Prospects for versatile phase manipulation in the TEM: Beyond aberration correction

Giulio Guzzinati, Laura Clark, Armand Béché, Roeland Juchtmans, Ruben Van Boxem, Michael Mazilu, Jo Verbeeck

EMAT, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium
SUPA, School of Physics and Astronomy, University of St Andrews, St Andrews KY16 9SS, UK

ABSTRACT

In this paper we explore the desirability of a transmission electron microscope in which the phase of the electron wave can be freely controlled. We discuss different existing methods to manipulate the phase of the electron wave and their limitations. We show how with the help of current techniques the electron wave can already be crafted into specific classes of waves each having their own peculiar properties. Assuming a versatile phase modulation device is feasible, we explore possible benefits and methods that could come into existence borrowing from light optics where the so-called spatial light modulators provide programmable phase plates for quite some time now. We demonstrate that a fully controllable phase plate building on Harald Rose's legacy in aberration correction and electron optics in general would open an exciting field of research and applications.

1. Introduction

When Harald Rose started to work on the implementation of the spherical aberration (Cs) corrector he was trying to overcome the intrinsic limits of the simple circularly symmetric electron optical lenses. Scherzer himself had realized that using multipolar lenses could have allowed to break free of the restrictions imposed by the Scherzer theorem [1–3]. However, due to the formidable challenges posed by its realization, this concept was only successfully implemented 50 years later by the Haider–Rose–Urban project [4–6]. The goal of the Cs corrector however, despite all its flexibility, is to “flatten” the phase imposed by the optical systems as much as possible in order to make the smallest possible probe (in STEM) and to collect and interpret the highest possible frequencies (in TEM). This has, of course, had a groundbreaking impact on the field, pushing the resolution of both TEM and STEM beyond the angstrom limit, and making it possible to reach the very high current densities that have allowed analytical techniques to reach atomic resolution.

We would like to peek into the territory of phase manipulation. Optical microscopy techniques have greatly benefited from the invention of the spatial light modulator, a device which allows to tune a light wave in both phase and amplitude. A corresponding “Spatial Electron Modulator” however is far from being available to the TEM community. In this paper we will review the methods currently available for the manipulation of the electron wave, and outline possible applications.

2. Review: the state of the art

It is not yet possible to freely and arbitrarily tune the wave function of an electron beam. However there are a number of techniques which allow this, to some degree. Many of these have already proven useful, despite all their limitations, in order to expand the flexibility of the TEM or in the developing field of singular electron optics.

2.1. Electromagnetic fields in free space

An electron encountering a region of space containing an electric potential acquires a phase shift equal to

$$\Delta \Phi_e = \frac{\pi}{\lambda E_0} \int r \, V(r) \, dl.$$  (1)

Indeed electrostatic optical elements are widespread in lower energy applications, ranging from scanning electron microscopes, low energy or photoemission electron microscopes to Mott detectors, and electrostatic aberration correctors have been devised as well in the form of electron mirrors. While impractical for the...
manipulation of high energy electrons, tet rode mirror aberration correctors have been demonstrated to successfully correct spherical and chromatic aberrations in LEEM-PEEMs and SEMs [7–13]. Similarly, magnetic vector potentials also influence the phase of electron waves. When passing through a region of space with non-zero magnetic vector potential, an electron acquires an additional Aharonov–Bohm phase of
\[ \Delta \Phi_{AB} = \frac{e}{\hbar} \oint \mathbf{A} \cdot d\mathbf{l}. \] (2)
This phase is acquired even when the magnetic field \( \mathbf{B} = \nabla \times \mathbf{A} \) is zero. Magnetic lenses commonly employed in the TEM can also be explained in the light of this effect, as can multipolar aberration correctors.

The multipolar corrector aims to determine the phase plate impressed on the beam by manipulating the magnetic field by the boundary conditions at the ends of the magnetic poles. This is fundamentally possible because fields in free space are described by analytical functions, and as such an entire region of space can be entirely described by the properties of its boundary. While this is certainly true, the level of accuracy and stability required for this type of manipulation to be effective is extremely high, and was one of the factors that made the realization of aberration correctors so challenging.

