

Three-dimensional Stacking of Phase Plates for Advanced Electron Beam Shaping

Gianluca Ruffato^{1,2} , Marco Beleggia³, Amir Hossein Tavabi⁴, Enzo Rotunno^{5,*}, Lorenzo Viani^{3,5} , Paolo Rosi⁵ , Payam Habibzadeh Kavkani^{3,5}, Caterina Chiari^{3,5}, Stefano Frabboni^{3,5}, Gian Carlo Gazzadi⁵, Giulio Pozzi³, Giovanni Bertoni⁵ , Peter Tiemeijer⁶, Rafal Edward Dunin-Borkowski⁴, and Vincenzo Grillo^{5,*} 

¹Department of Physics and Astronomy 'G. Galilei', University of Padova, via Marzolo 8, Padova 35131, Italy

²Department of Information Engineering, University of Padova, via Gradenigo 6, Padova 35131, Italy

³Dipartimento FIM, Università degli studi di Modena e Reggio Emilia, via G. Campi 213/A, Modena 41125, Italy

⁴Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons, Forschungszentrum Jülich, Jülich 52425, Germany

⁵Istituto Nanoscienze, Consiglio Nazionale delle Ricerche, via G. Campi 213/A, Modena 41125, Italy

⁶Thermo Fisher Scientific, PO Box 80066, Eindhoven 5600 KA, The Netherlands

*Corresponding authors: V. Grillo, E-mail: vincenzo.grillo@nano.cnr.it; E. Rotunno, E-mail: enzo.rotunno@nano.cnr.it

Abstract

Tuneable phase plates for free electrons are a highly active area of research. However, their widespread implementation, similar to that of spatial light modulators in light optics, has been hindered by both conceptual and technical challenges. A specific technical challenge involves the need to minimize obstruction of the electron beam by supporting films and electrodes. Here, we describe numerical and analytical mathematical frameworks for three-dimensional stacks of phase plates that can be used to provide near-arbitrary electron beam shaping with minimal obstruction.

Key words: beam shaping, Bessel beams, conformal mappings, electron optics, iterative methods

Introduction

Electron beam shaping (Harris et al., 2015; Bliokh et al., 2017) holds great promise for tailoring the electron wavefunction in an electron microscope, as well as for the measurement and decomposition of an electron beam in any basis (Grillo et al., 2017; Fickler et al., 2020; Troiani et al., 2020; Ruffato et al., 2021; Tavabi et al., 2021, 2024). Despite analogies with single particle wave equations in light optics, there are differences in the practical implementation of electron beam shaping. In light optics, a spatial light modulator (SLM) can be used to achieve phase and amplitude shaping, with computer control used to produce desired beam shapes (Bolduc et al., 2013). In electron beam shaping, the manipulation of charged particles requires the use of apertures and electromagnetic fields and is subject to Maxwell's equations (Verbeeck et al., 2018; Pozzi et al., 2021; Yu et al., 2023). The most similar approach to the use of an SLM for electron beam shaping currently involves the fabrication of miniature Einzel lenses that can be used as programmable pixels (Verbeeck et al., 2018; Yu et al., 2023). However, electronic control of the lenses is implemented in the electron beam path, obstructing parts of it and reducing its overall intensity.

An alternative approach involves the use of harmonic or quasi-harmonic phase plates that have minimal or no material in the electron beam path and satisfy the criterion of overall charge neutrality. A harmonic constraint means that the phase

ϕ must satisfy the relationship $\nabla^2\phi = 0$ at all points other than at the locations of a small number of “allowed” obstructions. The quasi-harmonic case includes the possible presence of small obstructions. In mathematical terms, the field sources can be described as a series of Dirac δ functions, but are in practice approximated by small finite-sized elements for practical reasons.

The relationship between a three-dimensional electrostatic potential and the harmonicity of the corresponding phase distribution is discussed in Appendix A. Harmonic and quasi-harmonic phase plates can be realized using micro-electromechanical systems technology, which allows electrode shapes and biasing circuits to be designed and fabricated in a planar geometry (Tavabi et al., 2022). In the harmonic case, electrodes placed at the edges of a phase plate can be used to define “boundary conditions” for the potential and phase. Such phase plates can be described mathematically in terms of individual or finite sums of multipoles. The simplest approach for realizing the quasi-harmonic approach is to place needle-shaped electrodes in the path of the electron beam (Pozzi et al., 2020; Ruffato et al., 2021; Tavabi et al., 2021, 2022, 2024). This approach has been used to create an orbital angular momentum (OAM) sorter and to generate electron vortex beams. Figure 1 shows examples of currently available harmonic and quasi-harmonic phase plates that have been realized experimentally and can be combined.

Received: May 14, 2024. Revised: September 14, 2024. Accepted: October 14, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of the Microscopy Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

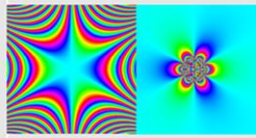
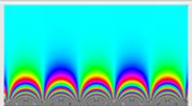
	Phase	Laplacian	Phase	Electrodes
Biprism	$\varphi = k x $	$\delta(x)$		
Vortex beam / spiral phase plate	$\varphi = \ell\theta$	0 except 2 lines		
n-fold astigmatism / multipoles	$\varphi = \rho^n \cos(n\theta + p)$	0 if $n > 0$ $\delta(r)$ if $n < 0$		
Sorter (1st element)	$a \left[y \operatorname{atan} \left(\frac{y}{x} \right) + x \ln \left(\frac{\sqrt{x^2 + y^2}}{b} \right) + x \right]$	0 except 1 line		
Sorter (2nd element)	$ab \exp \left(-\frac{x}{a} \right) \cos \left(\frac{y}{a} \right)$	0		

Fig. 1. Examples of currently available harmonic and quasi-harmonic phase plates. The first column provides analytical formulae for the phase, where $\rho = \sqrt{x^2 + y^2}$ and $\theta = \operatorname{atan}(y/x)$. The second column gives the Laplacian of the phase. The third column shows a colour representation of the phase (with the hue representing the magnitude of the phase). The fourth column shows a schematic representation of electrodes that can be used to generate the phase (with different colours representing different applied voltages).

