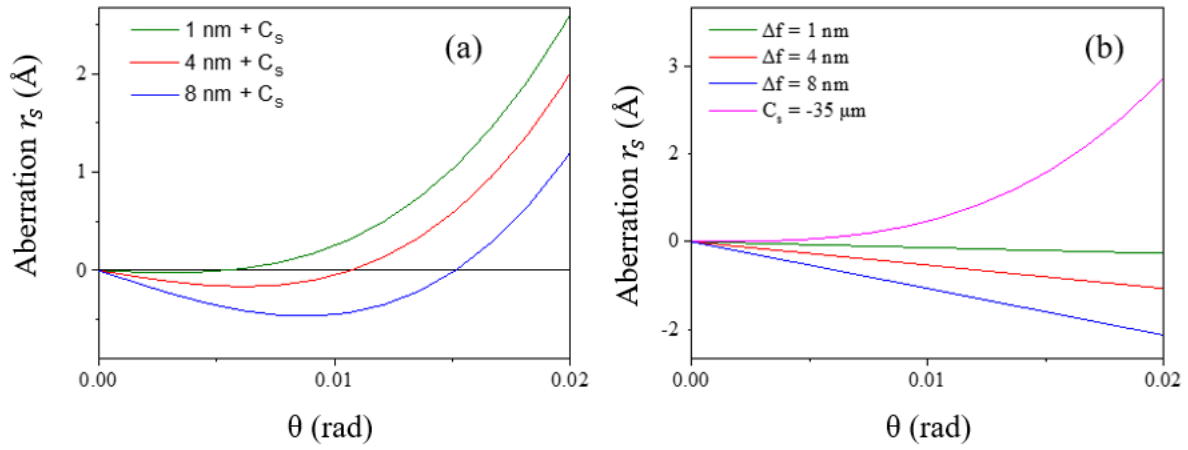


Figure 3.23. (a) Effect of positive Δf and negative C_s on r_s for different focus setting of $\Delta f = +1$, +4 and +8 nm for $C_3 = -35$ µm. (b) Effect of positive Δf and $C_s = 0$ are shown. Note the opposite effect of negative C_s and positive Δf on the r_s.

Experimental images obtained in HRTEM for monolayer MoS_2 and BN for different defocus values are shown in Figure 3.24. Depending on the electrostatic zones and associated

scattering angle, the diffraction-limited information will be blurred further for a given set of C_s and Δf. The peak intensity will be reduced due to aberration blurring for zones with scattering angles more than what can be adjusted through PCTF. The effect of Peak image intensity and contrast difference between Mo and B atoms after considering $C_s = -35\,\mu m$ and $\Delta f = 1, 4, \text{and } 8\ nm$ is shown in Figure 3.25 (b) and listed in Table 3.01. The modified peak intensity has been calculated based on the flux balance approach considering the ratio between the original interference field width to the final blurred area. Kindly note that the simulated peak intensity values at focus close to zero remain a little over reference background, one, irrespective of atomic numbers. This corresponds well with the experimental observation [Table 3.01].

Third-order spherical aberration $\mathbf{C_s}$ ($\boldsymbol{\mu m}$)	Defocus $\mathbf{\Delta f}$ (nm)	Scattering angle (rad)	Aberration figure (pm)
$\mathbf{-35}$	+1	0.011	31.15
$\mathbf{-35}$	+4	0.011	0.67
$\mathbf{-35}$	+8	0.011	43.10
0	+1	0.011	10.60
0	+4	0.011	42.42
0	+8	0.011	84.85

Table 3.09. Example scattering angle vs. aberration figure for various combinations of C_s and Δf as extracted from Figure 3.23.

The calculated intensity image of the Mo atom at three different focus values is shown in Figure 3.25 (a). The peak values are smaller for B than the Mo case as the outer zone's extent is reduced due to the smaller magnitude to the potential field. Figure 3.25 (a) is the graph of peak image intensity vs. atomic number for three different focus values. The peak intensity decreases slightly with a larger focus value. The peak intensity has a dependence of ~ aZ^b where a is a fitting constant and Z is the atomic number. Exponent b changes from 0.4 to 0.26 for change in focal length from 1 to 8 nm. However, the peak intensity values decrease significantly after considering image aberration (sec. 2.3.3 and Figure 3.25 (b)).

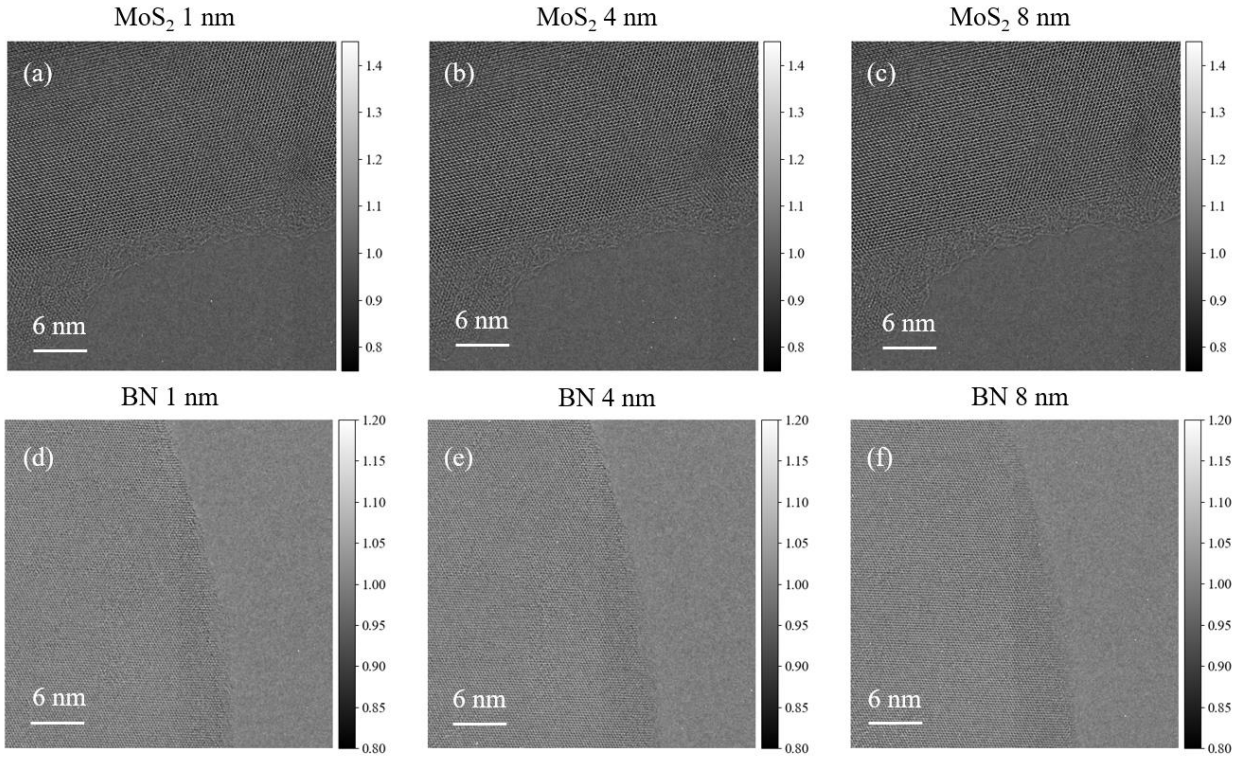


Figure 3.24. Experimental images at three different focus settings of $\Delta f = +1$, +4 and +8 nm for (a)-(c) MoS_2 and (d)-(f) BN.

Now, one can notice that after considering aberration blurring due to *psf* corresponding to each zone and scattering angle, the peak intensity falls significantly compared to the ideal lens case [Table 3.01 and Figure 3.25 (b)]. From the experimental image series recorded under similar conditions as simulation parameters, the peak intensity to reference vacuum ratio is ~ 1.23 (± 0.04) and 1.05 (± 0.02) for Mo and B atoms, respectively, after normalizing with respect to the vacuum intensity. For example, we have obtained a peak value of 1.19 and 1.09 for Mo and B atoms from the calculation for a focus setting of 4 nm (Table 3.01). The peak intensity does not vary significantly in the experimental image for a small variation of focus. Good matching is obtained with the simulated image after considering aberration blurring of intensity from different incoherent zones. Any discrepancy can be explained in terms of slight deviation in scattering angle and associated extent of zones, focus value, consideration of aberration figure while simulating the absolute intensity of atoms. The close agreement between experimental image and simulation results contradicts earlier prediction based on Stobb's factor, where a difference in contrast by a factor of 4 was reported [26].

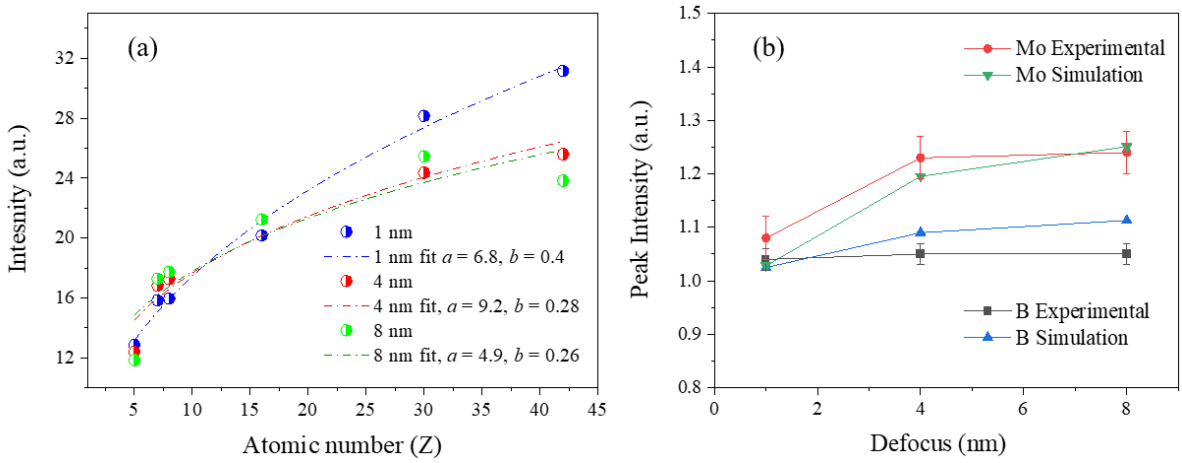


Figure 3.25. (a) Comparison of peak image intensity as a function of atom number for a focus setting of 1, 4, and 8 nm without considering any aberration. (b) Peak image intensity difference between Mo and B atoms after considering $C_s = -35$ µm and $\Delta f = 1, 4,$ and 8 nm.

3.5 Conclusion

In conclusion, an alternative approach for simulating atom images in HRTEM is presented. The method is based on the atom as an electrostatic charge center inducing wave interference along the radial direction and extending the conventional off-axis interference geometry with unidirectional electron biprism. The simulated results corresponding to the image intensity of various atoms are in close agreement with the experimental images of Mo and B atoms.

3.6 Bibliography

[1] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University Press, 1999).

[2] H. Rose, *Geometrical Charged-Particle Optics*, 2nd ed. (Springer, 2012).

[3] R. Erni, *Aberration-Corrected Imaging in Transmission Electron Microscopy* (Imperial College Press, 2015).

[4] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science*, 2nd ed. (Springer, 2009).

[5] J. C. H. Spence, *High-Resolution Electron Microscopy*, Third Edit (Oxford Science

Publications, 2003).

[6] W. O. Saxton, Ultramicroscopy. **55,** 171 (1994).

[7] U. Bhat and R. Datta, J. Appl. Phys. **125**, 154902 (2019).

[8] C. B. Carter and D. B. Williams, *Transmission Electron Microscopy: Diffraction, Imaging, and Spectrometry* (Springer, 2016).

[9] E. J. Kirkland, *Advanced Computing in Electron Microscopy: Second Edition*, 2nd ed. (Springer, 2010).

[10] C. L. Jia, M. Lentzen, and K. Urban, Science. **299**, 870 (2003).

[11] C. L. Jia, M. Lentzen, and K. Urban, Microsc. Microanal. **10**, 174 (2004).

[12] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Großß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser, Nat. Mater. **10**, 209 (2011).

[13] D. Van Dyck, J. R. Jinschek, and F. R. Chen, Nature **486**, 243 (2012).

[14] C. Jin, F. Lin, K. Suenaga, and S. Iijima, Phys. Rev. Lett. **102**, 195505 (2009).

[15] C. B. Eisenhandler and B. M. Siegel, J. Appl. Phys. **37**, 1613 (1966).

[16] O. Scherzer, J. Appl. Phys. **20**, 20 (1949).

[17] H. Lichte and M. Lehmann, Reports Prog. Phys. **71**, 016102 (2008).

[18] D. Geiger, H. Lichte, M. Linck, and M. Lehmann, Microsc. Microanal. **14**, 68 (2008).

[19] B. Fultz and J. M. Howe, *Transmission Electron Microscopy and Diffractometry of Materials*,(Springer Science & Business Media, 2012).

[20] U. Bhat and R. Datta, arXiv:2103.00538 (2021).

[21] D. Van Dyck and M. Op de Beeck, Proceedings of the XIIth International Congress on Electron Microscopy, pp. 26–27 (1990).

[22] P. Schiske, J. Microsc. **207**, 154 (2002).

[23] M. Linck, H. Lichte, and M. Lehmann, Int. J. Mat. Res. **97**, 890 (2006).

[24] B. Loukya, D. S. Negi, K. Dileep, N. Pachauri, A. Gupta, and R. Datta, Phys. Rev. B - Condens. Matter Mater. Phys. **91**, 134412 (2015).

[25] K. W. Urban, C. L. Jia, L. Houben, M. Lentzen, S. B. Mi, and K. Tillmann, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **367**, 3735 (2009).

[26] M. J. Hÿtch and W. M. Stobbs, Ultramicroscopy **53**, 191 (1994).

Chapter 4

Insights and alternative methods on the phase retrieval in HRTEM

This chapter proposes an alternative reconstruction method for retrieving the object exit wave function from the recorded image intensity pattern in high-resolution transmission electron microscopy (HRTEM). The method is based on applying a modified intensity equation representing the HRTEM image. Phase shift extracted from the experimental images of MoS_2, BN, and ZnO are in excellent agreement with the theoretical reference values. Additionally, it is shown that the Fourier series expansion of diffraction pattern is effective in retrieving the isolated and periodic image functions of a specific form directly. However, for aperiodic object information, e.g., defects, dopants, edges, etc., the first method works entirely.

Manuscript under preparation.