The phase of an electron probe can also be altered by interacting with light. When a charged particle is immersed in a strongly varying electromagnetic field as found e.g. in laser cavities it is subjected to a ponderomotive force. For an electron passing through the focal point of a laser beam, perpendicular to the electron propagation, Müller et al. [14] calculated the phase shift to be
\[ \Delta \Phi_B \approx \frac{P \lambda}{2 \beta \gamma}. \] (3)
where \( P \) and \( \lambda \) are the local energy density and wavelength of the laser, \( \beta = v/c \) and \( \gamma = (1 - \beta^2)^{-1/2} \), the Lorentz factors for the relativistic electron. By focusing the laser in the back focal plane of the objectives lens a Zernike phase plate can in principle be obtained [15–17]. In this case the unscattered wave component is phase shifted by the strongly varying light intensity in the focal point of the laser while the scattered wave is left unchanged. Efforts are being made to implement this technique [18]. Note that also here, the phase plate is defined by the standing optical wave which is determined by its boundary conditions (the cavity) making it far from trivial to produce a fully versatile phase plate this way.

2.2. Phase manipulation using aberration correctors

The development and commercialization of aberration correction for the minimization of the Seidel aberrations has been absolutely fundamental in enabling reliable, consistent access to sub-Angstrom resolution in modern electron microscopes. The electric and magnetic multipoles (quad/oc topole or hexapole systems) of an aberration corrector can be individually addressed and tuned to counteract both the intrinsic spherical (and recently chromatic) aberration and the remaining parasitic aberrations. Typically, for an electron beam to be considered “corrected”, there should be a variation of no more than \( \pi/4 \) in the phase, within a given opening angle. This leads to a good approximation to a plane wavefront, which can be focused to produce a small diffraction limited probe for STEM, or an aberration-free lens which can correctly transfer high frequencies in TEM.

While manipulating the phase of the wavefronts using aberration correctors to closely approximate the ideal plane wave is very well suited to normal TEM and STEM studies, with the degrees of freedom given by the free-lens mode of an aberration corrected microscope, the electron wave can be designed and optimized for the experiment at hand, with structures available beyond simple plane waves, or Airy disc probes.

The aberration function in Saxton notation is
\[ \chi = \frac{2\pi}{\lambda} (\theta A_0 \cos (\phi - \phi_1) + \frac{1}{2} \theta^2 (A_1 \cos (2(\phi - \phi_2)) + C_1) + \frac{1}{3} \theta^3 (A_2 \cos (3(\phi - \phi_3)) + B_2 \cos (3\phi - \phi_3) + C_3) + \cdots) \]

where \( \chi \) is the phase shift of the wave front at \((\theta, \phi)\).

The possibilities offered by the free tuning of the aberrations have been studied previously for optimizing phase contrast in atomic resolution HRTEM. Lentzen et al. have described how to combine defocus, third and fifth order spherical aberrations in order to obtain positive or negative Zernike-type phase contrast by employing a controlled amount of positive or negative spherical aberration [19,20]. In particular, the negative spherical aberration imaging (NCSI) technique has shown a remarkable enhancement of the contrast for light elements such as oxygen or nitrogen, allowing their direct imaging, while previously their position had been only accessible by performing exit-wave reconstruction [21–23].

A further example of probe restructuring using aberration correctors was demonstrated recently by Clark et al. [24], wherein free manipulation of the aberrations was combined with an annular aperture to produce a high intensity vortex beam. Thus, it can be seen that by selecting a single \( \theta \) value with the annular aperture, and minimizing all aberrations other than the \( A_i \) (where \( A_0 \) is beam tilt and \( A_1 \cdots A_n \) are the \(+1\)-fold astigmatisms), we can create a phase linearly increasing with azimuthal angle \( \chi \approx m\phi \) typical of a so-called vortex beam using the appropriate weights for the cosine series. This phase structure, and the resulting

Fig. 1. Creation of an electron vortex through aberration manipulation. \( A_1 \) and \( A_2 \) were optimized. Higher orders of astigmatism were beyond experimental limits. The annulus has an average radius of 7 mrad. Both intensity images represent 2 nm by 2 nm. (a) Phase map imposed by the \( A_i \) to produce a vortex phase, with overlaid annular aperture. (b) Resulting simulated intensity pattern in the sample plane. (c) Experimental intensity pattern in the sample plane, obtained in an FEI Titan3 at 300 kV.
Theoretical and experimental intensity profiles are demonstrated in Fig. 1. Such vortex beams have the peculiar property that they possess a quantized angular momentum around their propagation axis. They have been shown to be useful to apply rotational forces to nanoparticles [25], measure exchange of angular momentum [26], identify magnetic states up to the atomic scale [27–30] and to determine chiral structures on a local scale [31].