Although the resulting phase landscape, which exhibits negligible amplitude modulation when examined close to focus, appears to provide a perfect solution, it can be limited by an insufficient number of field sources in the illuminated area and by the smoothness of the phase across most of the region of interest. Here, we investigate a strategy that can be used to overcome these limitations by stacking phase plates above one another in the electron beam direction. Although current technology still limits the realization of a large number of superimposed phase plates, a dedicated hardware solution could make such a configuration possible in the future. It is therefore timely to explore this concept using mathematical frameworks taken from light optics. Inspiration can also come from the field of electron-light interactions (Madan et al., 2022), ultimately in three dimensions (García de Abajo & Ropers, 2023). We first explore the use of iterative methods to calculate an optimal sequence of phase plates to realize a desired wave function. We then explore analytical and semi-analytical closed-form solutions for interaction and propagation in the limit of the stationary phase approximation. We also use analytical arguments to consider limitations to the encoding of arbitrary wavefunctions.

Formalization of the Problem

An initial electron wavefunction is assumed to be a near-plane wave that can be written in the form

$$\psi(x, y, z) = Na(x, y)\exp(ik_z z), \quad (1)$$

where $a(x, y)$ represents either a smooth cut-off of the amplitude or a hard aperture (such as a condenser aperture) and

N is a normalization factor. The $k_z z$ term represents the phase factor for forward propagation in the z direction and is omitted hereafter.

The influence of a phase plate is assumed to take the form of a pure phase object transmission function

$$\psi(x, y; z_i + \delta z) = \psi(x, y; z_i) \exp(i\varphi_i(x, y)), \quad (2)$$

where z_i is the z position of the i th phase plate introducing phase φ_i , and δz represents the spatial extent of the phase plate. This simplification is valid only when the electromagnetic field of the phase plate extends over a limited distance δz in the z direction, i.e. when diffraction effects resulting from propagation over this distance can be neglected. If the field of the phase plate extends considerably in z , then a more complicated expression can be derived with complete 3D propagation over the fringing field of the phase plate. Here, we assume that the field above and below the phase plate decays over a distance that is smaller than the Fresnel distance $\delta z \ll d^2/\lambda$, where d is the beam diameter and λ is the electron wavelength.

Propagation between successive phase plates can be considered by selecting finite propagation distances or by stacking successive phase elements in specific Fourier-conjugated planes (where each plane is in the diffraction/far field plane of the previous plane). In the paraxial regime, Fresnel propagation can be modelled as a convolution (represented by the symbol $*$) based on a parabolic propagator of the form

$$\psi(x, y; z_{i+1}) = \psi(x, y; z_i) * \exp\left(\frac{i\pi}{\Delta z \lambda} \rho^2\right), \quad (3)$$

where $\Delta z = z_{i+1} - z_i$. In more complicated scenarios, round lenses can be placed between successive phase plates, or a

convergent incident electron wave can be used. Lenses can serve to magnify a probe from the condenser aperture to a desired size. A transfer lens can also increase the effective propagation distance without the necessity to have a physically long element. Although such considerations are important for real device fabrication, lensing terms are incorporated mathematically into the phase plates described below. Focusing terms are required to adapt the initial and final beam diameters. This demand can be reduced by appropriately choosing similar lateral dimensions for the initial aperture and final state.

It should be noted that the alternating nature of interaction and propagation between phase plates is similar to the numerical simulation of dynamical electron-specimen interaction using the multislice algorithm (Cowley & Moodie, 1957). This similarity should be kept in mind when using inversion methods. Techniques such as ptychography also deal with the problem of finding the most appropriate potential that connects the probe with diffraction by the specimen (Maiden et al., 2012; Chen et al., 2021).

Iterative Numerical Approach to Wave Synthesis

The main problem of wave synthesis is to find an appropriate phase distribution for each phase plate, such that the initial and final state resembles the initial wave (as in Eq. (1)) and desired wave function, respectively.

Considering the similarity of the problem to multislice ptychography (Maiden et al., 2012; Chen et al., 2021), a natural choice for finding intermediate phase plates is to use an approach based on the Gerchberg–Saxton algorithm (Gerchberg & Saxton, 1971), which is an example of more general iterative “error reduction” algorithms (Fienup, 1982). Here, the error to be minimized is the sum of the errors of fitting the initial and final waves, both of which are known. In its standard form, it calculates missing phase information by making use of amplitude measurements in two Fourier-conjugated planes. The wavefunction is propagated forwards and backwards between a chosen plane and its Fourier conjugate, starting from a guess for the phase distribution of the phase plate to be determined. As the iterations proceed, the amplitude in each plane obtained by transformation is substituted by its measured counterpart. Even if the phase is not strictly correct at each step, amplitude substitution typically reduces the difference between the calculated and true phase in both spaces monotonically. Such error reduction algorithms are not restricted to Fourier conjugation and have also been used in phase reconstruction based on focal series in high-resolution transmission electron microscopy (Allen et al., 2004).

It is not difficult to imagine an extension to the Gerchberg–Saxton algorithm, in which further planes are introduced, with the amplitude remaining unaltered at the positions of successive intermediate phase plates. In the case of many-step propagation, it is possible to proceed iteratively between the first and last plane by propagation, with the update rule for each intermediate plane involving adjusting the phase by the phase difference between forward and backward propagation.

If both the amplitude and the phase need to be encoded and if the phase at each step is arbitrary, then the problem can be reduced by assuming that only the final phase plate takes care of the phase, with no further propagation. Conversely, if only

quasi-harmonic phase plates are used, then a large number of intermediate phase plates may be required.

The iterative method with arbitrary phase plates has been demonstrated in light optics in the “wavefront matching” algorithm (Fontaine et al., 2019). In this scheme, forward A and backward B propagating waves sampled on a matrix are calculated. The forward wave starts (e.g., from the aperture beam) and propagates to the final plane, while the backward wave propagates from the desired probe back to the aperture plane. At every iteration, the phase of each phase plate i is updated, with the update rule at the i th plane calculated using the expression

$$M'_i = M_i + A_i \cdot B_i^* \exp(iP) + c, \quad (4)$$

where M is a complex matrix that contains the phase and is related to the phase of the plate through the phase matrix $\phi_i = \arg\{M_i\}$. The scalar factor P is the average (over all matrix elements) phase difference between $A_i \cdot B_i^*$ and ϕ_i , where the $A_i \cdot B_i^*$ operation is an elementwise product and $*$ denotes conjugation. The forward and backward waves are recalculated using the new phase plate, with a small constant c (formally a constant matrix) used to achieve a smoother solution in the part of the image where the amplitude is small. The value of c is chosen to be a fraction of the maximum wave intensity, so that only the less intense part of the wave is not encoded.