4.1 Introduction

Phase (ϕ) is the fundamental information derived from the atomic resolution phase-contrast microscopy [1–6]. The phase-contrast technique enables a detailed examination of the object based on the atomic arrangement and electron density. As the electron beam passes through the ultra-thin specimen in TEM, the phase is altered because of the specimen potential. The change in phase ($\Delta\phi$) of the probe electron wave leads to the formation of specific intensity patterns in the respective image and diffraction planes. The phase information is not recorded directly, as the detectors are only capable of recording the intensity. The intensity patterns can be recorded through any suitable imaging device, e.g., a charge-coupled device (CCD) camera. In case of electron diffraction, $\Delta\phi$ carry information not only on the crystallographic phase of the material along high symmetry orientation but also on the electrostatic potential that is essential for the identification and counting of atoms, extracting information on the chemical bonding from the experimental images [7–13]. To get the phase shift, we have to obtain the wave function of the electron, which is not so straightforward. The method to extract the wave function of the electron wave from the intensity is known as the object exit wave reconstruction (OEW). There are a few existing experimental and associated numerical phase retrieval methods in high-resolution transmission electron microscopy (HRTEM), e.g., through focal image series reconstruction based on HRTEM [2,14,15], atomic resolution off-axis electron holography [16,17], fitting object function directly by intensive computer simulation or so-called direct method [9], transport of intensity equation (TIE) [18], and phase velocity [7], etc. Some of the techniques mentioned above work both at medium and atomic resolution. The TIE method was developed for medium resolution applications.

OEW reconstruction is the inverse process of image simulation. Among various methods, complexities in OEW reconstruction based on conventional through focal HRTEM image series are addressed in the present discussion. Its analogy and differences to off-axis electron holography are highlighted. Defocus HRTEM image is equivalent to in-line holography [19]. Even though HRTEM is the most used technique to visualize atoms, it fails to obtain the direct phase information of the electron waves scattered from the object. In off-axis electron holography, the wave interference occurs at an angle between the reference and the object waves. The retrieval of OEW is performed at first by Fourier transformation (FT) of the image containing electron

interference pattern, then selecting one of the two sidebands (SBs) which are complex conjugate (or twin image) to each other followed by inverse-FT [4]. This procedure isolates the central band (CB) and the twin image wavefunctions from the recorded image. The phase and amplitude can be evaluated either by the arctan function corresponding to inverse-FT or fitting the inverse-FT pattern with image simulation. As the starting data is the image, the FT procedure does not lead to the loss of any information in the crystallographic phase. Inverse FT can return the image intensity pattern. Deconvolution of coherent aberration envelope can be performed posterior that modifies the aberration figure in the image plane.

In HRTEM, the CB and twin image wave functions overlap in the diffraction plane, and the FT procedure cannot separate them in the frequency space, unlike off-axis electron holography [19,20]. Most of the reconstruction methods in HRTEM involve multiplying the image intensity recorded at different focus settings with a complex filter function. The filter function consists of a coherent aberration envelope corresponding to each focus. Summing up all the images will eliminate the unwanted twin image and non-linear components and result in the desired OEW function. The complexities of reducing the unwanted components from the preferred OEW function led to several reconstruction algorithms [15].

However, in the present chapter, it is demonstrated that by marginally modifying the intensity equation describing the HRTEM image, it is possible to retrieve the phase information directly from the atomic resolution images. The same equation in the wavefunction formalism is used for existing OEW reconstruction procedures in the in-line and off-axis electron holography. Two types of phases are of importance here. One is the crystallographic phase describing the scattering distribution of potential corresponding to the isolated or periodic arrangement of atoms forming a specific image or diffraction pattern in the respective planes. The second one is the change in phase of probe electron wave due to strength of atomic potential, which is equivalent to electron density in X-ray crystallography, determining the intensity of the dots in the image pattern. A comparative discussion is provided between the existing reconstruction methodologies of OEW in HRTEM, off-axis electron holography, and the present proposal [21–23]. The retrieval of OEW function due to object, lens imperfections, and geometry of interference depends on how the phase change is incorporated and the equations used in the mathematical formulations. However, it is shown that the phase information in terms of crystallography and object potential

is not lost in the image for both in-line and off-axis electron holography. The presence of a reference wave ensures the modulation of the intensity pattern and retaining the phase information in the image plane. The results obtained for MoS_2, BN, and ZnO are in excellent agreement with the theoretical reference values within the specified resolution limits. The method follows the Born rule of probability density and addresses the twin image issue in in-line electron holography. However, phase information is completely lost for the intensity pattern recorded in the diffraction plane. But it is shown that it is possible to retrieve the complete phase information directly from the diffracted intensity. This latter method is based on the cosine-based Fourier series expansion of the diffraction pattern similar to the zero-phase retrieval based on Patterson's function in X-ray crystallography for small molecular systems [24,25]. This method works accurately for a particular type of function, both in isolated and periodic form. However, for aperiodic object information, e.g., defects and dopants, the first method works in its entirety.

4.2 Experimental techniques

TEM samples of MoS_2 and BN layered materials are prepared by ultrasonication of respective powders (Sigma Aldrich) for 40 mins to exfoliate monolayers and few layers, followed by drop-casting on a holey carbon grid [26]. A cross-sectional TEM specimen of ZnO epitaxial thin film is prepared by first mechanical thinning and then Ar ion polishing to perforation. HRTEM images are recorded in an aberration-corrected FEI TITAN^{3TM} 80-300 kV transmission electron microscope operating at 300 kV with optimum C_S and defocus settings [27,28]. ZnO epitaxial thin film is grown homoepitaxially on a (0001) ZnO substrate by pulsed laser deposition (PLD) following a specific growth procedure described in Ref. [6,29].

4.3 Existing methods of exit wave reconstruction

In this section, existing methods of OEW reconstruction are discussed for both HRTEM and off-axis electron holography methods.

4.3.1 Various schemes on HRTEM focal series reconstruction

4.3.1.1 Focal variation method

Reconstruction methods associated with the HRTEM through focal image series aim to retrieve the unknown phase ϕ of the OEW function of the form $\psi_i = A(x,y)e^{i\phi(x,y)}$Where $A(x,y)$ represents the amplitude of the object exit wave function. The OEW function can interpret the object structure $f(x,y,z)$ based on the model methods. Generally, the illumination is considered monochromatic, and partial coherence theory appears into the stage for semi-monochromatic waves. The theoretical model behind the off-axis electron holography technique to retrieve the OEW function is straightforward (sec. 4.3.2). The methods corresponding to in-line holography using through focal HRTEM image series are much more elaborate, involving intensive data refinement and fitting procedures [17]. This is due to the nature of the fitting equation considered to retrieve the ψ_i that attempts to reduce the effect of complex conjugate ψ_i^* and non-linear image components $\psi_i\psi_i^*$ from the recorded image set.

There exist few different focus variation methods e.g., Wiener formulation by Schiske, 3D paraboloid method (PM), maximum likelihood method (ML), and various numerical schemes associated with them [2,3,30]. Saxton showed the equivalence between the different reconstruction techniques in terms of equivalence in restoring filter applied to the intensity expression in Fourier space to retrieve the wave function. The simplified form of the restoring filter function fon *N* number of images is given as

$$r_n(k) = \frac{1}{N}\exp\{i\gamma(k)\} \tag{4.01}$$

and the associated aberration function

$$\gamma_n(k) = \pi C_s\lambda^3k^4 - \pi\lambda z_n k^2 \tag{4.02}$$

finally, the restored wave function is written as

$$\psi(k) = \sum_n I_n(k)r_n(k) \tag{4.03}$$

Where, C_s is the third-order spherical aberration, λ is the wavelength, z_n is the defocus corresponding to the n^{th} image, and k is the spatial frequency. The application of restoring the filter function is similar to deconvolving the effect of aberration. The inverse-FT of Eq. 4.03 will return

the wave function in the image plane. This is only possible if one starts with the intensity pattern recorded in the image plane with sufficient spatial resolution, and during inverse-FT, spatial information on the crystallographic phase is preserved through $Re(x,k)$ (real part) and $Im(x,k)$ (imaginary part) of Fourier waves. The pattern of the wave function in terms of phase ϕ and amplitude A thus obtained is required to be fitted with the model calculation for further interpretation.

However, Saxton derived the final form of the wave function without considering any restoration filter from a different start in the intensity expression given in Eq. 4.04. The approach provides insight into the effect of restoration on dispersing conjugate wave function and non-linear image components. Under perfect coherence, the image intensity at some focus *z* near the Gaussian image plane is written as

$$i(x,z) = |1 + \psi_i|^2 = 1 + \psi_i(x,z) + \psi_i^*(x,z) + h(x,z) \tag{4.04}$$

Where, $\psi_i(x,z)$ is the desired part or the OEW, $h(x,z) = \psi_i(x,z)\psi_i^*(x,z)$ and $\psi_i^*(x,z)$ are the unwanted non-linear term and complex conjugate, respectively. $\psi_i^*(x,z)$ is the twin image of $\psi_i(x,z)$. Kindly note that the reference wave ψ_0 is set to one by considering axial illumination ($k_0 = 1$) in the above expression.

The 2D Fourier transformation of Eq. 4.04 and then explicitly adding the dependence on defocus results in the following expression.

$$I(k,z) = \delta(k) + \psi(k)\exp(\pi i\lambda z k^2) + \psi^*(-k)\exp(-\pi i\lambda z k^2) + H(k,z) \tag{4.05}$$

Where, $H(k,z)$ is the non-linear component that describes the autocorrelation between the two linear terms in the reciprocal space. As proposed by van Dyck, the paraboloid method can be derived from the above expression by taking 3D-FT with respect to defocus where the wave function and its complex conjugate follow the reflected parabola from the reference diffraction plane [2].

Now, within the coherent detection, i.e., multiplying by corresponding phase conjugate and summing over *N* images after assuming k is non-zero, which allows omitting the delta function and for a constant k, the image intensity becomes

$$\sum_n I_n \exp(-\pi i\lambda z_n k^2) = \psi \sum_n 1 + \psi^* \sum_n \exp(-2\pi i\lambda z_n k^2) + \sum_n H_n \exp(-\pi i\lambda z_n k^2) \tag{4.06}$$

From the above expression, it was concluded that the wave functions ψ and ψ^* accumulate to *N* and $\sqrt{N}$ times to its original value, respectively. The non-linear term H_n behaves randomly with the defocus. Thus, by dividing the above equation by *N*, the wanted OEW function may be recovered. Eq. 4.06 is widely considered in almost all the through-focus image series OEW reconstruction. All the efforts to develop image reconstruction codes primarily deal with eliminating the non-linear and complex conjugate terms and finding the best fit with the model calculation. However, complex conjugate and non-linear terms will always be present depending on how their weights are subdued.

HRTEM intensity is given by,

$$I_{in\ line} = |\psi_0 + \psi_i|^2 = A_0^2 + A_i^2 + 2A_0A_i cos(\phi_i - \phi_0) \tag{4.07}$$

Where $\psi_0 = A_0(x,y)e^{i\phi_0(x,y)}$ is the initial wave function of the electron with amplitude $A_0(x,y)$ and phase $\phi_0(x,y)$. The image intensity pattern is based on self-interference between reference incident and scattered waves within the picture of single-electron wave interference phenomena (Eq. 4.08 and 4.09). The expression is similar to the wave interference between the reference and object waves in off-axis electron holography (sec. 4.3.2). However, in off-axis geometry, there is an additional phase term Qx due to wave interference at an angle that gives spatial modulation in the interference field [5]. Moreover, Eq. 4.04 is in an intermediate state to eliminate the effect of twin image and non-linear terms by working in the diffraction plane. However, the final form of the expression implies that fitting the intensity equation alone and evaluating the phase term should, in principle, allow extracting the relative phase change from the image plane. For more details on associated twin image wave functions and applicability of Eq. 4.07, see sec. 4.4.1. The results based on the above schemes can be found in Ref. [6,17].

Finally, the approach based on partial coherence theory considers the effect of finite source size, chromatic defocus spread, current-voltage fluctuation of the instrument, objective aperture size, and wave aberration function [1,31,32]. The reconstruction method based on partial coherence theory is an iterative linear restoration that addresses the residual non-linear term. Repeated application of linear restoring filter from the subtraction of the calculated non-linear term improves the initially guessed wave function. Various derivations available based on partial coherence theory are given in sec. 4.3.1.3. One can notice that the coherence and interference phase

shift in formalism originates in the convolution procedure in real space and cross-correlation in Fourier space.

Eq 4.04 can also be written following the subsequent description of holography given in Born and Wolf (p 504, section 8.10.1) [21]. The illuminating wave function ψ_0 considered to be 1, similar to Schiske's axial illumination condition for which $k_0 = 0$.

$$\begin{aligned} i(x,z) = |\psi_0 + \psi_i|^2 &= \psi_0\psi_0^* + \psi_0^*\psi_i(x,z) + \psi_0\psi_i^*(x,z) + \psi_i(x,z)\psi_i^*(x,z) \\ &= A_0^2 + A_i^2 + A_0A_i \exp i\,(\phi_i - \phi_0) + A_0A_i \exp i\,(\phi_0 - \phi_i) \\ &= A_0^2 + A_i^2 + 2A_0A_i cos(\phi_i - \phi_0) \end{aligned} \quad (4.08)$$

After following Born and Wolf,

$$\begin{aligned} i(x,z) = |\psi_0 + \psi_i|^2 &= |A_0 \exp(i\phi_0) + A_i \exp(i\phi_i)|^2 \\ &= \exp(i\phi_0) \exp(-i\phi_0)\, |A_0 + A_i \exp i(\phi_i - \phi_0)|^2 \\ &= A_0^2 + A_i^2 + 2A_0A_i cos(\phi_i - \phi_0) \end{aligned} \quad (4.09)$$

Two slightly different starting arrangements give the same result. The quantities are all real and different from the expression in Eq. 4.04. It depends on the vector nature of the momentum direction whether we will have in-line or off-axis-like geometry. In the off-axis geometry, another phase term appears, a component of wave vectors Qx inside the cosine term, which describes carrier frequency and a reference point for a fringe shift.

The second line of Eq. 4.08 is an intermediate state where imaginary parts will automatically cancel each other and yield only real parts. It means that by fitting the intensity equation and evaluating the cosine term, the relative phase change in the image can be extracted.

In the above discussion, the effect of Cs is ignored but can be incorporated through the exponential phase function or transfer function. Experimentally this is akin to applying a frequency filter formed by a known complex transfer function corresponding to every focus value to the Fourier space intensity. It is then summing to retrieve the most dominant wave function compared to other contributions, namely the conjugate and non-linear parts. Thus, the complex quantity can be used to calculate the amplitude and phase part of the OEW function by the arctan function.

4.3.1.2 Paraboloid method

Van Dyck discusses two different approaches. The second one has a similarity with the Kirkland method.

In the first approach by Van Dyck [2], what is widely known as the 3D paraboloid method, the 2D Fourier transform $\psi(k)$ of the wave function $\psi(x)$ is recovered via a summation over *N* number of images.

$$\psi(k) = \frac{1}{N}\sum_n I_n(k)\exp\{-\pi i\lambda z_n k^2\} \tag{4.10}$$

Based on a large set of images recorded at equally and closely spaced focus levels on either side of the Gaussian focus. After summation over *N* images, the simplified restoring filter function takes the form of Eq. 4.01.