Furthermore, we present here new results, demonstrating the flexibility of non-standard aberration corrector manipulation, to create an Airy beam. A wavefront defined as

$$\Psi(x, z) = \text{Ai} \left( \frac{x^2}{x_0^2} \right) \times \exp \left( i \left( zx_0^2 - i \frac{x^3}{12k^2x_0^4} \right) \right)$$

(5)

(where $x_0$ is a scaling parameter of the transverse modulation of the wave) is called an Airy wave, and has a number of unusual properties. Of these, the most interesting are that such a beam remains diffraction-free during propagation, and that it accelerates transversely to the optical axis, without the presence of external forces. However, as such a wave function is not normalizable, a true Airy beam cannot be created. Experimental approximations to an ideal Airy beam can be produced using truncated wavefronts [32], and have been found to have reduced diffraction, and transverse acceleration of the main lobe of the beam, in both light optics, and TEM [32,33]. The reduced experimental diffraction is similar to that of a Bessel beam, as discussed by Durnin and Miceli [34], while the transverse acceleration of the main lobe of the beam remains physically possible, as this is countered by the acceleration of the subsidiary lobes in the opposite direction – and hence the center of mass of the whole wavefront maintains a rectilinear trajectory [35].

Electron Airy beams have been demonstrated previously in TEM [33], but the holographic mask technique which was used has a limited efficiency and leads to the undesired creation of conjugate and higher order beams. Production with direct phase manipulation using aberration correctors enables a swift and effective bypass of these drawbacks.

It can be seen that the Fourier transform of (5) is characterized by a cubic phase, and therefore a phase plate $\chi = k_x^2 + k_y^2$ is required (where $k_x = (2\pi/\lambda)\cos \phi$ and $k_y = (2\pi/\lambda)\sin \phi$ are the components of the transverse momentum), which can be achieved by minimizing all aberrations other than three-fold astigmatism ($A_2$) and third-order axial coma ($B_2$). We adjust their relative strength such that $B_2 = 3A_2$ and relative direction such that $\Phi_{A_2} = \Phi_{B_2} + (\pi/3)$. The resulting theoretical and experimental results are displayed in Fig. 2.

Fig. 2a shows the phase distortions applied to the electron wavefront by these aberrations. The simulated and experimental beam intensity profiles in the sample plane are shown in Fig. 2b.

As a demonstration of the interesting self-accelerating property of the Airy beam, we record the beam intensity profile through a defocus series from 0 nm to +160 nm, in steps of 10 nm. This is represented in Fig. 2, as both a 3D visualization, and as a 2D $xz$-plane slice. The acceleration of the main lobe of the beam is clearly seen in the parabolically curving trajectory.

2.3. Electromagnetic fields in matter

Electrical fields are ubiquitous in matter at the microscopic scale, and even in neutral matter the spatial distribution of the fields results in a finite volume average of the electric potential known as the mean inner potential. When an electron beam passes through a thin weakly scattering sample the amplitude variation is minimal, but a phase modulation is impressed on the beam. It is then easy to see from Eq. (1) that a transparent film modulated in thickness acts as a phase plate modulating the wave with an energy-dependent phase shift. Phase plates expand the TEM possibilities and have been studied for several decades now. Zernike phase plates have been realized preparing a film of carbon or silicon nitride of appropriate thickness and milling a hole in its center through which the unscattered component can pass unaltered. This film can then be inserted into the TEM, in the back focal plane of the objective lens [36]. Although the approach has been proven to work, the limited lifetime and stability of the phase plates, which are damaged and contaminated by the irradiation while in use, have until recently prevented widespread adoption of the technique [37,38]. Boersch phase plates have been invented in order to overcome these limitations [39], and have been successfully demonstrated in recent years. This type of phase plate aims to introduce a $\pi/2$ phase shift by altering the unscattered wave component rather than the scattered one, and is implemented as a tunable electrostatic lens, which has the advantage of not being specific to one single electron energy [40–42].

The Aharonov–Bohm effect has also been exploited both to create Zernike and spiral magnetic phase plates.

In the former case, a magnetic ring functions as a Zernike phase plate. The electrons passing inside the ring acquire a relative Aharonov–Bohm phase shift of $\pi/2$ with respect to the electrons passing outside of the ring [43,44]. The thickness and material of the magnetic ring determine the flux density through its section, which in turn determines the phase difference between passing electron waves.

In the latter case, the tip of a thin magnetic needle is used to obtain a spiral phase plate. The needle is in essence nothing more than an elongated dipole magnet. By separating the two ends of this dipole by a large distance, and placing an aperture over one end, a passing electron beam can be made to experience a magnetic monopole-like field [45]. There are some artifacts due to the needle, which is not infinitely thin and also has some surface imperfections which affect the shape of the resulting field. For a sufficiently thin needle, one can ensure it has the proper magnetic flux which will spread out radially in 3 dimensions at the tip. This local field distribution leads to the electron beam gaining a linear, azimuthally dependent phase, due to the (approximately) spherical distribution of magnetic flux emanating from the needle’s tip in all directions. This can most easily be seen by looking at the electron beam traveling past the monopole along the $z$-axis, as illustrated in Fig. 3a.