Qualitatively, the iteration substitutes something more “correct” for the previous guess at each step. Here, the phase that matches the forward A and backward B wave is more correct at each step. The corrections occur in both directions, i.e., from the first phase plate to the last and from the last to the first, ensuring uniform convergence. In the convergent form, A and B describe the same wavefunction evolution connecting the initial form of Eq. (1) to the synthesized wave in the final plane.

The use of 200 iterations and 3–7 phase plates typically produces a reasonable result that improves with the number of phase plates.

The original paper (Fontaine et al., 2019) suggests that such an iterative technique could be used to build a universal sorter of orthogonal quantum states that is able to separate a wave function into orthogonal component functions. Such transformations are “lossy,” meaning that some of the intensity is scattered diffusely, in part due to the presence of discontinuities in the phase or its derivatives. The application of such a sorting approach in transmission electron microscopy will be considered in a separate article.

Figure 2 shows an example of the use of three intermediate phase plates to form a Laguerre–Gaussian $L_{3,5}$ beam (McMorran et al., 2011; Rosi et al., 2022) starting from a hard aperture. In this example, the chosen beam is a vortex beam, so the phase plates contain vortex-like phase distributions. The beam also contains a π phase discontinuity in the radial direction. The target and reconstructed beams are shown in Figs 2(b) and 2(c), respectively. The propagation distance between the phase plates was set to be on the order of the Fresnel distance, with a beam of radius 200 nm corresponding to a distance between phase plates of ~ 1 mm at 300 keV. Threshold constants of $c = 10^{-5}$ and $c = 0$ were used for Figure 2(a–c) and (d, e), respectively; 200 iterations were used.

From the images, it appears that the first phase plate primarily introduces the vortex phase, the second phase plate primarily reassembles the amplitude and the third phase plate

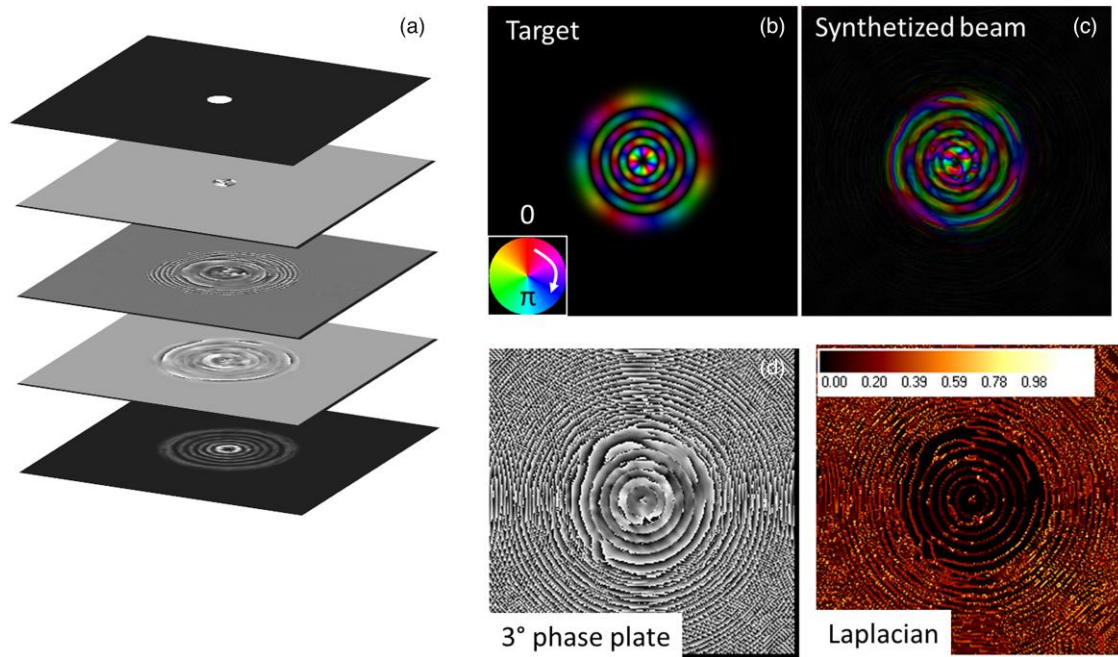


Fig. 2. Application of a wavefront matching algorithm to a three-phase-element configuration. **(a)** Representation of three phase plates and initial and final wave amplitudes, showing in sequence the initial amplitude of a hard aperture, the phases of three phase plates and the amplitude of a reconstructed $L_{3,5}$ wavefront. **(b)** and **(c)** The phases of the target and synthesized wavefunctions. Hue and brightness refer to phase and intensity, respectively (see the inset colour wheel). **(d)** The phase of the third phase plate (without any threshold). **(e)** The absolute value of the Laplacian of (d).

primarily takes care of the radial discontinuities. The solution is unlikely to be unique, as the roles of the phase plates could be swapped or a more non-centrosymmetric solution could be used. For a sufficient number of phase plates, the first one could be chosen arbitrarily and a reasonable solution could still be found. However, the solution would be non-optimal in terms of diffuse intensity.

Although the phase plates do not have explicitly “harmonic” phases, the Laplacian of the third phase plate [Figure 2(d, e)] shows that the phase is close to being quasi-harmonic, with the nonzero parts mainly concentrated in circular discontinuities of close to π radians in the radial direction. The discontinuities are required to reproduce the discontinuities that are inherent to an $L_{3,5}$ beam, where a real function describes the radial part of the wavefunction. The constant c in Eq. (4) can be used to achieve smoother (and therefore closer to harmonic) solutions.

Discontinuities and vortices are aspects of wavefunction generation that cannot be introduced easily using purely harmonic elements. Interestingly, a stigmator can be used to transform a vortex beam into a Hermite–Gaussian-like beam with a π radian discontinuity (Kramberger et al., 2019). However, for smooth, continuous (not necessarily harmonic) phase plates, there is no clear mechanism based on three-dimensional stacking to make singularities appear or disappear without using discontinuity-carrying phase plates. In detail, propagation described by convolution with a complex Gaussian kernel cannot directly introduce discontinuities. This statement has formal similarities to the fact that a phase retrieval algorithm cannot introduce phase vortices when using the transport of intensity equation for phase retrieval (Lubk et al., 2013). It also suggests that purely harmonic phase plates cannot be stacked in three dimensions to generate arbitrary waves.