And the restored wave function becomes

$$\psi(k) = \sum_n I_n(k)\exp\{i\gamma_n(k)\} = \exp(\pi i C_s\lambda^3 k^4) \times \frac{1}{N}\sum_n I_n(k)\exp(-\pi i\lambda z_n k^2) \tag{4.11}$$

In the above expression, a spherical aberration phase shift is included explicitly. For perfect coherence, after appropriate approximation, the restoring filter function has the same form as Eq. 4.01 (based on Wiener formulation).

Saxton pointed out that in the 3D paraboloid method, the picture can be obtained by 3D Fourier transformation of Eq. 4.06, i.e., to the third image space dimension *z*.

$$\delta(k,k_z) + \psi(k)\delta\left(k_z - \frac{1}{2}\lambda k^2\right) + \psi^*(-k)\delta\left(k_z + \frac{1}{2}\lambda k^2\right) + \int H(k,z)\exp(-2\pi i k_z z)\,dz \tag{4.12}$$

It was shown by Van Dyck that the wanted wave function $\psi(k)$ is localized on the paraboloid at $k_z = \frac{1}{2}\lambda k^2$ in reciprocal space, the complex conjugate will lie on a reflected paraboloid and completely separated except at small *k*.

At $k_z = \frac{1}{2}\lambda k^2$, the transformation with respect to *z* gives the OEW.

$$\psi(x) = \int I(k,z)\exp\{-\pi i\lambda k^2 z\}\,dz \tag{4.13}$$

4.3.1.3 Partial coherence theory

In the second approach by Van Dyck et al. [1], the method is based on partial coherence theory considering the effect of the finite size of the source, chromatic defocus spread, current-voltage fluctuation of the instrument, objective aperture size, and wave aberration function [31,32]. The reconstruction method based on partial coherence theory is known as an iterative linear restoration.

According to Van Dyck, in frequency space, the specimen to image interference is given by

$$I(G) = \phi(0)\phi^*(-G)T(0,-G) + \phi^*(0)\phi(G)T(G,0) + \int_{G'\neq 0, G'\neq G} \phi(G+G')\phi^*(G')T(G+G',G')dG' \quad (4.14)$$

G (G≠0) is the two-dimensional frequency vector, T is the transmission cross coefficient (TCC). The first and second terms in Eq. 4.14 represent the linear interference between the transmitted electron beam and one of the diffracted electron beams. The third term is the non-linear term involving interference between diffracted beams. The overlap between the two wave functions describes the correlation and contributes to the Fourier amplitude in the image. Thus, one can notice that coherency and interference have an origin in convolution operation.

The above equation can be true if there is an integration outside the two wave functions term. The TCC can then be written equal to the convolution of two such wave functions within the integration. See the derivation for the same by Kirkland [23].

The same equation above has been reproduced by Saxton in his review under various other methods as follows,

$$I_n(k) = \delta(k) + \psi(k)w_n(k) + \psi^*(-k)w_n^*(-k) + \sum_{k'} \psi(k')\psi^*(k'-k)m_n(k',k'-k) \quad (4.15)$$

Where $w(k)$ are the transfer functions and $m(k_1, k_2)$ is a mutual transfer function. The coherence effect is considered in the above equation.

The method described above, also known as iterative linear restoration, addresses the residual non-linear term. The repeated application of the linear restoring filter from the subtraction of the calculated non-linear term improves the initially guessed wave function.

4.3.1.4 Partial coherence theory-based formulations

The Fourier transformation of the above equation is written as according to Ref. [15]

$$I(g) = \sum_k \psi(g+k)\psi^*(k) \tag{4.16}$$

The above expression has been interpreted as a pairwise summation of interference between two beams with wave vectors (g+k) and k. In the reciprocal plane, the wave function description is the autocorrelation between waves propagating along different directions governed by reciprocal lattice vectors. The above result is due to the direct Fourier transformation of the image intensity in real space as following

$$I(r) = \psi(r)\psi^*(r) \tag{4.17}$$

According to Kirkland (p 99) [23], if one adds lens response

$$g(x) = |\psi_t(x) \otimes h_0(x)|^2 = [\psi_t(x) \otimes h_0(x)][\psi_t^*(x) \otimes h_0^*(x)] \tag{4.18}$$

The Fourier transform of the above equation leads to

$$\begin{aligned} G(k) &= [\psi_t(k)H_0(k)] \otimes [\psi_t^*(k)H_0^*(k)] \\ &= \int \psi_t(k')H_0(k')\psi_t^*(k'+k)H_0^*(k'+k)d^2k' \\ &= \int T_{cc}(k', k'+k)\psi_t(k')\psi_t^*(k'+k)d^2k' \end{aligned} \tag{4.19}$$

Where, $$T_{cc}^{coh}(k', k'+k) = \exp[-i\chi(k') + i\chi(k'+k)]A(k')A(k'+k) \tag{4.20}$$

Overlap between the two wave functions describes the correlation and contributes to Fourier coefficient k in the image. Thus, one can notice that the coherence is because of convolution.

4.3.1.5 Schiske filter function

Schiske's filter function aimed to carry out the posterior correction of photograph records [30]. According to Schiske, the weakly scattered wave function $\psi_s(x)$ produced while reading the photographic records with incident wave ψ_0, which produces complex amplitude $a(x)$.

Scattered wave is written as

$$\psi_s(x) = a(x)\psi_0(x) = A_0 e^{i(k_0.x_0)} \tag{4.21}$$

In Fourier space, $a(x)$ has the following form

$$a(x) = \int c(k) e^{i(k.x)} dk \tag{4.22}$$

Where $c(k)$ is the amplitude of Fourier waves in frequency space and is written in terms of a complex function and its conjugate with the consideration of aberration function as

$$j(k) = c(k)\exp[-i\gamma(k_0 + k) + i\gamma(k_0)] + c^*(-k)\exp[i\gamma(k_0^* - k^*) - i\gamma(k_0^*)] \tag{4.23}$$

Finally, the Intensity in the image plane immediately beneath the object is written as

$$I(Mx) = I_0 + I_0 \int j(k)\, e^{i(k.x)} dk \tag{4.24}$$

Where the first term on the right-hand side will give the central band or direct component, and the second term will form the scattered wave $\psi_s(x)$ and it's conjugate. $\psi_s(x)$ is now known as the OEW function in modern HRTEM literature. $c(k)$ is recovered by recording *n* images at various focus settings and simplified by considering $k_0 = 0$,

$$c(k) = \frac{i}{2} \frac{\sum_m j_m(k)\, sin\, \gamma_n}{\sum_n sin^2 \gamma_n} \tag{4.25}$$

Where, γ_n is the aberration function with n^{th} defocus.

Once, $c(k)$ is determined, $a(x)$ can be known. $a(x)$ is imaginary for pure phase object like weak phase object. And then the $\psi_s(x)$ can be determined without any residual aberration. This $\psi_s(x)$ has the information on the object structure.

What Schiske showed was to get rid of aberration from the information wanted, i.e. $\psi_s(x)$. But did not address twin images, non-linear components, and the amplitude of the scattered wave, contributing from the conjugate and direct component (DC) part.

4.3.2 Reconstruction scheme in off-axis electron holography

In the standard formalism of off-axis electron holography, one of the two sidebands (SBs) performs inverse Fourier transformation to recover the wave function. The SBs are the convolution between the FT of the wave function and the delta function. The delta function ensures shifting of

the FT of wave functions in frequency space, thus separates it from the central band (CB) and conjugate wave function part.

According to James Loudon [33] (for more details, see Ref. [4,5]), the geometry of interference in off-axis electron holography is given in Figure 4.01.

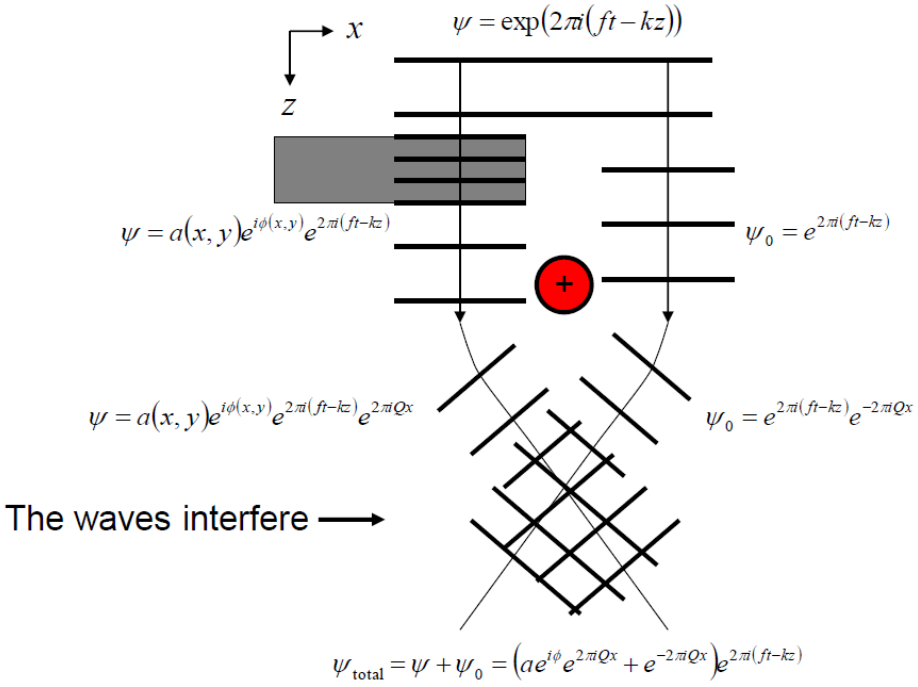


Figure 4.01. Schematic of transmission Off-axis electron holography interference geometry. The probe plane wave split into two halves is brought together for interference. One of the two partial waves goes through the sample and another through vacuum serving as a reference wave. The sample induces an additional phase shift which can be used to interpret the object structure [33].

$$\psi_{total} = e^{2\pi i(ft-kz)}\left(ae^{i\phi}e^{2\pi iQx} + e^{-2\pi iQx}\right) = e^{2\pi i(ft-kz)}e^{i\frac{\phi}{2}}(ae^{i\frac{\phi}{2}}e^{2\pi iQx} + e^{-i\frac{\phi}{2}}e^{-2\pi iQx}) \tag{4.26}$$

$$I_{total} = |\psi_{total}|^2 = \psi^*_{total}\psi_{total} = (ae^{-i\frac{\phi}{2}}e^{-2\pi iQx} + e^{i\frac{\phi}{2}}e^{2\pi iQx})(ae^{i\frac{\phi}{2}}e^{2\pi iQx} + e^{-i\frac{\phi}{2}}e^{-2\pi iQx}) \tag{4.27}$$

$$I_{total} = 1 + a^2(x,y) + 2a(x,y)\cos\left(4\pi Qx + \phi(x,y)\right) \tag{4.28}$$

$$I(x) = 1 + a^2(x) + 2a(x)\cos\big(4\pi Qx + \phi(x)\big)$$
$$= 1 + a^2(x) + a(x)e^{i\phi(x)}e^{4\pi iQx} + a(x)e^{-i\phi(x)}e^{-4\pi iQx} \quad (4.29)$$

Eq. 4.29, according to Ref. [15], is

$$I_{hol}(\vec{r}) = I_0 + I_{ima,inel}(\vec{r}) + I_{ima,el}(\vec{r}) + 2|\mu|A_0A_{el}(\vec{r})\cos\left(2\pi\vec{q}_c\vec{r} + \phi(\vec{r})\right) \quad (4.30)$$

The intensity expression in Eq. 4.27 is Fourier transformed and using $\delta(q-Q) = \int_{-\infty}^{\infty} e^{2\pi i(q-Q)x}dx$ that separates the DC part and two twin image components (SBs) in the frequency plane.

$$F.T.[I(x)] = \delta(q) + F.T.[a^2(x)] * \delta(q) + F.T.\left[a(x)e^{i\phi(x)}\right] * F.T.\left[e^{4\pi iQx}\right] + F.T.\left[a(x)e^{-i\phi(x)}\right] * F.T.\left[e^{-4\pi iQx}\right]$$

$$F.T.[I(x)] = \delta(q) + F.T.[a^2(x)] * \delta(q) + F.T.\left[a(x)e^{i\phi(x)}\right] * \delta(q-2Q)\phi + F.T.\left[a(x)e^{-i\phi(x)}\right] * \delta(q+2Q)$$

$$F.T.[I(x)] = \delta(0) + F.T.[a^2(x)] * \delta(0) + F.T.\left[a(x)e^{i\phi(x)}\right] * \delta(q-2Q) + F.T.\left[a(x)e^{-i\phi(x)}\right] * \delta(q+2Q)$$

(4.31)

Also, Eq. 4.29, according to Ref. [15], can be written as

$$FT[I_{hol}(\vec{r})] = I_0\delta(\vec{q}) + FFT\left[I_{ima,el}(\vec{r}) + I_{ima,inel}(\vec{r})\right] \otimes \delta(\vec{q}) + \hat{V}FFT\left[A_0A_{el}(\vec{r})e^{i\phi(\vec{r})}\right] \otimes \delta(\vec{q}-\vec{q}_c) + \hat{V}FFT\left[A_0A_{el}(\vec{r})e^{-i\phi(\vec{r})}\right] \otimes \delta(\vec{q}+\vec{q}_c)$$

(4.32)

Now select one of the two SBs → $F.T.\left[a(x)e^{i\phi(x)}\right]$

Inverse FT of the sideband → $\left[a(x)e^{i\phi(x)}\right]$ gives the wave function.

The spatial resolution of the technique is determined by the size of the mask placed around the sideband.

4.3.3. Comparison between the HRTEM and off-axis electron holography reconstruction

There is a fundamental difference in the intensity expression given by Born and Wolf on in-line holography and the off-axis electron holography. In off-axis geometry, an additional phase term in the form of wave vector component $Q = \vec{q}_c = 2k_x$ appears due to wave interference at an angle. This defines the carrier frequency in the off-axis electron hologram and changes with the angle of interference controlled by the biprism voltage. The larger the angle of interference, the larger the horizontal wave vector and finer the fringe spacing. Information about the object phase is carried through $\phi(x)$ and appears as a shift on the hologram fringes. In the case of in-line holography, it is the $(\phi_i - \phi_0)$ which appears inside the cosine function, and only the relative phase change in terms of intensity between two object points can be observed in the intensity pattern.