The main challenge with a magnetic material is that one must ensure a single-magnetic domain object to prevent unwanted stray fields disrupting the ideal magnetic phase distribution. In both cases, this was achieved by making the magnetic object sufficiently thin. It is worth pointing out that the Aharonov–Bohm phase shift is independent of the electron energy, and the phase plates based on this principle are equally effective for beams of any energy. Optical deflection on the other hand does depend on the energy via the wavelength.

2.4. Holographic reconstruction

Ironically the simplest way to manipulate the phase of a wave is by manipulating its intensity through computer generated holography (CGH). Despite having several limitations this method has been employed extensively in electron singular optics as it is effective, simple to apply and highly general. Bessel, Airy and vortex beams have been demonstrated with CGH [46,47,33,48,49] with a reference wave such as a tilted plane wave or a spherical wave. The resulting interference pattern can then be used to manufacture an aperture by standard nanofabrication techniques, that can then be inserted in the TEM to produce the desired wave. In the simplest case the reference wave used is a tilted plane wave and the interference pattern calculated is then binarized by applying a threshold on the
intensity and then used to mill an absorbing metal (such as gold or platinum) film in order to produce a binary aperture. The resulting mask acts as a specialized diffraction grating, in which the first diffraction order will take the form of the target wave. Due to the binarization higher diffraction orders are also present which possess a phase which is a multiple of the target wave, while a non-binary hologram preserving the full intensity modulation of the calculated interference would only produce first order diffracted beams.

Many variations of this technique have been demonstrated. If a spherical reference wave is used the different diffraction orders are separated along the beam axis rather than in the transverse direction. Holograms can also be fabricated out of light materials such as C or Si3N4 to obtain pure phase holograms that are highly transparent, guaranteeing a higher transmitted intensity. If the hologram is given a blazed design, then in principle all the intensity can be directed into the desired diffraction order and experimental demonstrations have reached a 25% efficiency [50].

On-axis holograms have also been demonstrated. A hologram calculated from the target wave through an iterative Fourier transform algorithm is imprinted on a transparent film to obtain a pure phase mask. While this approach allows to sculpt the intensity distribution of the beam with a high flexibility, its efficiency is low. Due to the physical impossibility of producing subwavelength structures for electrons at TEM energies, most of the wave intensity still forms an on-axis central spot, which is about 400 times more intense than the target wave produced, as well as speckling in the pattern and multiple diffraction orders [51].

3. Dreamscape: applications

But what about the future? If such a “spatial electron modulator” were to be invented, would it be of any use for electron
microscopy? Here we propose several possible applications of electron wave manipulation techniques.

3.1. Specific probes for specific results

The flexibility provided by manipulating the wave structure of the probe can be exploited by matching the beam to the property of interest in the sample.

For instance, the helical character of a vortex beam can be used to detect the handedness of chiral space groups, in which atoms are arranged on a helix. By performing convergent beam diffraction (CBED) with a vortex beam of appropriate size and topological charge, the radial distribution and symmetry of the higher order Laue zones are sensitive to the chirality of the crystal, and the information can be retrieved by careful analysis of the diffraction pattern or of the HAADF-STEM signal [31]. This technique provides local chirality information and does not require extensive dynamical simulations.

Fig. 3. (a) An apertured long dipole as an approximation to a monopole. Electrons travel from the reader’s eyes into the page past the needle. In order to produce a perfect electron vortex, the magnetic needle must be magnetized such that it contains just enough flux to induce a relative Aharonov-Bohm phase shift of \(2\pi\) for electrons passing along either side of the needle (marked with points A and B). Due to the continuity of the wave function over the region C, electrons traveling along paths in this region will obtain an azimuthal phase. From the symmetry of the monopole-like field emerging from the tip of the needle, this phase will increase linearly when going round the tip from point A to point B. (b) and (c) Experimental electron holograms recorded in order to measure the magnetic phase-plate generated by a magnetized needle. The magnetization of the needle has been reversed between the two images. (d) and (e) Magnetic phases extracted by analyzing the holograms.
Vortex beams have also been used to rotate nanoparticles in the TEM. It was in fact found that illuminating a 3 nm gold nanoparticle with a vortex beam of comparable size induced a rotation whose direction depended on the handedness of the vortex. The rate of angular momentum transfer was estimated by numerical simulations to be about $0.1\hbar$ per electron. A typical beam current of about 50 pA would then yield a rotational acceleration around $10^{14}$ rad/s$^2$ in the absence of friction. The observed speed of rotation however was found to be approximately 0.01 rad/s owing to the strong friction between the particle and the substrate. Electron vortex beams appear to be a useful tool to study rotational friction at the nanoscale, a field which had thus far been inaccessible [25].