At this point, it is tempting to explore the idea of implementing a (quasi) harmonicity constraint. The simplest approach would be to substitute a partial solution with a harmonic counterpart at each step in the error reduction scheme. Unfortunately, this procedure is not guaranteed to monotonically reduce the difference between the iterative solution and a solution (if it exists) that produces a desired set of harmonic phase plates and a final probe. At the moment, the problem of finding a set of (quasi) harmonic phase plates for arbitrary wave shaping does not have a general iterative solution. In particular, it is not possible to determine whether a solution exists and how many phase plates are needed for its implementation. The results presented in the next section provide guidance on this point. We also confirmed in tests that, by increasing the number of stacked phase plates, the quality of the resulting wavefunction could be improved to provide a more faithful reconstruction and a less lossy result, in agreement with the work of Fontaine et al. (2019).

Analytical Solutions for Three-dimensional Phase Plates

In this section, the problem of finding a set of (quasi) harmonic phase plates that can be used to transform a near-plane-wave beam into a near-arbitrary beam is tackled for the specific case when an analytical solution exists. We show here that an analytical solution can be calculated if the stationary phase approximation is used to solve the propagation integral between the phase plates. This requirement assumes that the in-plane gradient of the phase plate is much stronger than the gradient of the phase of the incoming beam. In this approximation, the evolution of the wave can be predicted by assuming that part of the wavefunction (in both phase and amplitude) is transported to another position, which is

dictated by the direction of the local gradient of the phase plate in the case of Fourier conjugation or by the local coordinate and local gradient direction in the case of Fresnel conjugation. This assumption corresponds to the propagation of an input wavefront by a point-to-point projection onto a destination plane, i.e., a coordinate transformation (Hossack et al., 1987).

Conformal mappings have a special place in coordinate transformations, where the word “conformal” refers to the fact that such a transformation conserves angles between coordinate axes locally. If complex spatial coordinates $g = x + iy$ are used, then a transformation $(x, y) \rightarrow (x', y')$ can be expressed in complex form as a $g \rightarrow g'$ function $q(g)$. The conditions for a coordinate transformation to be a conformal mapping are described in Appendix B.

Conformal Mapping in Fourier Conditions

A well-known implementation of conformal mapping in Fourier conjugation is provided by the OAM sorter (Berkhout et al., 2010; Grillo et al., 2017; Pozzi et al., 2020; Tavabi et al., 2021, 2024), which involves the application of a *log-polar* transformation using two Fourier-coupled phase elements. The *log-polar* change of coordinates is

$$(x, y) \rightarrow \left(a \operatorname{atan}(y/x), -a \log\left(\frac{\sqrt{|x^2 + y^2|}}{b}\right) \right), \quad (5)$$

where a and b are scaling parameters. In complex coordinates, this expression can be written in the form $g \rightarrow a \log(\frac{g}{b})$. This transformation maps angular onto linear coordinates, converting azimuthal phase gradients typical of OAM beams into tilted plane waves. Upon free space propagation, these, in turn, are sorted into spatially separated points in the far field.

We previously investigated the generation of near-arbitrary conformal mappings by combining two phase plates in Fourier-conjugated planes (Ruffato et al., 2021). Whereas the first phase plate imparts a desired transformation, the second phase plate compensates for phase distortions due to propagation and completes the mapping. An important aspect of conformal mapping is that a coordinate transformation q is only conformal if the phase plates that introduce the transformation are harmonic or quasi-harmonic.

The first step of the mapping, including the integral that takes the wavefunction to the intermediate plane, can be written as a function of the phase of the first element φ_1 in the form

$$\psi_1(u, v) = \frac{1}{i\lambda z} \int \psi_0(x, y) \exp(i\varphi_1) \exp\left(-ik \frac{ux + vy}{z}\right) dx dy. \quad (6)$$

The stationary phase approximation can be applied and the integral in (x, y) can be substituted by values of the integrand calculated at stationary (saddle) points of the total phase, namely $\nabla(\varphi_1 - k \frac{ux + vy}{z}) = 0$. Here, stationary points are indicated by a tilde symbol as $(\tilde{x}_j, \tilde{y}_j)$, and the approximate version of Eq. (6) is

$$\psi_1(u, v) \approx \sum_{\{(\tilde{x}_j, \tilde{y}_j)\}} k\sigma \frac{\psi_0(\tilde{x}_j, \tilde{y}_j) \exp(i\varphi_1(\tilde{x}_j, \tilde{y}_j)) \exp\left(-ik \frac{u\tilde{x}_j + v\tilde{y}_j}{z}\right)}{\sqrt{|H(\tilde{x}_j, \tilde{y}_j)|}}, \quad (7)$$

where $k = 2\pi/\lambda$, and H is the determinant of the Hessian of the

transformation calculated at the stationary points. That is,

$$H = \frac{\partial^2 \varphi_1(x_j, y_j)}{\partial x^2} \frac{\partial^2 \varphi_1(x_j, y_j)}{\partial y^2} - \left(\frac{\partial^2 \varphi_1(x_j, y_j)}{\partial x \partial y} \right)^2 \quad (8)$$

and the coefficient $\sigma = \operatorname{sgn}\left(\frac{\partial^2 \varphi_1(x_j, y_j)}{\partial x^2}\right)$ when $H > 0$, otherwise $\sigma = -i$. The intensity of the output field is not merely a rearrangement of the input field according to a coordinate transformation, but is also modulated by the Hessian in the denominator. This property can be used to shape the beam by setting H to fit the desired amplitude (Scholes et al., 2020).

In practice, geometrical rays depart from each point of the wavefront and converge to a few more dense areas. The density of rays in these areas is controlled by the Hessian determinant.

Unfortunately, it is not easy to invert Eq. (7) to define the phase φ_1 that produces the appropriate Hessian at the saddle points to control the beam amplitude. For a specific example where a solution can be found, it is possible to demonstrate that a conformal mapping based on two quasi-harmonic phase plates applied to an initial plane wave can produce a strongly non-harmonic solution. This is the case for zero-order Bessel beams (Grillo et al., 2014, 2016), which can be generated by using two spiral phase plates (SPPs) with opposite phase patterns. An SPP has an azimuthal phase distribution described by the expression $\varphi = m \cdot \operatorname{atan}(y/x)$. In the stationary phase approximation, such a phase term is equivalent to the conformal mapping

$$(x, y) \rightarrow \frac{fm}{k\rho}(-y, x), \quad (9)$$

where f is the focal distance of the lens and k is the electron wavenumber (Ruffato & Romanato, 2020). The same result can be described in complex notation by the expression $q = \frac{fm}{k} \frac{i}{g^*}$, where $*$ indicates conjugation and again $g = x + iy$. In polar coordinates, this transformation can be written in the form $(r, \theta) \rightarrow (fm/k\rho, \theta + \pi/2)$. Based on Eq. (7) and assuming a Gaussian beam of input waist w_0 , the intensity of the output field is