We describe the importance of some components and minor issues with the description in terms of amplitude term $a^2(x)$ and Q in Eq. 4.29 and 4.31. Eq. 4.29 can be understood in comparison to Eq. 4.08 and Eq. 4.07 as follows. There is a difference between Eq. 4.29 and 4.31 describing holographic phase shift or fringe bending in terms of $\phi(x)$, separation of twin image wave functions and the way actual FT works on the experimental HRTEM image intensity embedded with the atomic resolution hologram, i.e., $I(x)\ or\ I(x,y)$ as given in Eq. 4.07 and 4.08. What Eq. 4.29 is stating is that the object phase information $\phi(x)$ appears as fringe bending in the recorded hologram with carrier frequency Q in the image plane. At atomic resolution, the terms $a^2(x)$ and $F.T.[a^2(x)]$ in Eq. 4.28 or 4.31 do not describe the HRTEM intensity pattern like Eq. 4.08 or 4.09, which has an exclusive cosine term $A_0^2 + A_i^2 + 2A_0A_icos(\phi_i - \phi_0)$ in HRTEM intensity expression with object phase term ϕ_i. The interference pattern carries the information in a lattice structure and interprets the associated intensity in terms of object potential.

Though it was mentioned appropriately in Ref. [15] through $I_{ima,el}(\vec{r})$ (Eq. 4.30), however, $I_{ima,el}(\vec{r})$ is written based on OEW of the form $A(\vec{r})e^{i\phi(\vec{r})}$ which is similar to the one used to derive Eq. 4.29. The factor $2A_0A_i$ outside, the cosine term is non-linear, and this cannot be accessed directly and can be used as a uniform background (described in sec. 4.4.1), where the overall intensity value is used due to such term. Moreover, the amplitude of individual complex

waves like $a(x)$ or $A(\vec{r})$ can only be read through $Re[a(x)\cos\phi_i]$ or $Im[a(x)\sin\phi_i]$ forming interference patterns and the intensity or modulus square corresponding to it in the form of $a^2(x)$ should not have any spatial variation in intensity and only contributes to the uniform background. Similarly, the modulus of the individual wave function ψ_0 or ψ_i also will not show any structure and give only constant modulus term contributing to the background.

Therefore, it is only through the trigonometric functions upon interference between waves show up intensity pattern. However, Eq. 4.31 and 4.32 predicts a structure in spatial intensity through $\delta(q) + F.T.[a^2(x)] * \delta(q)$ to explain FFT patterns in CB, which is not the true situation. Now the $F.T.\left[e^{4\pi iQx}\right]$ or $\delta(q-2Q)$ and $F.T.\left[e^{-4\pi iQx}\right]$ or $\delta(q+2Q)$ terms ensure displacement of $F.T.\left[a(x)e^{i\phi(x)}\right]$ a term on the frequency axis having information on spatial modulation of intensity and without constant background term. The inverse FT of $\delta(q+2Q)$ (it should be Q and not $2Q$, see Eq. 4.32 and further details) term gives a uniform background in the image plane and inverse FT of $F.T.\left[e^{-4\pi iQx}\right]$ returns the information on carrier frequency and fringe bending with respect to object phase in the image plane if part of this phase term is added with $F.T.\left[a(x)e^{i\phi(x)}\right]$ before performing inverse FT. Therefore, one can see that the use of particular terms can yield different information upon inverse FT. Thus, the transfer of information on fringe bending is not captured in the FT pattern. It is only captured in the image plane! It is the $F.T.\left[a(x)e^{i\phi(x)}\right]$ that contains the object phase information. The correct form of the object exit wave function (OEW) of the form $a(x)e^{i\phi(x)}$ used should have a relative phase term due to self-interference (see Eq. 4.08). Absolute of this would show a copy of abs-FT pattern of HRTEM image intensity around the two SBs similar to CB and any abs-FT of HRTEM image intensity alone. We elaborate more and describe the difference below.

In Figure 4.01, the OEW due to one-half of the wave should have a form described in Eq. 4.08 or 4.09. Therefore, we modify Eq. 4.29, 4.30, 4.31, 4.32, slightly different forms, and finally, compare it with the actual FT of HRTEM image with hologram. $a(x)e^{i\phi(x)}$ becomes $A_0\exp(i\phi_0) + A_i\exp(i\phi_i)$, we can call it ψ_{inline} or ψ_{HRTEM} and the intensity $I_{HRTEM} = A_0^2 + A_i^2 + 2A_0A_icos(\phi_i - \phi_0)$, is the modulus square of ψ_{HRTEM} as already given in Eq. 4.08 or 4.09. The $I_{ima,el}(\vec{r})$ term used in Eq. 4.30 and 4.32 should be considered as I_{HRTEM}.

$$\psi_{HRTEM} = \psi_{inline} = A_0\exp(i\phi_0) + A_i\exp(i\phi_i) \tag{4.33}$$

Thus, Eq. 4.26 modifies to

$$\psi_{total} = \{(A_0 e^{i\phi_0} + A_i e^{i\phi_i}) e^{2\pi iQx} + A_0 e^{i\phi_0} e^{-2\pi iQx})\}$$

$$= \{(\psi_{HRTEM} e^{2\pi iQx} + A_0 e^{i\phi_0} e^{-2\pi iQx})\} \quad (4.34)$$

And Eq. 4.27modifies to

$$I_{total} = |\psi_{total}|^2 = \psi^*_{total}\psi_{total}$$

$$= (\psi_{HRTEM} e^{2\pi iQx} + A_0 e^{i\phi_0} e^{-2\pi iQx})(\psi^*_{HRTEM} e^{-2\pi iQx} + A_0 e^{-i\phi_0} e^{2\pi iQx}) \quad (4.35)$$

$$I_{total} = A_0^2 + \psi_{HRTEM}\psi^*_{HRTEM} + A_0 e^{-i\phi_0}\psi_{HRTEM} e^{4\pi iQx} + A_0 e^{i\phi_0}\psi^*_{HRTEM} e^{-4\pi iQx}$$

$$= 1 + \{A_0^2 + A_i^2 + 2A_0 A_i cos(\phi_i - \phi_0)\} + (A_0^2 + A_0 A_i e^{i(\phi_i - \phi_0)}) e^{4\pi iQx} + (A_0^2 + A_0 A_i e^{-i(\phi_i - \phi_0)}) e^{-4\pi iQx} \quad (4.36)$$

And

$$F.T.(I_{total}) = F.T.(1) + F.T.(I_{HRTEM}) + F.T.(A_0 e^{-i\phi_0}\psi_{HRTEM}) * \delta(q - 2Q) +$$

$$F.T.(A_0 e^{i\phi_0}\psi^*_{HRTEM}) * \delta(q + 2Q) \quad (4.37)$$

Eq. 4.37 is the correct form in terms of $\delta(q + Q)$ as used in Eq. 4.32 [15].

$$F.T.(I_{total}) = F.T.(1) + F.T.(I_{HRTEM}) + F.T.(A_0 e^{-i\phi_0}\psi_{HRTEM} e^{2\pi iQx}) * \delta(q - Q) +$$

$$F.T.(A_0 e^{i\phi_0}\psi^*_{HRTEM} e^{-2\pi iQx}) * \delta(q + Q) \quad (4.38)$$

Eq. 4.37 now correctly describes the HRTEM image recorded with atomic resolution hologram in the CB and the wavefunction ψ_{HRTEM} has plane wave term $A_0 e^{-i\phi_0}$ outside of it. However, the absolute square should now have a structure observed around SB [Figure 4.02]. Earlier consideration of the wavefunction in the form $a(x)e^{i\phi(x)}$ will not return any structure unless a cosine term is involved upon the absolute square. This is only ensured by writing the I_{HRTEM} in the form of Eq. 4.08 or 4.09.

Saxton's equation, as the primary wave function is considered 1, can avoid integration and convolution between wave function and transfer function multiplication in reciprocal space. Kindly see below for a unified description of the above two equations

$$g(x) = \{(\psi_0 + \psi_i) \otimes h\}\{(\psi_0^* + \psi_i^*) \otimes h^*\}$$

$$G(k) = \{\psi_0(k)H(k) + \psi_i(k)H(k)\} \otimes \{\psi_0^*(k)H^*(k) + \psi_i^*(k)H^*(k)\}$$

$$= \int \{\psi_0(k')H(k') + \psi_i(k')H(k')\}\{\psi_0^*(k'+k)H^*(k'+k) + \psi_i^*(k'+k)H^*(k'+k)\}d^2k' \quad (4.39)$$

Without transfer function, the above expression becomes

$$g(x) = \{(\psi_0 + \psi_i)\}\{(\psi_0^* + \psi_i^*)\}$$

$$G(k) = \{\psi_0(k) + \psi_i(k)\} \otimes \{\psi_0^*(k) + \psi_i^*(k)\}$$

$$= \int \{\psi_0(k') + \psi_i(k')\}\{\psi_0^*(k'+k) + \psi_i^*(k'+k)\}d^2k' \quad (4.40)$$

In Kirkland method, also called ML, MAP and MIMAP, attempts to solve the non-linear imaging problem exactly for the following equation set

$$I_n(k) = \sum_{k'} \psi\,(k')\psi^*(k'-k)m_n(k', k'-k) \quad (4.41)$$

The above equation is re-written from Eq. 4.36, with the primary beam treated separately. As there is no obvious relationship with the Paraboloid method, a long focal series may not be required. Eq 4.38 suggests that while performing inverse FT, cosine fringes will be available in the real space image, and the location of SBs will be at Q and not 2Q. The Q position is more appropriate as in the calibrated fast Fourier transform (FFT) pattern, SBs appear at Q, i.e., ~ at 2.5 $Å^{-1}$ for holography fringe spacing of 0.4 Å [Figure 4.02]. Though the correct form of this is used in Ref. [15], additional phase term $e^{2\pi iQx}$ is not considered. However, after absolute FT, the terms $e^{-i\phi_0}$ and $e^{2\pi iQx}$ will be canceled out and only ${A_0}^2 I_{HRTEM}$ will exist. This is a more realistic situation if one considers how the FT of the experimental HRTEM image, including the hologram, works. The experimental hologram combines (HRTEM pattern + cosine hologram) the sample area and cosine hologram in the vacuum area with constant background for both types of areas.

By examining how the FT process works on the experimental image, we can understand CB and two SBs in off-axis electron holography and compare the origin of the twin images between off-axis and in-line holography. The generation of the abs-FFT (fast Fourier transform) pattern works as follows. For any function $f(x)$, in the present example, an HRTEM image with atomic resolution hologram will be expanded into various cosine and sine harmonics. It is the

integration value of $f(x)\cos(2\pi kx)$ or $f(x)\sin(2\pi kx)$ for a given spatial range that will determine the pattern. The intensity in a FT spot means the integration is non-zero, and zero means integration is zero for a given frequency. Thus, the origin of two SBs is due to the non-zero value of $\int_{x1}^{x2} Image \times \cos(2\pi kx)\, dx$ for $k = Q$. As the hologram has a cosine fringe with some background and a different higher amplitude than the FT cosine function, multiplication with another cosine function with the same frequency match will yield a high integration value (see illustration in Figure 4.03). That is why SB central spot intensity is symmetric to CB [Figure 4.02]. Now the frequency $Q + ng$, where g is the reciprocal lattice vector and n = 1, 2, 3…., will also yield spots similar to spots around CB but with fewer numbers with higher frequencies as the oscillation of trigonometric functions are high compared to frequencies near CB, giving a smaller value of integration.

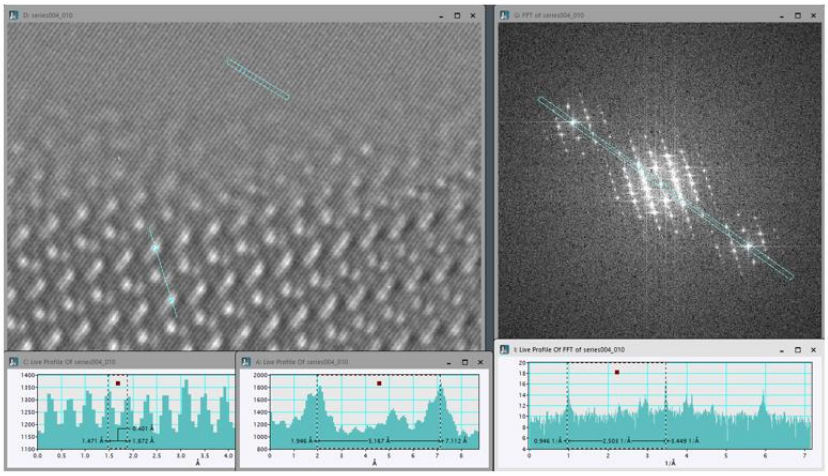


Figure 4.02. ZnO lattice image with atomic resolution hologram with the fringe periodicity of 40 pm. The sideband location is at 1/40 pm in the absolute FFT image on the right-hand side.

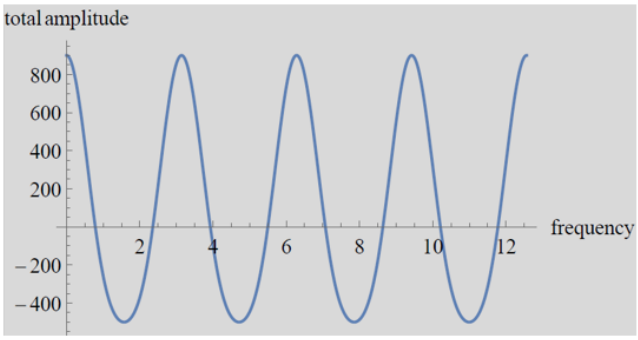
From the illustration in Figure 4.02, one can see that the information on the wave functions given by FFT spots around the SBs, has contributed not only from the intensity of crystal periodicity but also from cosine periodicity in the recorded image. The same is true for FFT spots around CB. This is because the $f(x)$ or image function is the same for CB and SB while performing FT.

Now FT of the experimental image can be written as

$$F.T.(I_{total\ HRholo})\ for\ all\ frequencies = F.T.(I_{total\ HRholo})\ for\ frequency\ g = 0\ and\ \pm ng$$

$$+F.T.(I_{total\ HRholo}) * \delta(q-Q)\ for\ frequency\ g = Q\ and\ Q \pm ng$$

$$+F.T.(I_{total\ HRholo}) * \delta(q+Q)\ for\ frequency\ g = -Q\ and\ -Q \pm ng \quad (4.42)$$



Plot[{(700 + 200 * Cos[2x]) * Cos[2x]}, {x, 0,4Pi}]

Figure 4.03. Schematic showing multiplication of two cosine functions having the same frequency will result in non-zero integration value for a given period. This is similar to the formation of SB intensity while performing FT of HRTEM image embedded with hologram.

4.4 Alternative methods of phase retrieval

In this section, the alternative proposals are introduced to retrieve the phase (ϕ) and amplitude (A) of the OEW function from the image and diffraction planes. The methods are straightforward and do not require through focal image series acquisition. Like off-axis electron holography, only a single image embedded with the hologram is sufficient.