Other types of beams are also being studied. One class of structured beam which shows promise are non-diffracting beams. A prime example is the Bessel beam, which cannot be experimentally produced as it is inherently unnormalizable. A practical approximation can be achieved with real-world (electron) optics. The non-diffracting property is then reduced to a strongly extended depth of focus, thus enabling high resolution HAADF-STEM images of highly tilted samples [48]. Furthermore, the limited angular spread of such a beam reduces the effect of spherical aberrations offering some potential for higher resolution in non-Cs-corrected instruments.

Tailored wavefronts can also be used to investigate inelastic interactions. Electron vortex beams show promise for the study of magnetic materials at atomic resolution. By exciting a spin polarized transition in a magnetized atom with an electron vortex beam the usual selection rules for electron energy loss are altered. This makes it possible to selectively study one spin polarized transition by applying an appropriate post-selection step [27–30]. Recent work seems to indicate that a similar magnetic signal can also be obtained by introducing probe aberrations with a symmetry matching the symmetry of the crystal [52]. It has also recently been suggested that electron energy loss spectroscopy performed with vortex beams can be used to probe chirality in plasmonic excitations in nanoparticles and biological molecules [53].

### 3.2. Structured illumination

Structured illumination microscopy is a method to achieve super-resolution imaging, used in light microscopy, and might present an alternative to other super-resolution schemes developed for electron microscopy [54–58].

When a sinusoidal grating pattern is projected on the sample, the superposition of this fringe pattern with the spatial frequencies of the sample will give a Moiré effect. A Moiré pattern is modulated by frequencies equal to the difference between the projected grating’s frequency and those of the sample. These modulation frequencies can be used to recover frequencies above the optical band limit of the instrument (Fig. 4).
By using a grating frequency equal to the upper band limited frequency of the microscope one can double the instrument’s resolution at the cost of taking more images requiring a high stability [59–61]. Different super-resolution imaging methods based on structured illumination have also been demonstrated. It has been shown that employing speckle patterns as illumination patterns allows to relax the stability requirements, at the cost of further increasing the dose [62]. Recently, structured illumination making use of a STEM probe in CTEM was shown to beat the information limit of a commercial TEM in a technique termed ISTEM [63].

Structured illumination is also useful for optical sectioning. The sinusoidal pattern will undergo diffraction as it propagates. If a thick sample is imaged only the part where the grating is in focus will display the sinusoidal intensity modulation, while parts of the sample above and below that will appear out of focus. By recording multiple images with a slightly displaced grating it is possible to extract an optical section of the sample [64,65].

3.3. Adaptive TEM

The ability to fully control the waveform in amplitude and phase could be employed to improve resolution in thick samples. Imagine trying to take a STEM image of a plane within a sample. Even though the beam is optically focused on the plane of interest the scattering from the top part of the sample broadens the electron probe compromising the resolution. Perhaps it would be possible in the future to cancel out the diffractive effects on the beam above or below the plane of interest. We can explore this concept by imagining a plane wave impinging on a sample. The wave will propagate through the sample, interacting with the local potential of the material and will reach the other side of the sample, as the familiar exit wave. Because of the time inversion symmetry of an electron’s elastic propagation through a crystal, we can reverse propagate the exit wave, now entering at the far side of the sample, and emerging on the other side as the initial plane wave. This conceptually illustrates that the (elastic) effect of the sample on the wave can be undone if an appropriate wave modulation is applied to the probe. In fact, a probe can be tuned to deliver a desired wave at a given depth inside a crystal as long as the right starting wave can be produced. This might be useful to overcome the scattering of the probe in confocal microscopy, making use of iterative techniques to shape the input wave for a given position and depth.

4. Conclusion

In conclusion, we have explored the possibility and desirability of a device to freely control the phase of the electron wave inside a TEM. We have reviewed some examples on how approximations to this versatile device can be achieved with current technology and how they offer interesting new classes of waves with potential for applications. Neglecting the technological difficulties in making this device, we take a look at techniques borrowed from optics where phase and amplitude manipulation of waves are common and suggest that many new and attractive imaging techniques could become available to the TEM community in the future.

Acknowledgments


References