$$I_{p1} \propto \frac{1}{\rho^4} \exp\left(-\frac{2}{w^2 \rho^2}\right), \quad (10)$$

where $w' = \frac{w_0 k}{fm}$. The intensity is zero on the optical axis and decays with the fourth power of the radial coordinate at large distances from the z -axis. This behaviour is consistent with a rigorous analytical calculation (Anzolin et al., 2009) of the integral in Eq. (6) based on Kummer beams with a ring-like profile whose size increases with topological charge m and exhibits the same trend for high radial values. Equation (7) suggests that the transformed wavefunction carries the azimuthal phase $m\theta$, representing the transfer of orbital angular momentum to the input beam. This transformation phase can be compensated by a second spiral phase plate of opposite sign, i.e. $-m\theta$. In the limit of large m , the intensity profile in the intermediate plane becomes a narrow ring. Further free space propagation is expected to transform this intensity distribution into a Bessel beam in the far field.

Figure 3 shows the more realistic case of using a top hat instead of a Gaussian beam as an initial state, calculated numerically using the Fresnel propagator. An additional limiting

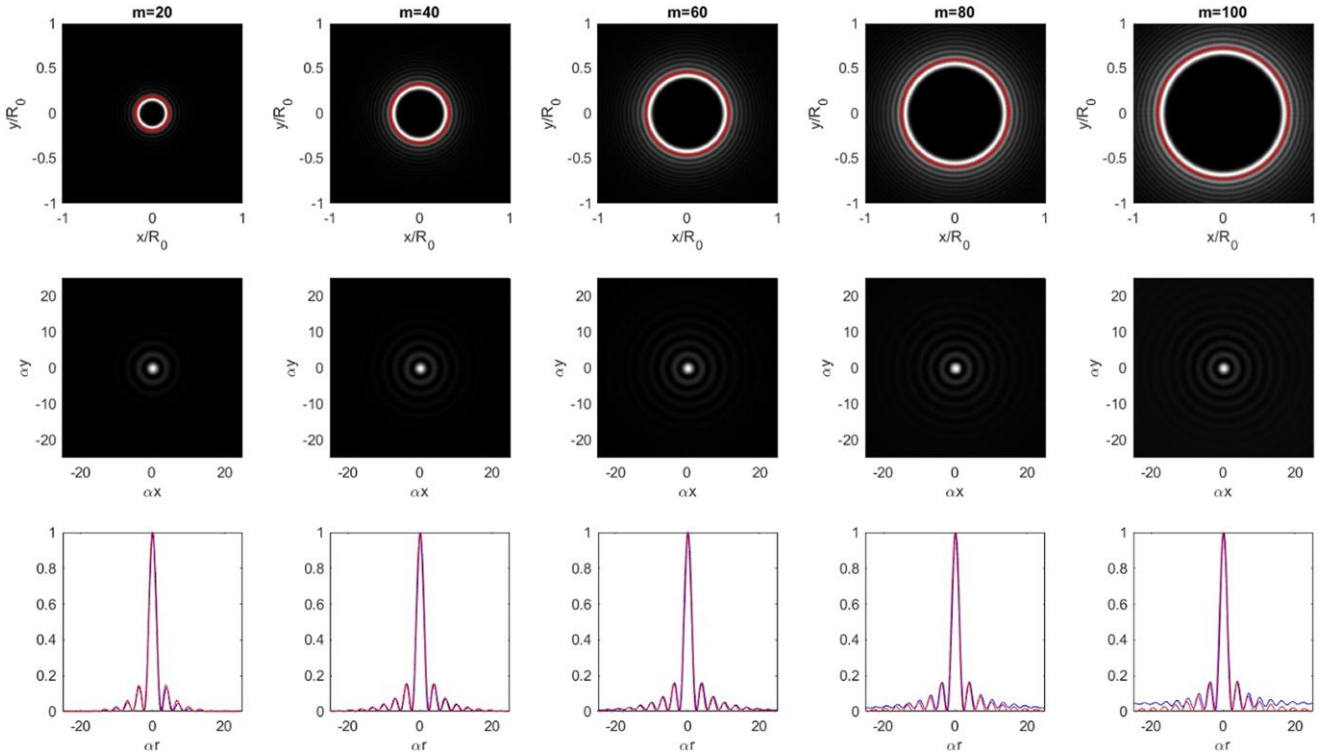


Fig. 3. Top hat electron beams shaped by SPPs with increasing topological charge m between 20 and 100 in steps of 20. (a) Intensity in an intermediate plane, in which a second SPP is placed for phase compensation. The spatial coordinates are normalized to the top hat radius R_0 . If an aperture is used to isolate the first high-intensity ring (red circle), then the final beam resembles a zero-order Bessel–Gaussian (BG) mode (the subscript after BG denotes its order, e.g., BG₀ is 0th order), as shown in (b). (c) Comparison between numerical cross-sections and theoretical profiles.

aperture is used in the intermediate plane to remove the long tails and to isolate the first high-intensity ring. Under these conditions, the beam tends to be an ideal Bessel beam. Since the ring has a finite width, it is more appropriate to describe the result as a Bessel–Gaussian beam, i.e. a Bessel mode modulated by Gaussian damping. If R and w_R are the radius and width of the intensity ring in the intermediate plane, then the final wavefunction is given by the expression

$$\psi_2 \propto J_0(\alpha r) e^{-\frac{r^2}{w^2}}, \quad (11)$$

where $\alpha = kR/f$ and $w = 2f/(kw_R)$. Conversely, if the phase $m\theta$ in the intermediate plane is not compensated exactly, as shown in Figure 4, then high-order Bessel–Gaussian beams (Gori et al., 1987) of the order $m_f = |m - n|$ are obtained, where n is the topological charge of the second SPP. Good agreement can again be seen between the simulated and ideal Bessel–Gaussian profiles.