4.4.1 Recovering phase from the HRTEM image intensity pattern

The method described here works on atomic resolution HRTEM image recorded under suitable imaging conditions, i.e., with a particular combination of spherical aberration coefficient C_s and defocus Δf that sets the optimum contrast and resolution. A complete evaluation of Eq. 4.04, based on wavefunction formalism widely used for standard HRTEM image reconstruction, yields Eq. 4.07. A further modification leads to Eq. 4.43. The object phase can be recovered by

applying Eq. 4.43, describing the image intensity pattern in HRTEM within few nanometers from the Gaussian image plane.

$$I_{in\ line}(x,y) = |\psi_0 + \psi_i|^2 = A_0^2 + A_i^2 + 2A_0A_i\ sin\{\phi_i(x,y)\}$$

$$= \alpha I_0 + \beta I_0 + 2\sqrt{\alpha I_0 \times \beta I_0}\ sin\{\phi_i(x,y)\} \quad (4.43)$$

Where, I_0 is the mean vacuum intensity. The factors α and β represent the fraction of direct and scattered part of the intensity and can be determined by analyzing the image pattern, where $\alpha + \beta = 1$.

From the example micrograph in Figure 4.04, one can determine the mean values of vacuum and mean intensity of image area where periodic dots are present with high and low intensity. The standard deviation will be much smaller for the vacuum area compared to the image area. A typical example of the vacuum mean intensity of MoS_2 image is Mean=13133.8, standard deviation = 423.878, and image area mean 13023.9 with standard deviation 1695.06. standard deviation in the image area is approximately equal to $I_{\max dot}$ or $I_{\min dot}$. Now the values of β and α are determined by $\frac{I_{\max dot}}{I_{vacuum\ mean}}$ and $1-\beta$, respectively. Typical values of β and α are found to be ~ 0.88/0.95 and 0.12/0.05, respectively, from the experimental images of MoS_2/BN.

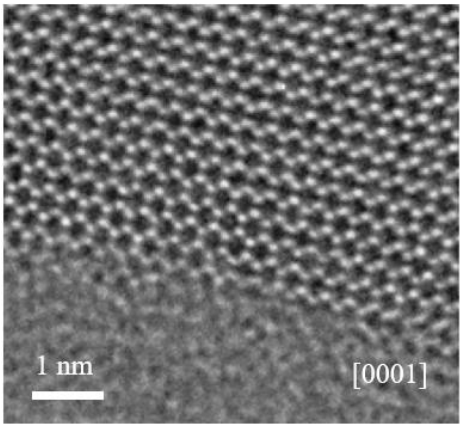


Figure 4.04. High-resolution transmission electron micrograph of MoS_2 along [0001] Z.A.

Intensity expression corresponding to the off-axis electron hologram is given by

$$I_{off\ axis} = 1 + a^2(x,y) + 2a(x,y)\cos\{2\pi Qx + \phi(x,y)\} \quad (4.44)$$

The appearance of sinusoidal function in Eq. 4.43 in contrast to Eq. 4.07 is due to the OEW phase term ϕ_0 has a relative phase term $\pi/2$, i.e., $\Delta\phi = \frac{\pi}{2} - \phi_i$, with respect to the vacuum phase, in addition to ϕ_i due to object potential in the case of HRTEM, which sets the vacuum phase value to zero. This phase difference is absent in off-axis electron holography (compare Eq. 4.43 and 4.44). Two halves of the wave on either side of the biprism carry the same vacuum reference phase term and are eliminated in the final intensity expression. An intuitive physical picture of the origin of such a $\pi/2$ phase shift between diffracted and primary incident waves based on interference geometry is provided in Figure 4.05.

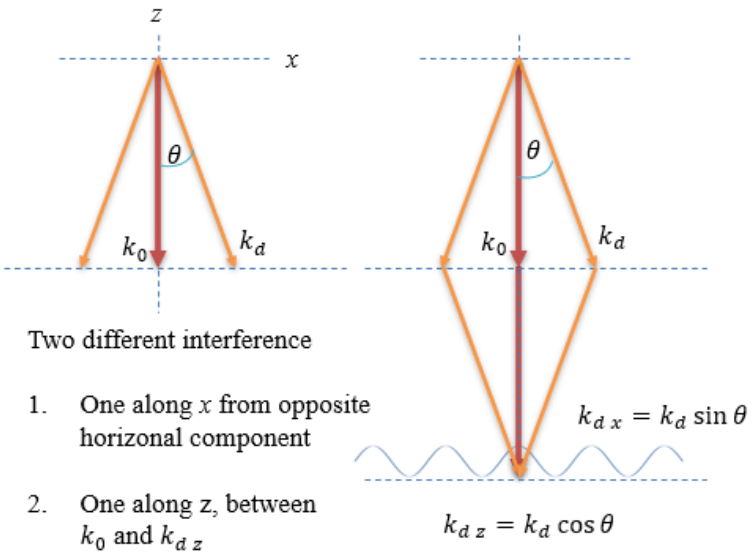


Figure 4.05. Schematic representation of wave vector components taking part in the interference. Horizontal components are responsible for intensity patterns, while vertical components will give background. A pattern corresponding to vertical components can only be observed with focus variation if focus steps of the order of the electron's wavelength are available in an instrument.

90˚ between diffracted wave and the primary incident wave has origin in Fresnel-Huygens construction. We can understand the physical picture of the origin of the 90˚ phase shift considering diffraction or interference geometry given in Figure 4.05. The first figure showing the example diffracted wave vector k_d propagating at some angle θ with respect to the primary wave vector k_0. The second illustration shows interference pattern formation when the diffracted and primary waves meet due to focusing on the image plane. The horizontal component of diffracted

wave vectors will interfere according to the off-axis holographic principle and produce an interference pattern in the image plane. The *z* components will produce an interference pattern that can be observed only by changing focus and monitoring intensity for a fixed location on the image plane. As the carrier frequency of the interference fringe depends on $k_d \cos(90-\theta)$ and $k_d \sin(90-\theta)$ on in-plane and along z directions, a small angle will produce a typically fringe wavelength of the order of 10 Å, for a wavelength of electron 2 *pm* and scattering angle 0.1 degrees. The fringe wavelength along the z-direction will be ~ 2 *pm*. Therefore, it will be difficult to see the pattern with focus change along z-direction unless instrument focus settings have the required step resolution and only contribute to uniform background. Thus, the waves contributing to two different interference patterns have 90˚ angular separation and can be considered a 90˚ phase shift between diffracted and primary waves and the resulting flux distribution.

The physical picture of self-interference of single-electron wave function involves interaction with the object potential [20]. The interaction is mostly elastic due to fast probe electron, and a small probability of inelastic interaction is useful for the analytical techniques. The primary difference between Eq. 4.07 and 4.44 is $2\pi Qx$, which appears in the trigonometric function of off-axis electron holography due to wave interference at an angle (sec. 4.3.2). This particular phase term oscillates with the spatial coordinate *x* over the field of the electron hologram. The phase term $\phi(x,y)$ corresponding to OEW phase is acquired by another half of the wave while passing through the sample. This OEW phase manifests as holographic fringe bending in the image plane, e.g., relative fringe bending between vacuum and MgO crystal and differently striped thickness of object structure, that allows to determine the mean inner potential (MIP) at medium resolution [Figure 4.06] [4,34].

Another difference between the two holography techniques is the twin image. It is well known that for in-line holography, separation of twin images is an issue, whereas, for off-axis geometry, they get separated at $\pm Q$ in the frequency space. However, in the image plane, they superimpose on top of each other for both techniques. In the standard practice of OEW reconstruction, the Eq. 4.07 and 4.44 are generally written in an intermediate state (see Eq.4.04 for in-line and Eq. 4.29 for off-axis), having both OEW function ψ and its conjugate ψ^* inconsistent with the holographic principle of image formation [21]. The motivation behind this is based on the concept that the twin image components can be separated from the direct component (DC) if allowed to propagate along the scattering direction, i.e., from the image plane to the diffraction

plane and emphasizing the direct retrieval of the wave function. However, it is argued here that this does not pose an issue if one wishes to work in the image plane of the intensity pattern to retrieve both the phase and amplitude of OEW directly.

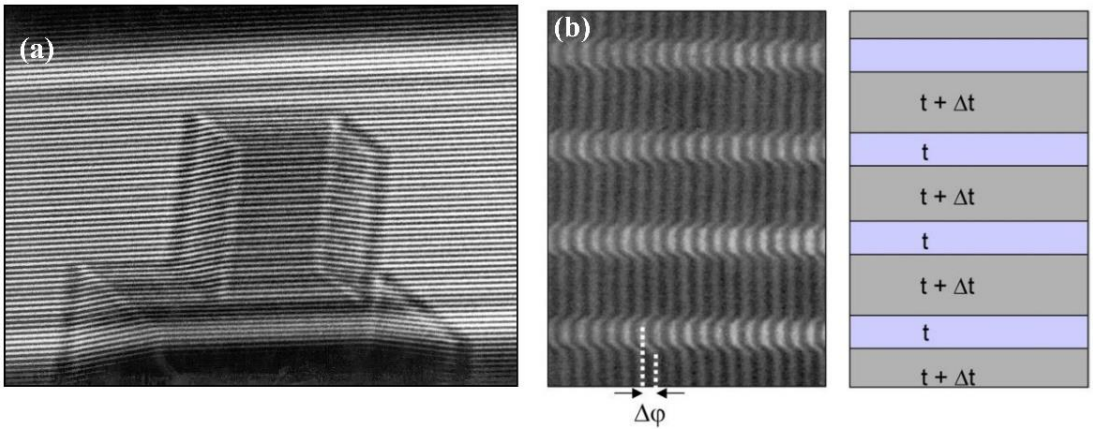


Figure 4.06. Off-axis electron holography fringe bending after encountering object potential in the image plane, (a) between vacuum and MgO crystal. Copyright (2002) by Elsevier [35]. (b) Object structured with different thickness stripes. Copyright (2007) by Institute of Physics publishing group [4].

From Eq. 4.43 and 4.44, it is evident that the phase information is preserved in the final form of the intensity equation responsible for the intensity modulation as described by the trigonometric function. Retaining the phase term is inherent to self-interference and resulting modulation in intensity pattern. It is similar to the off-axis electron holography fringe bending due to object potential in the image plane. Simultaneously, no information on the wave functions or twin images and DC components are available. The final form of intensity expression instead of the wave function-based intermediate state of the equation can be guided by the Born rule in Quantum Mechanics. The Born rule states that it is the probability density $\psi\psi^*$ the square of the probability amplitude of particle's wave function ψ that is real and observable quantity during the measurement and not the associated complex wave function which acts as a state vector [36,37]. There are controversies in the literature on measuring such state vectors, which is a complex quantity but can be constructed based on probability density [38–40]. The Born rule is an important link between the abstract mathematical formalism of quantum theory based on the complex wave function and the experimental measurement. It constitutes an essential part of the Copenhagen interpretation of quantum mechanics [41]. Therefore, the final form of the intensity equation is

equivalent to measuring the total probability density due to contribution from all the three components of wave functions, i.e., DC part, wanted wave function ψ and its complex conjugate ψ^*. And in the process of recording, the information on the wave functions is lost, recovering of which is emphasized in all the reconstruction approaches through the intermediate state of the equations.

Retaining the phase term is inherent to self-interference and resulting modulation in intensity pattern. It is similar to the off-axis electron holography fringe bending due to object potential in the image plane. Evaluating the local phase information from real space images in fringe bending primarily comes from spatial resolution. It will be addressed in the forthcoming discussion, along with additional aspects already mentioned earlier. In standard practice, the image is Fourier transformed, and then a digital aperture function is used in that plane that takes care of the resolution. The use of intermediate state of the equation to disperse the twin images and DC part in the diffraction plane made the earlier methods elaborate and complex for HRTEM reconstruction where twin images overlap on the same reciprocal space and required to enhance the weight of ψ over the ψ^* and $\psi\psi^*$ terms through summation over many images. Though the information on the wave functions is lost, however, the phase term is preserved in the final state of the equation, which is the same for both ψ and ψ^* but having an opposite sign.

The twin images are exact copies of each other and superimpose at the Gaussian image plane and do not cause any loss in information in terms of phase and amplitude. This can be understood with off-axis electron holography, where the inverse FT of each sideband preserves the same intensity pattern. In in-line holography, the twin images propagate along the opposite direction with defocus similar to the parabola picture of OEW reconstruction. Complex conjugate pairs lie on a reflected parabola with respect to the diffraction plane and focus variation. This picture is reminiscent of the transactional interpretation of quantum mechanics. The wave function and its complex conjugate experience phase change in the opposite directions equivalent to the forward and backward propagation in time [42,43].

Now proceeding for the experimental reconstruction of the OEW function, the phase can be determined directly by applying Eq. 4.43. Suppose the quantity inside the arcsine function is more than 1, then this should be divided by 1. The computer program should read the quotient as phase jump by the amount (quotient × $\pi/2$) plus the phase corresponding to the remainder. Then,

calculate the total phase change and corresponding counts of the atoms accordingly. This problem does not appear for one-layer thick MoS_2 and BN but is encountered in 3D crystals like ZnO, where thickness includes several atoms in the column. Next, the amplitude corresponding to ψ needs to be evaluated from the image intensity. The total amplitude of the recorded image is simply the square root of it but not equal to the amplitude of the single OEW function. As already mentioned, the total amplitude in the image intensity pattern has three components, direct component (DC), and two twin image components for off-axis and DC and only diffracted components in in-line holography. In the case of in-line holography, the twin image components overlap with the DC part. The DC component has approximately one order of magnitude higher total intensity than individual twin images from SBs and diffracted FT spots as measured directly from the FT of the images by placing a region of interest around various spots and SBs and evaluating the sum. Thus, the intensity of the OEW function is approximately given by the factor β for HRTEM, and this will be further divided into halves between two twin image components in case of off-axis electron holography. The amplitude of $\psi_{HRTEM} = \psi_0 + \psi_i$ is given by the square root of $I_{in\ line}(x, y)$ and only amplitude component A_i associated with trigonometric function is shown in reconstructed wavefunctions, where A_i and A_0 are the magnitudes of individual diffracted and reference complex wave functions, respectively.

The total intensity of CB is at least one order of magnitude higher than two SBs together [Figure 4.02]. In HRTEM, it is the same by considering all the FFT spots with respect to the direct spot at the center. Therefore, approximately 1% of the mean vacuum intensity is assigned to the image wave function (Eq. 4.43). Now from Eq. 4.43, after removing all the constant background terms and then a division by $2\sqrt{\alpha I_0}$ term retain only $\sqrt{\beta I_0}$ a term associated with the trigonometric function will give the amplitude of the object exit wave function (OEW). In the case of off-axis electron holography, it is given by $2a(x, y)$ divide over two for one of the SBs (Eq. 4.44). Individual $A_0 \exp(i\phi_0)$ or $A_i \exp(i\phi_i)$ wave function components will not give any intensity pattern, only the phase change and associated overall amplitude $2A_0A_i$ associated with the cosine term of $\psi_{HRTEM}\psi^*_{HRTEM}$ can be known from the interference experiment. Kindly note that the subtraction procedure to adjust the intensity value is automatically taken care of by Eq. 4.43 while calculating the phase. Therefore, it is possible to reconstruct the OEW function from the information obtained in terms of phase and amplitude from the measurement described above.

We have analyzed selected experimental images of MoS_2, BN, and ZnO using Eq. 4.43, and the reconstructed phase and amplitude images are presented in Figure 4.07 and 4.08 for MoS_2 and BN respectively. One can see from the phase images that the vacuum phase is zero, and phase values are high and low depending on the brightness of the dots, which represents the periodic arrangement of atoms in the lattice. Theoretical peak phase shift values extracted from Ref. [44] for the resolution range 1-0.5 Å are given in Table 4.01 for the atoms of interest in the present investigation. According to this table, the peak phase shift $\varphi_{max}(rad)$ follows ~ $Z^{0.6} - Z^{0.7}$, where *Z* is the atomic number [44]. The exponent of Z increases with increasing spatial resolution results in a higher peak phase shift value for a given atomic number.

From the present reconstruction method, a peak phase shift value of ~ 0.51 and 0.45 rad are obtained from the Mo and S atom positions, respectively, throughout the image, which are slightly higher than the theoretical values. Kindly note that for MoS_2, there are two S atoms along with the [0001] projection. Therefore, the phase shift values are almost double compared to the single S theoretical phase shift value. The values are 0.07 and 0.10 rad for B and N atoms, respectively, matching the theoretical estimation. Results on ZnO and bi-layer MoS_2 are given in Figures 4.09 and 4.10. As the sample thickness increases from the specimen's vacuum edge for ZnO thin film, a systematic increase in phase and atoms numbers can be observed.

Atoms (Z)	**Phase shift (rad)**
B (5)	0.076-0.078
N (7)	0.100-0.102
O (8)	0.106-0.108
S (16)	0.164-0.23
Zn (30)	0.242-0.41
Mo (42)	0.300-0.51

Table 4.01. Theoretical peak phase shift extracted for resolution in the range 1-0.5 Å [44].

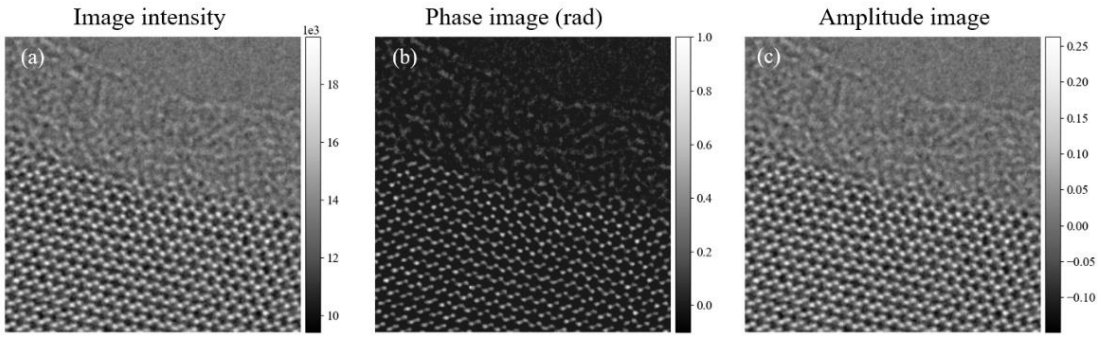


Figure 4.07. (a) HRTEM image of MoS_2 recorded under negative $C_S = -35$ µm and positive defocus of $\Delta f = 8$ nm. (b) phase and (c) amplitude of the OEW. A peak phase value of ~ 0.51 and 0.45 rad are obtained for Mo and two S atoms, respectively.

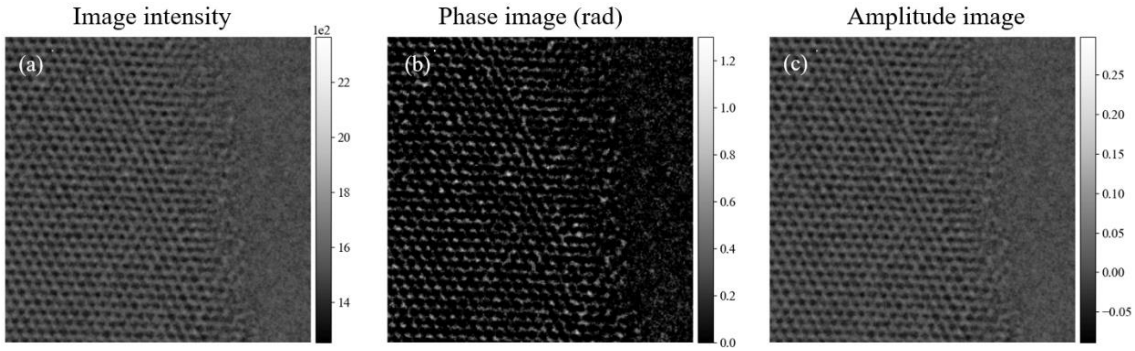


Figure 4.08. (a) HRTEM image of BN monolayer recorded under negative $C_S = -35$ µm and positive defocus of $\Delta f = 8$ nm. (b) phase and (c) amplitude of the OEW. A peak phase value of ~ 0.07 and 0.10 rad are obtained for B and N atoms, respectively.

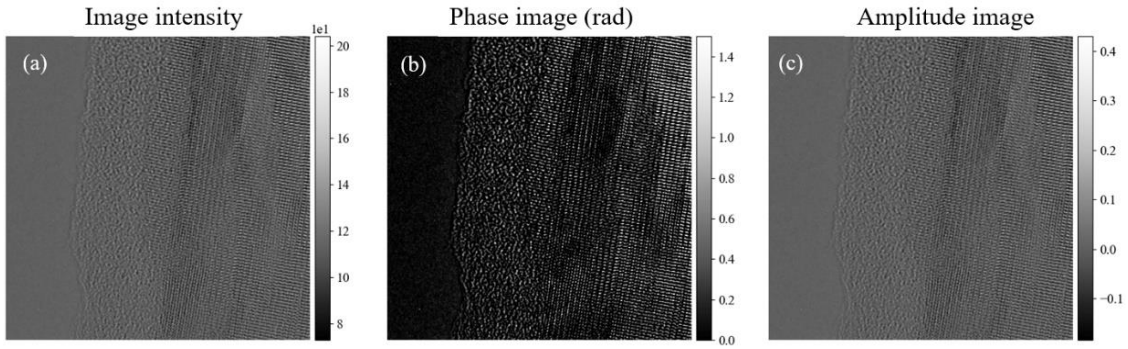


Figure 4.09. (a) HRTEM image of ZnO epitaxial thin film monolayer recorded under $C_S = -35$ µm and positive defocus of $\Delta f = 8$ nm. (b) phase and (c) amplitude images of the OEW. Notice the systematic increase in phase and associated atom number for Zn and O atoms away from the sample edge with the vacuum.

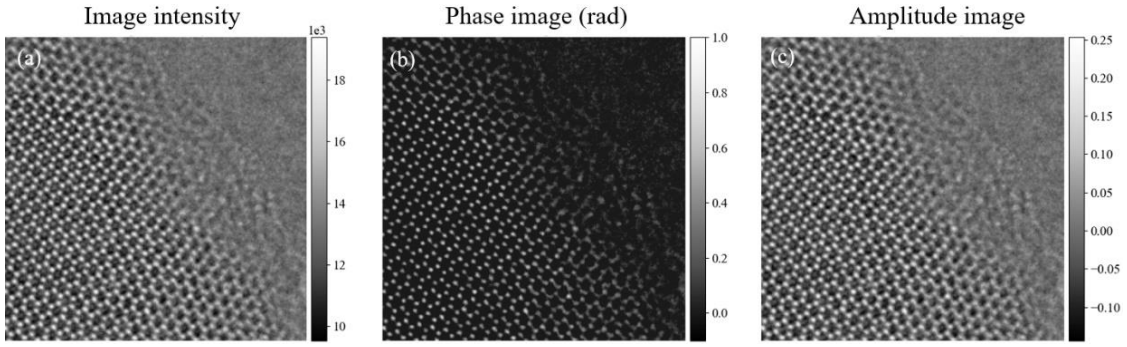


Figure 4.10. (a) Intensity image of bilayer MoS_2 under $C_S = -35$ µm and positive defocus of $\Delta f = 8$ nm. Reconstructed (b) phase and (c) amplitude images of OEW.

The method described above, using Eq. 4.43, translates the intensity information into a phase value. And an excellent trend is obtained for images with low and high Z compounds giving lower and higher phase shift values, respectively, without additional data refinement procedures. Reference vacuum wave is not required as one can approximate the mean of atomic resolution image intensity as mean vacuum intensity. Fresnel fringe, electron interference fringe in off-axis electron holography should be removed from the image data before applying Eq. 4.43 and are generally removed before any analysis in case of off-axis electron holography.

4.4.2 Recovering image function from the diffracted intensity

In this section, the recovery of image function containing information on both crystallographic and object phases is described if the information is available only in the diffraction plane. The real image function $f(x)$ of both symmetric non-periodic and periodic in particular forms, e.g., Gaussian, reciprocal, etc., can be retrieved by cosine Fourier series expansion of the absolute FFT of such functions followed by summation over all frequencies. The absolute FFT and Fraunhofer patterns are equivalent to each other. The procedure is similar to the zero-phase reconstruction using the Patterson function in X-ray crystallography applicable to smaller-sized molecules [24,25]. Example Gaussian function in 1D both in the isolated and periodic form is given in Figure 4.11.

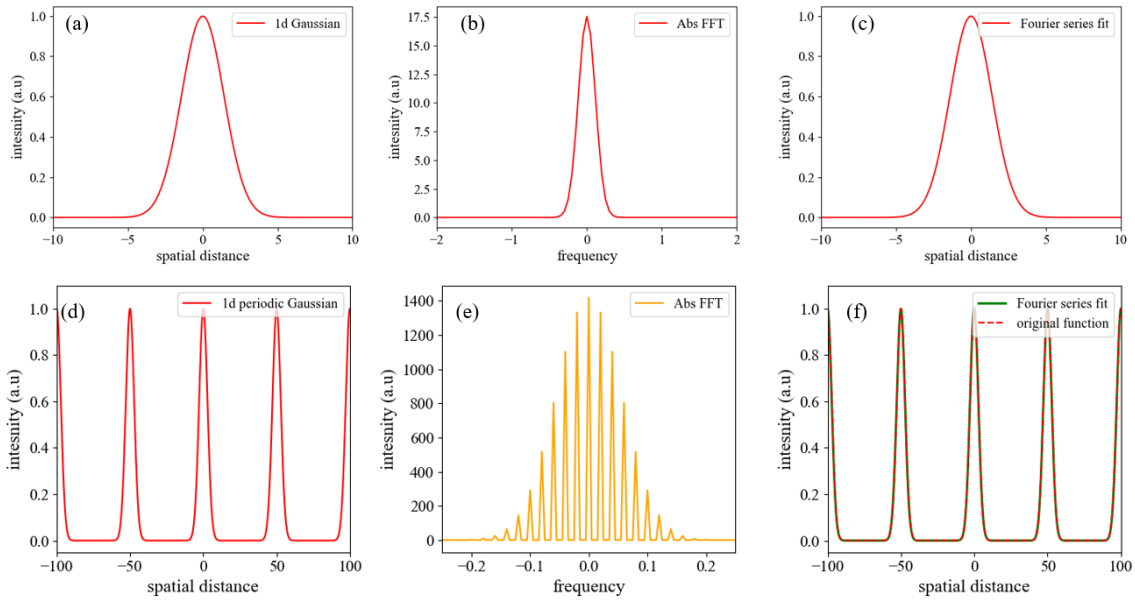


Figure 4.11. (a) single Gaussian function in 1D, (b) absolute FFT of the function, and (c) Fourier series expansion followed by summation over all the frequencies (d) Periodic Gaussian function in 1D, (e) absolute FFT of the function, and (f) Fourier series expansion followed by summation over all the frequencies.

The unique feature underlying this reconstruction method's workings is that for most functions, the integration of imaginary part over spatial parameter or abs-*Im* component is negligible compared to real part or abs-*Re* [Figure 4.12]. Therefore, the original function $f(x)$ can be retrieved entirely by the following real cosine Fourier series alone. If $F(k)$ is the absolute FFT of $f(x)$, then $f(x)$ can be retrieved from the $F(k)$ via following real Fourier series expansion

$$f(x) = absolute\left(\sum_{n=-k}^{n=k} \frac{1}{n} C_k \cos(2\pi k x)\right) \tag{4.45}$$

Where n = number of data points in the frequency axis, k is the frequency, C_k= absolute of Fourier transformations or diffracted intensity at some frequency k and x is the 1D spatial coordinate over which real image function will be defined.

For periodic function in 1D, $f(x) = f(x \pm px)$, where p =0, 1, 2, …, one need to ensure that the range of n should be $= \pm \frac{2p}{M}$ where M = number of data points both in real and diffraction space. In the case of a 2D isolated function, Eq. 4.45 will have another summation over n_y for the

second orthogonal axis. Figure 4.13 shows example reconstruction for 1D reciprocal ((a)-(c)) and Gaussian function ((d)-(f), respectively.

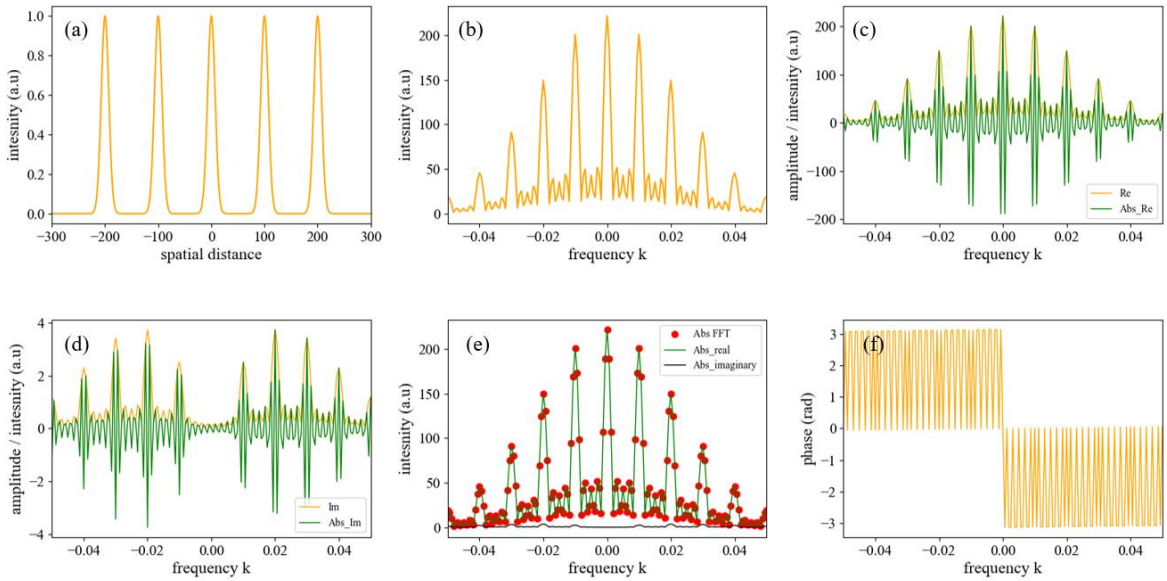


Figure 4.12. (a) Example gaussian function, (b) Absolute FFT, absolute of (c) Re, and (d) Imaginary parts. (e) contribution to absolute FFT is mostly from the absolute Re part, and contribution from the absolute imaginary part is negligible. (f) The phase of the Fourier wave.