“Inverse Sorting” as Beam Shaping

A second application of conformal mapping exploits the fact that it can be used to transform a class of functions into a series of peaks upon two consecutive stages of diffraction (one of which is related to the conformal mapping). Whereas such an approach is normally used to decompose states into a specific basis, it is possible to invert the relationship to combine states. An example involves the use of an OAM sorter to produce a combination of states with different OAM values and multiplex them with each other (Fickler et al., 2014). Figure 5 illustrates the concept of combining three OAM-radial states. If the logarithmic radial state is

labelled with quantum number k [expressed in units that are consistent with OAM (Tavabi et al., 2021, 2024)] and the OAM is labelled ℓ , then the combination is $|k=0, \ell=3\rangle + |k=1, \ell=0\rangle + |k=0, \ell=1\rangle$. Given the discrete nature of OAM states, a hypothetical pixelated amplitude controller as input could couple with specific OAM states. The fact that the generated beam has many singularities is not surprising, since one of the two sorter phase plates is based on a charged needle and is, therefore, not purely harmonic.

This concept is limited to two phase plates. Although it is possible to combine two conformal mappings using four phase plates, the number of possibilities is only improved marginally because each conformal mapping can be expressed in terms of two coupled quasi-harmonic phase plates. The primary reason for using two conformal mappings is the number of obstruction elements that violate harmonicity, which can become large with the increasing complexity of the conformal mapping. Interestingly, two successive conformal mappings can be achieved using only three elements, since the middle element can accommodate the second and first phase plates of the two conformal mappings. In order to extend the concept in a more general way, it is possible to make use of a more complex scheme that is known as transformation optics (Pendry et al., 2006). This approach is not discussed further here.

Fresnel Conformal Mapping

In this section, another analytical approach, involving Fresnel propagation between two phase plates, is considered. We show that the use of two identical but opposite phase plates

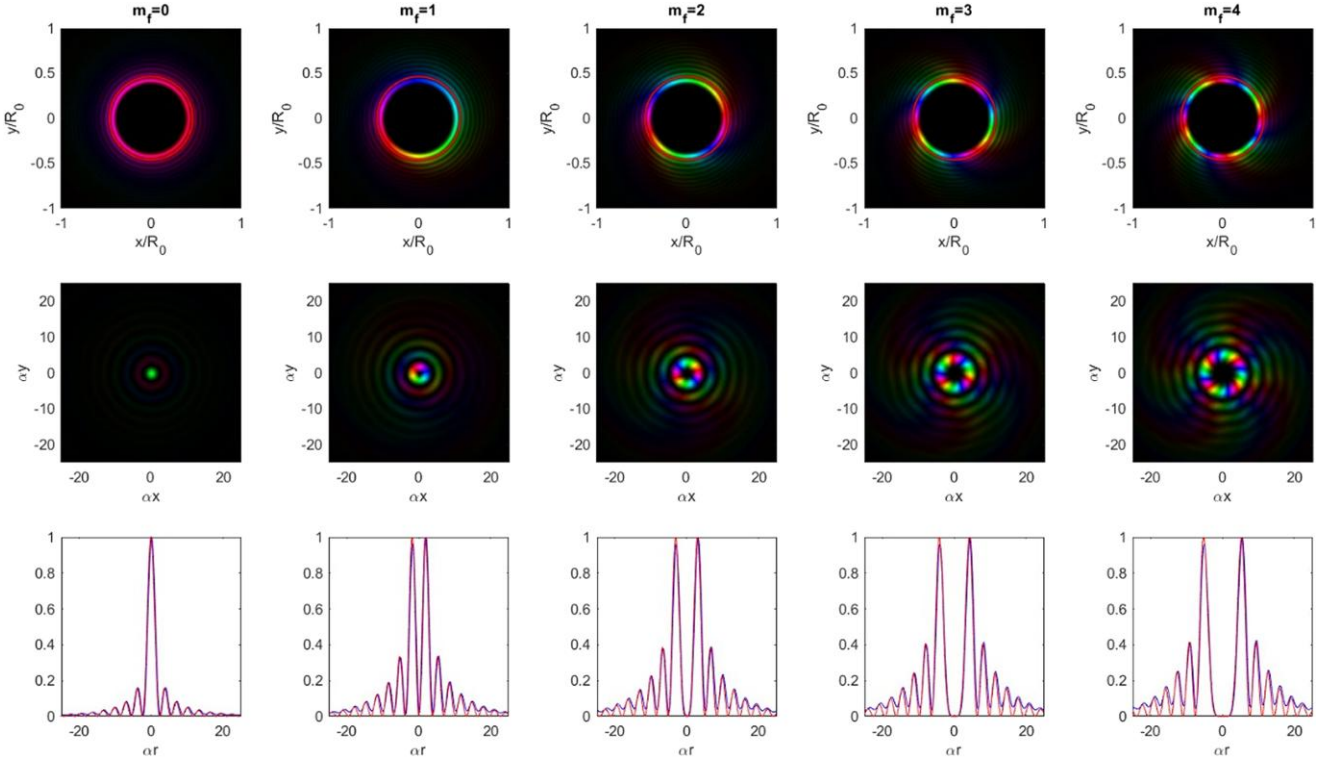


Fig. 4. Top hat electron beams shaped by SPPs with topological charge $m = 60$. (a) Intensity in an intermediate plane, in which a second SPP is placed with mismatched charge n . The spatial coordinates are normalized to the top hat radius R_0 . If an aperture is used to isolate the first high-intensity ring (red circle), then the final beam resembles a high-order Bessel–Gaussian (BG) mode of order $m_f = m - n$, (the subscript after BG denotes its order, *e.g.*, BG₀ is 0th order), as shown in (b). (Hue and brightness refer to phase and intensity, respectively, as in Fig. 2) (c) Comparison between numerical cross-sections and theoretical profiles. This example is important, as it demonstrates the formation of Bessel beams, which can be useful for overcoming spherical and chromatic aberration with only a modest loss of intensity (Grillo et al., 2015, 2016).

with a sufficient propagation distance Δz between them (typically larger than 1% of the Fresnel distance) can be used to impart a phase distribution of the form

$$\varphi_{TOT} = \frac{\lambda}{4\pi} |\nabla \varphi_0|^2 \Delta z. \quad (12)$$

Equation (12) indicates that such a combination of two harmonic phases provides a richer phase landscape than a single isolated harmonic phase plate, as it overcomes the harmonic constraint.

A basic example of the generated non-harmonic phase is that of two opposite quadrupoles generating a parabolic phase or two hexapoles generating a quartic phase $\varphi_{TOT} \propto \rho^4$.

This result of Eq. (12) can be explained in two ways. An intuitive approach is based on Fresnel conformal mapping. If the first phase plate is $\varphi_0(x, y)$, then the evolution of the wavefunction based on geometrical optics results in the transformation $(x, y) \rightarrow (x, y) + \frac{\lambda}{4\pi} \nabla \varphi_0(x, y) \Delta z$, where the displacement of each point $\delta x, y = \frac{\lambda}{4\pi} \nabla \varphi_0 \Delta z$. If this expression is substituted into $\varphi_0(x, y)$ and written in the former coordinates, then a Taylor series of the form $\varphi_0(x', y') = \varphi_0(x, y) + \delta x, y \nabla \varphi_0$ can be developed, leading to the above expression. A more rigorous approach involves the use of explicit commutation between the propagator and the phase term. In formal terms:

$$\psi_{TOT} = \exp(i\varphi_0) \exp\left(i \frac{\lambda}{4\pi} \Delta z \nabla^2\right) \exp(-i\varphi_0). \quad (13)$$

Expansion of the propagation operator to first order leads to the expression

$$\psi_{TOT} = 1 + \frac{\lambda}{4\pi} \Delta z \nabla^2 \varphi_0 + \frac{i\lambda}{4\pi} |\nabla \varphi_0|^2 \Delta z. \quad (14)$$

The middle term primarily affects the amplitude of the wave and disappears in the case of a harmonic phase plate. The rest of the expression is a first-order approximation of the phase effect described by Eq. (12). The expression is, in fact, also valid beyond the first order.

It should be noted that this idea is not new. For example, two opposite quadrupoles are known to produce a focusing effect, while two hexapoles are known to produce spherical aberration. As the sign of the induced aberration is always positive, a hardware spherical aberration corrector uses a circular lens doublet that introduces an effective negative Δz .

Defocus and spherical aberration are only two examples. A harmonic function is, in general, a combination of multipoles of the form

$$\varphi_0 = \sum_m a_m \rho^m \sin(m\theta + \theta_m) \quad (15)$$

and the total phase is

$$\varphi_{TOT} = \frac{\lambda \Delta z}{4\pi} \left| \sum_m m a_m \rho^{m-1} \sin(m\theta + \theta'_m) \right|^2 \quad (16)$$

where potentially θ'_m is different from θ_m . Unfortunately, even though this expression is more flexible than that for a single harmonic phase plate, there are limitations. One example is the positivity constraint mentioned above, which means that

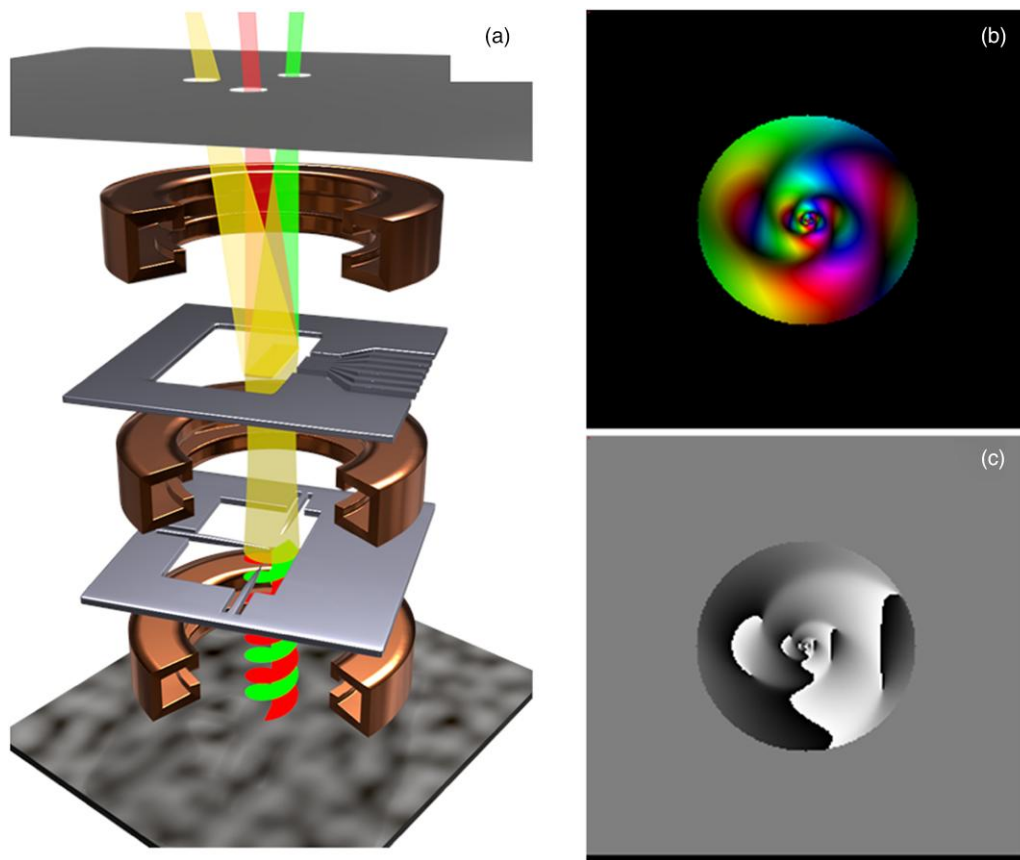


Fig. 5. Electron beam shaping by using an OAM sorter in the inverse direction. The scheme in (a) shows that the selection of a few beams in a first plane using hard apertures results in a combination of near-plane waves approaching the elements of an OAM sorter in inverse order and produces a defined combination of vortex beams. By controlling the phase and amplitude in the first plane, it is possible to generate complex non-harmonic patterns that contain discontinuities. The resulting complex colour representation (with hue and brightness representing phase and intensity, respectively) and phase alone are shown in (b) and (c) for the example described in the text. The same colour scale is used as in Fig. 2. The phase in (b) indicates that the beam is characterized by a non-trivial set of singularities.

a single doublet cannot produce a combined convex and concave phase landscape. Moreover, larger gradients concentrate close to sources and, therefore, outside or at the edges of the beam section for a perfectly harmonic phase plate.

In general, it would be possible to create a matrix formalism to treat all multipole contributions at different planes and to describe the results in terms of Zernike polynomials. However, the absence of discontinuities on normal multipoles limits the effective phases to very smooth landscapes unless a very large number of planes and orders of multipoles are used.

A simple extension involves allowing electrodes to be introduced to create a quasi-harmonic configuration. Although Eq. (14) cannot be simplified directly, the $\frac{\lambda}{4\pi}\Delta z\nabla^2\varphi_0$ term is non-zero in the region where electrodes are present and which is (usually) not electron-transparent.