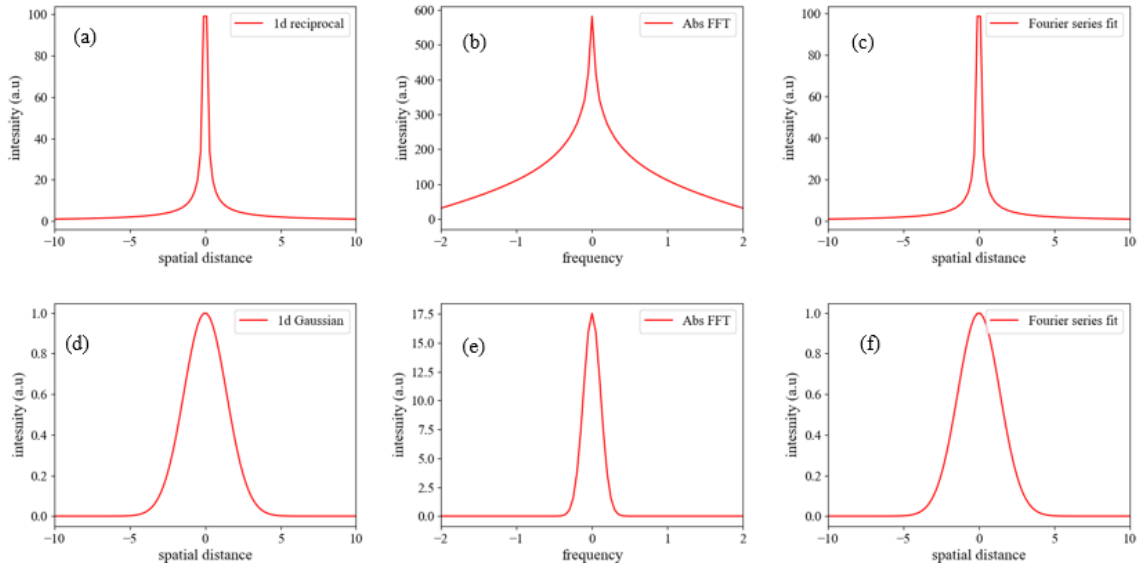


Figure 4.13. Example reconstruction for (a)-(c) reciprocal and (d)-(f) Gaussian function in 1D form. (a) A reciprocal function and (b) it's abs_FT and (c) the reconstructed function. (d) An example Gaussian function with it's abs_FT in (e), and (f) the reconstructed function.

For isolated 2D function following Eq. 4.45 is modified to

$$f(x,y) = abs\left(\sum_{n=-k}^{n=k} \frac{1}{n} C_{k_x k_y} \cos\left(2\pi k_{k_x k_y} r\right)\right) \tag{4.46}$$

Where, $r = \sqrt{x^2 + y^2}$. The 2D reconstruction for Gaussian and reciprocal functions is given in Figures 4.14 and 4.15, respectively, with corresponding 1D line profiles.

Modified Fourier series equation for a periodic 2D function is given in Eq. 4.47. The functional form of cosine changes depending on the isolated 2D and periodic 2D functions to get an exact fit.

$$f(x,y) = \sum_{n_{i,j}=k_{i,j}} \frac{1}{n_{i,j} \times n_{i,j}} C(k)_{i,j} \frac{1}{2} [\cos(2\pi k_{i,j} x) + \cos(2\pi k_{i,j} y)] + \frac{1}{(n_{i,j} \times n_{i,j})} \frac{1}{4} C(k)_{i,j} [\cos(2\pi k_{i,j}(x+y)) + \cos(2\pi k_{i,j}(x-y))] \tag{4.47}$$

One needs to ensure to have more sampling points to have an exact fit. The example for the periodic Gaussian function in Figure 4.16 used 400×400 sampling points.

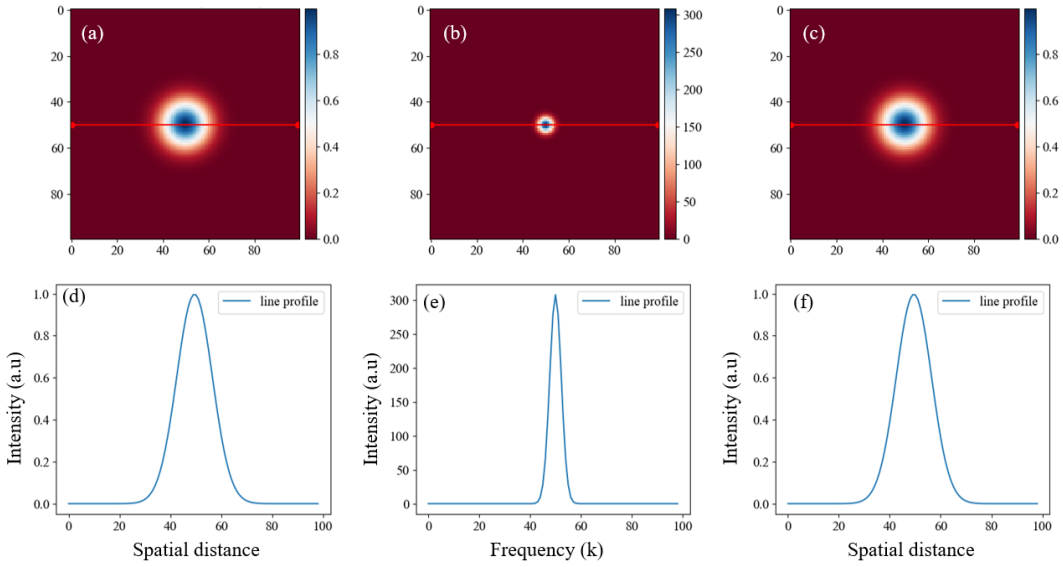


Figure 4.14. Example reconstruction for Gaussian function in the 2-dimensional form with corresponding line profiles. (a) Gaussian function in 2D, (b) absolute Fourier transform and (c) Fourier series fit from the absolute FT, using Eq. 4.46 and the corresponding line profile (d), (e), and (f), respectively.

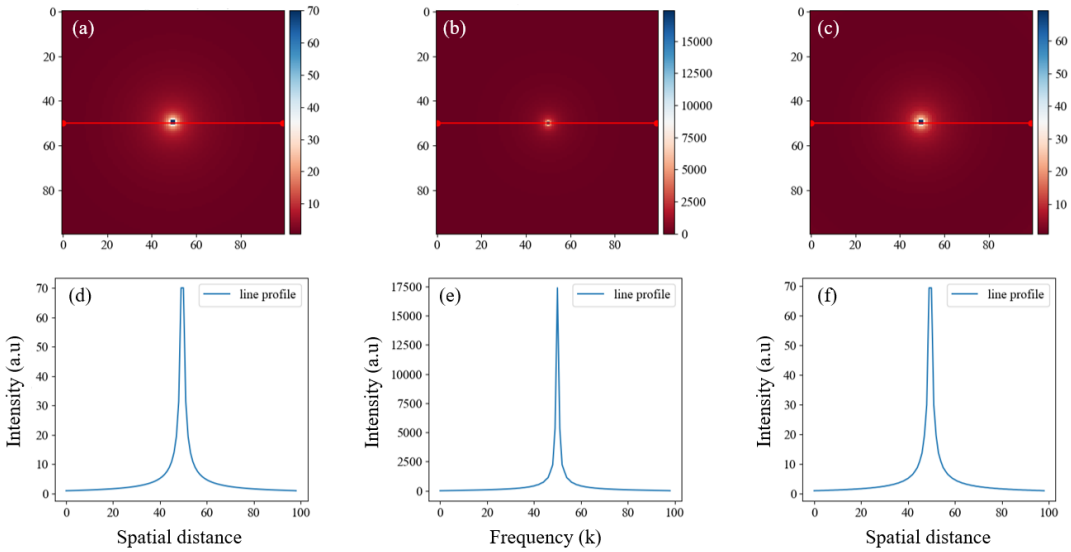


Figure 4.15. Example reconstruction for reciprocal function in the 2-dimensional form with corresponding line profiles. (a) Reciprocal function in 2D, (b) absolute Fourier transform and (c) Fourier series fit from the absolute FT, using Eq. 4.46 and the corresponding line profile (d), (e), and (f), respectively.

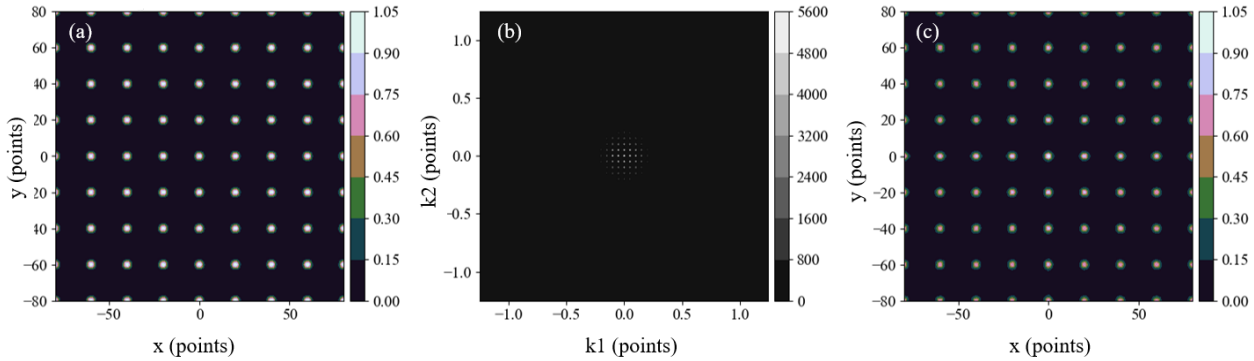


Figure 4.16. Example reconstruction for periodic Gaussian function in the 2-dimensional form. (a) A 2D periodic gaussian function, (b) absolute Fourier transform, and (c) Fourier series fit from the absolute FT, using Eq. 4.47.

The method described above works well for both isolated and periodic functions. From the image intensity function, one can then apply the first method to retrieve the OEW and associated object phase related to potential. However, suppose there is any inhomogeneity in the distribution of periodic function, e.g., dopants, edges, interfaces, etc. In that case, the accompanying

information will be required from the image intensity pattern the method described in the previous section works suitably.

4.5 Conclusion

In conclusion, alternative reconstruction methods to retrieve the phase and amplitude of OEW in HRTEM imaging are introduced. The first method is based on directly applying a modified HRTEM intensity equation to retrieve the phase of OEW from the image. The reconstruction results in terms of peak phase shift values are in excellent agreement with the theoretical estimation. The second method described is applicable for retrieving the image intensity function from the information available in the diffraction plane for both isolated and periodic functions.

4.6 Bibliography

[1] W. Coene, G. Janssen, M. Op De Beeck, and D. Van Dyck, Phys. Rev. Lett. **69**, 3743 (1992).

[2] D. Van Dyck and M. Op de Beeck, Proceedings of the XIIth International Congress on Electron Microscopy, pp. 26–27 (1990).

[3] W. O. Saxton, Ultramicroscopy. **55,** 171 (1994).

[4] H. Lichte and M. Lehmann, Reports Prog. Phys. **71**, 016102 (2008).

[5] D. Geiger, H. Lichte, M. Linck, and M. Lehmann, Microsc. Microanal. **14**, 68 (2008).

[6] U. Bhat and R. Datta, J. Appl. Phys. **125**, 154902 (2019).

[7] D. Van Dyck, J. R. Jinschek, and F. R. Chen, Nature **486**, 243 (2012).

[8] C. Jin, F. Lin, K. Suenaga, and S. Iijima, Phys. Rev. Lett. **102**, 195505 (2009).

[9] C. L. Jia, S. B. Mi, J. Barthel, D. W. Wang, R. E. Dunin-Borkowski, K. W. Urban, and A. Thust, Nat. Mater. **13**, 1044 (2014).

[10] C. L. Jia, M. Lentzen, and K. Urban, Science **299**, 870 (2003).

[11] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Großß, A. Chuvilin, G.

Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser, Nat. Mater. **10**, 209 (2011).

[12] C. L. Jia, S. B. Mi, K. Urban, I. Vrejoiu, M. Alexe, and D. Hesse, Nat. Mater. **7**, 57 (2008).

[13] P. A. Midgley and R. E. Dunin-Borkowski, Nat. Mater. **8**, 271 (2009).

[14] R. R. Meyer, A. I. Kirkland, and W. O. Saxton, Ultramicroscopy **92**, 89 (2002).

[15] C. B. Carter and D. B. Williams, Transmission Electron Microscopy: Diffraction, Imaging, and Spectrometry (Springer, 2016).

[16] F. Genz, T. Niermann, B. Buijsse, B. Freitag, and M. Lehmann, Ultramicroscopy **147**, 33 (2014).

[17] T. Niermann and M. Lehmann, J. Phys. D. Appl. Phys. **49**, 194002 (2016).

[18] K. Ishizuka and B. Allman, J. Electron Microsc. (Tokyo). **54**, 191 (2005).

[19] J. C. H. Spence, *High-Resolution Electron Microscopy*, Third Edit (Oxford Science Publications, 2003).

[20] B. Fultz and J. M. Howe, *Transmission Electron Microscopy and Diffractometry of Materials* (Springer Science & Business Media, 2012).

[21] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University Press, 1999).

[22] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science*, 2nd ed. (Springer, 2009).

[23] E. J. Kirkland, *Advanced Computing in Electron Microscopy: Second Edition*, 2nd ed. (Springer, 2010).

[24] A. L. Patterson, Phys. Rev. **46**, 372 (1934).

[25] J. Rius, IUCrJ **1**, 291 (2014).

[26] R. Sahu, U. Bhat, N. M. Batra, H. Sharona, B. Vishal, S. Sarkar, S. Assa Aravindh, S. C. Peter, I. S. Roqan, P. M. F. J. Costa, and R. Datta, J. Appl. Phys. **121**, 105101 (2017).

[27] K. Dileep, R. Sahu, S. Sarkar, S. C. Peter, and R. Datta, J. Appl. Phys. **119**, 114309 (2016).

[28] R. Datta, S. Kanuri, S. V. Karthik, D. Mazumdar, J. X. Ma, and A. Gupta, Appl. Phys. Lett. **97**, 071907 (2010).

[29] B. Loukya, P. Sowjanya, K. Dileep, R. Shipra, S. Kanuri, L. S. Panchakarla, and R. Datta, J. Cryst. Growth **329**, 20 (2011).

[30] P. Schiske, J. Microsc. **207**, 154 (2002).

[31] J. Frank, Optik (Stuttg). **38**, 519 (1973).

[32] K. Ishizuka, Ultramicroscopy **5**, 55 (1980).

[33] J. Loudon, *Advanced Transmission Electron Microscopy Lecture 2: Electron Holography*, (unpublished).

[34] M. Linck, B. Freitag, S. Kujawa, M. Lehmann, and T. Niermann, Ultramicroscopy **116**, 13 (2012).

[35] H. Lichte, and M. Lehmann. Advances in imaging and electron physics **123**, 225 (2002).

[36] M. Born, Zur Zeitschrift Für Phys. **37**, 863 (1926).

[37] M. Born, Science **122**, 675 (1955).

[38] J. M. Zuo, M. Kim, M. O'Keefe, and J. C. H. Spence, Nature **401**, 49 (1999).

[39] C. J. Humphreys, Nature **401**, 21 (1999).

[40] B. Q. Pham and M. S. Gordon, J. Phys. Chem. A **121**, 4851 (2017).

[41] S. Siddiqui and C. Singh, Eur. J. Phys. **38**, (2017).

[42] R. E. Kastner, *The New Transactional Interpretation of Quantum Mechanics : The Reality of Possibility* (Cambridge University Press, 2012).

[43] J. G. Cramer, *Quantum Nonlocality and the Possibility of Superluminal Effects*, Propuls. Phys. Work. Proc. 1 (1997).

[44] M. Linck, H. Lichte, and M. Lehmann, Int. J. Mat. Res. **97**, 890 (2006).

Chapter 5

Summary and Future perspective

This chapter outlines the important conclusions drawn from the present thesis. Our contributions towards quantitative imaging are summarized. Quantitative counting of Zn and O atoms in ZnO thin film is carried out by different routes, and the results are compared. Atomic potential models and phase-detection limits for off-axis electron holography and HRTEM are discussed. Comparative analysis of existing image simulation as well as reconstruction methods has been carried out. A new image simulation method in HRTEM based on the atom as an electrostatic interferometer similar to the off-axis electron holography biprism is introduced. Alternative methods of phase retrieval in HRTEM have been introduced based on real space and reciprocal space. The subsequent section discusses the promising future of quantitative imaging in the context of recent development.

5.1 Summary of the thesis

The thesis contains three aspects: i) Quantitative atom counting of Zn and O atoms by different routes and comparing the results. ii) Alternative image simulation method in high-resolution transmission electron microscopy (HRTEM) and iii) new phase retrieval methods in HRTEM images.

1. Quantitative counting of Zn and O atoms is carried out by off-axis electron holography and HRTEM. In the off-axis electron holography, reconstruction of the images is carried out based on both central band (CB) and sideband (SB) [1].
2. Potential models and phase-detection limits in off-axis electron holography and high-resolution transmission electron microscopy (HRTEM) are discussed [1].
3. In off-axis electron holography phase retrieval, phase number increases with an increase in the sample thickness up to a certain distance. SB and CB yield similar results for less than extinction distance but above which they do not match. A similar number of Zn and O atoms are obtained based on their stoichiometry. A three times higher number of atoms is obtained from the mean phase shift than the peak phase shift. A poor match between CB and SB in a thicker region is due to dynamical scattering and different formalism in the reconstruction method [1].
4. In HRTEM reconstruction, no linear trend of phase shift with varying thickness is found. Throughout the sample area, a phase shift corresponding to a single number of Zn (0.138 rad) and O (0.09 rad) is obtained [1].
5. A new method of simulating atom image in HRTEM is introduced, considering the atom as an electrostatic charge center. Interference along the radial direction with different zones results in the image of an isolated atom [2].
6. Based on the new method, image intensity varies as a function of Z^b where b decreases with an increase in defocus. A range of 0.4 to 0.26 is obtained for b for defocus1 to 8 nm. Simulated intensities with consideration of aberration match the experimental data for Mo and B atoms [2].
7. Alternate methods of phase retrieval in HRTEM are proposed based both on real space and reciprocal space information. The first method is based on applying a modified HRTEM intensity equation to retrieve the phase of object exit wave (OEW) from the image, which contains only the background and pattern information due to the phase component but no direct

wave function information. The peak phase shift obtained by this method agrees well with the theoretical estimation. The beauty of this work is that with a single HRTEM, image reconstruction can be carried out without a tedious process of through focal series reconstruction [3].

8. The second method involves retrieving the image intensity function from the diffraction plane information for isolated and periodic arrangements [3].

5.2 Future perspectives

Quantitative imaging is a versatile area of research emerging with new and powerful techniques day by day. There is a spectrum of opportunities in quantitative imaging with aberration-corrected microscopes and better detection capabilities.

Valance electron information is essential to understand the bonding and deduce many more interesting properties of the material. Because of the strong scattering from the core electrons, valence electrons cannot be measured or imaged easily. Meyer Jannik C. et al. [4] demonstrated an experimental analysis of charge redistribution due to chemical bonding by HRTEM and matched it with first-principles electronic structure calculation. Charge transfer on the single-atom level for nitrogen-substitution point defects in graphene was studied, and iconicity of the single-layer hexagonal boron nitride was confirmed. Figure 5.01 shows charge distribution, projected potentials, and TEM simulations used to study nitrogen-doped graphene [4].

There are several approaches to understand the potential distribution around the atom by examining the phase change. Off-axis electron holography is one of the most used techniques at atomic resolution. The phase change is incorporated in the hologram as fringe bending. The fringes between the atomic columns in the holograms give a unique opportunity to extract valence information with the atomic columns. In a study by Linck et al. [5], the use of a new high-brightness Schottky field electron emitter in the holography experiment enhances the phase signal resolution significantly. From the study of grain boundary in the gold, enhancement of information by a factor of 2.88 is reported compared to previous results. Other quantitative studies of charge density in monolayer MoS_2 and potential mapping of graphene sheets are based on atomic resolution off-axis electron holography [6,7].

Quantitative imaging helps to analyze the material at atomic precision. There is an example of a three-dimension reconstruction of the material from a single HRTEM image [8]. Occupancy of the atoms present, defects present, and the position of a particular type of atom in the atomic column of MgO crystal are reconstructed in this study. Quantitative imaging helps to map the electron phase to the potential. One such attempt towards this is given in Ref. [9]. Experimental off-axis electron holography and simulated images of a few-layer WSe_2 are compared. Electron-optical phase images are in good agreement with the electrostatic potential-based calculations, which include the bonding effects.

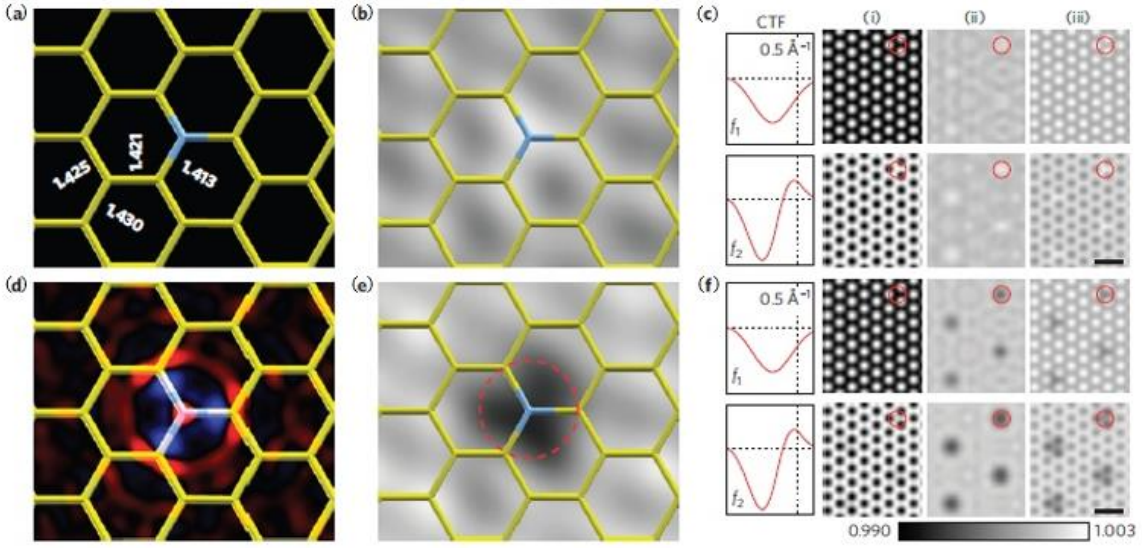


Figure 5.01. (a) Relaxed atomic configuration for a nitrogen substitution in graphene. Bond lengths are given in angstroms. (b) Projected potential based on the independent atomic model (IAM). Dark contrast corresponds to higher projected potential values, in accordance with TEM imaging conditions. (c) TEM simulation based on the IAM potential for two experimental conditions f_1 and f_2 (for more details, see Ref. [4]). Filters are: (i) unfiltered, (ii) periodic components removed by a Fourier filter, and (iii) low-pass filtered. (d) Atomic structure (same bond lengths), with the projected electron density changes due to bonding shown in color. Blue corresponds to a lower red to a higher electron density in the DFT result compared with the neutral-atom (IAM) case. (e) Projected potential, filtered as in (b), based on the all-electron DFT calculation. (f) TEM simulations using the DFT-based potentials. The greyscale calibration bar applies to columns (ii) and (iii), all shown on the same greyscale range for direct comparison. The scale bars are 5 Å. Copyright (2011) by Nature publishing group [4].

There are only a few examples that show that quantitative imaging offers more than material characterization. It aims to detect the position of defects and dopants in the atomic columns [Figure 5.02 (a)] [10]. Gathering valence electron information and imaging them is a highly challenging task in transmission electron microscopy [Figure 5.02 (b)] [11]. But with better detectors, it does not seem so farfetched as it did a few decades ago.

Improvement in instrument and detection capacities of the microscope opens direct detection (K3 camera) of individual electrons [12]. This next-generation camera is optimized for low-dose applications in life science and material science. It offers a maximum signal-to-noise ratio without damaging the specimen [Figure 5.02 (c)] [12]. It facilitates high-quality signals with minimum Poisson noise and provides the best resolution and contrast. K3 is the highest performance detector available for cryo-electron microscopy. Unlike other detectors, it has a detection quantum efficiency beyond 80%. It counts individual electrons rather than the traditional integration of the analog signal. It has a field view of 24 megapixels (5760×4092). It offers a high recording speed of 1500 full frames per second. With better signal quality, the study of material properties will be better with atomic-level precision.

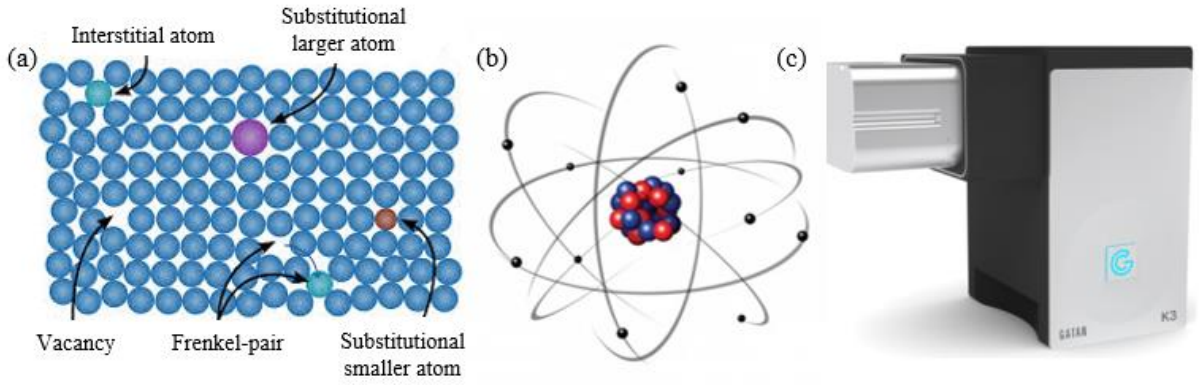


Figure 5.02. (a) Schematic of a crystal lattice with different types of defects [10]. (b) Atomic model [11] and (c) K3, direct detection camera [12].

Quantitative imaging at atomic resolution offers opportunities not only in materials study but also fundamental science from a sub-atomic length scale. With advanced techniques and instrumentation, the area is evolving and opening a new horizon. Eventually, it is to understand and extract the valence electron information, potential, and field around it. With advanced and emerging

methods and instrumentation, the future is to unravel the details which material holds. Ultimately quantitative imaging aims to experimentally obtain the information, which so far is available only by the density functional theory or the simulation.

5.3 Bibliography

[1] U. Bhat and R. Datta, J. Appl. Phys. **125**, 154902 (2019).

[2] U. Bhat and R. Datta, J. Phys. Commun. https://doi.org/10.1088/2399-6528/ac1839 (2021).

[3] U. Bhat and R. Datta, arXiv:2103.00538 (2021).

[4] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Großß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser, Nat. Mater. **10**, 209 (2011).

[5] M. Linck, B. Freitag, S. Kujawa, M. Lehmann, and T. Niermann, Ultramicroscopy **116**, 13 (2012).

[6] V. Boureau, B. Sklenard, R. McLeod, D. Ovchinnikov, D. Dumcenco, A. Kis, and D. Cooper, ACS Nano **14**, 524 (2020).

[7] D. Cooper, C. T. Pan, and S. Haigh, J. Appl. Phys. **115**, 233709 (2014).

[8] C. L. Jia, S. B. Mi, J. Barthel, D. W. Wang, R. E. Dunin-Borkowski, K. W. Urban, and A. Thust, Nat. Mater. **13**, 1044 (2014).

[9] S. Borghardt, F. Winkler, Z. Zanolli, M. J. Verstraete, J. Barthel, A. H. Tavabi, R. E. Dunin-Borkowski, and B. E. Kardynal, Phys. Rev. Lett. **118**, 086101 (2017).

[10] http://sarahlambart.com/teaching/mineralogy-05.pdf

[11] https://www.scmp.com/lifestyle/technology/article/1251477/bohrs-atomic-model-100- years-old

[12] http://www.gatan.com/k3-camera.