Figure 6 shows an example of the use of single needles, of the type used in an OAM sorter, in coupled phase plates. The phase plates were separated by a distance on the order of the Fresnel distance to increase the second-order phase effect. The distance between the phase plates plays the role of a linear phase scaling term, but does not alter the overall shape of the result. For a needle of size $10\ \mu\text{m}$, the distance would be on the order of metres. However, it can be scaled down to dimensions of mm by applying a larger bias voltage to the

needle. The result is an overall phase plate with large phase maxima. The solution is similar but not identical to that of a single needle with a non-uniform charge distribution, especially because the non-harmonic region extends beyond the needle area. The result is compared to an exact calculation in the figure. The difference arises from the breakdown of the validity of Eq. (12) close to the needle, which violates the harmonic constraint.

If the first phase plate is not the complex conjugate of the second phase plate, then the above formula can be generalized to $\varphi_{TOT} = \frac{\lambda}{4\pi}|\nabla\varphi_1|^2\Delta z + \varphi_1 - \varphi_2$. This situation has the effect of superimposing an additional harmonic phase plate.

A more complex solution involves stacking a larger number of phase plates. However, this approach can become complicated, as a stack of two phase plates produces an overall non-harmonic phase and the above equation cannot be applied to a larger cascade. For example, a combination of three harmonic phase plates contains terms such as $\nabla^2|\nabla\varphi_0|^2$ and $|\nabla|\nabla\varphi_0|^2|^2$.

In general, phase derivatives tend to be small in material-free parts, resulting in a limitation on what can be achieved using a perfectly harmonic phase plate. The quasi-harmonic approach appears to be the only promising approach for achieving near-universal programmable phase control with reduced obstructions.

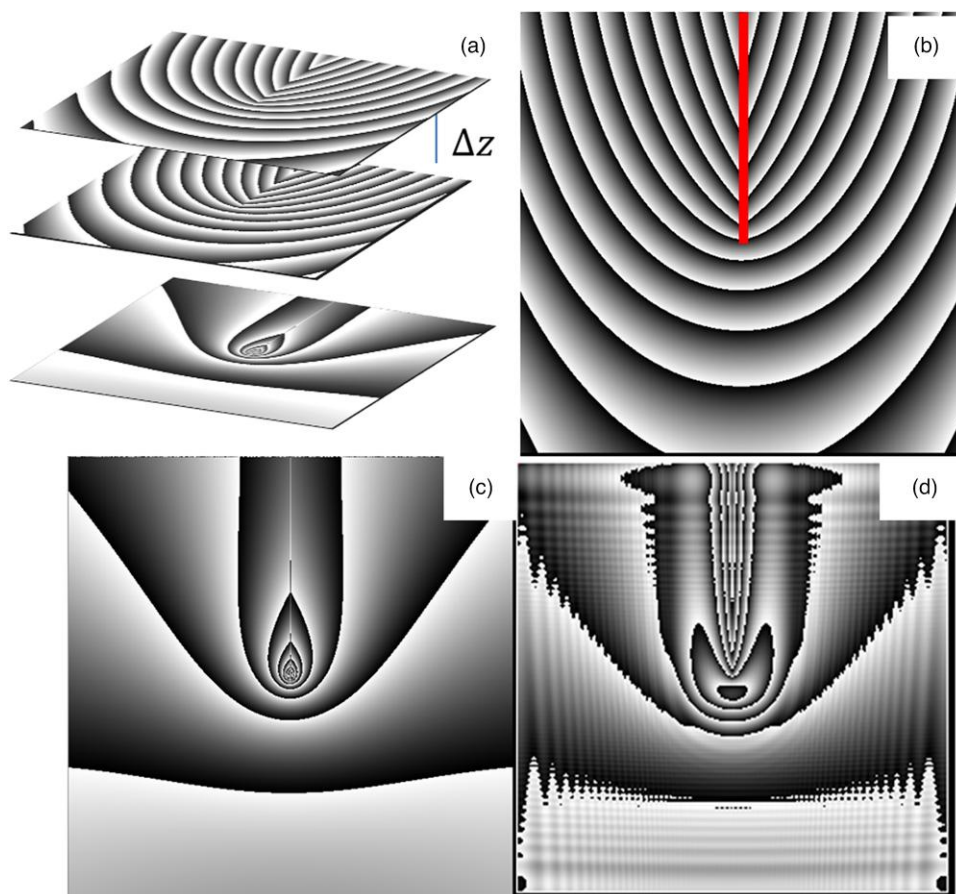


Fig. 6. Example of the effect of stacking two opposite harmonic phase plates separated by a distance Δz , as shown in (a). The position of a uniformly charged needle is marked with a red line in the phase plate shown in (b). If the gradient is calculated analytically, then the result is the phase landscape shown in (c). Instead, (d) shows the phase for explicit propagation by one Fresnel distance. The small difference around the tip arises from a breakdown of harmonicity. In this specific example, we used a propagation Δz of $\sim 10\%$ of the Fresnel distance and increased the final phase by a factor of 10.

Outlook

The numerical and analytical approaches presented in Sections 3 and 4 represent two different families of approaches to arbitrary beam shaping using (quasi) harmonic phase plates. A combined approach can be envisaged that uses the two approaches to obtain a more general solution, while retaining the framework of harmonic phase plates and conformal mapping. For example, it is possible to approximate a generic but relatively regular phase plate by a combination of needles. Such an approach could form a basis for further regularization of wavefront matching. The conditions under which the solution converges are under investigation. Taken together, our results provide confidence that, for a large enough number of phase plates, an iterative quasi-harmonic solution can be found.

Conclusions

This paper describes arbitrary electron beam shaping based on the stacking of harmonic or quasi-harmonic phase plates. Such phase plates offer the prospect of near-lossless implementation in electron optics. A solution to create an arbitrary wavefront based on an iterative approach inspired by light optics is shown to be promising but difficult to implement in the absence of a harmonic constraint on the phase. As harmonic phase plates are related to conformal mappings, an analytical

framework can be used to design a class of non-harmonic phase elements, starting from quasi-harmonic phase elements. An analytical formalism for both Fraunhofer and Fresnel mapping has been described. Both numerical and analytical approaches show that there is an important difference between quasi-harmonic and perfectly harmonic phase plates, as only the former can produce a sufficiently general phase landscape in a three-dimensional stacking concept. Our results show that there is a wide range of possibilities for generalized near-harmonic three-dimensional phase plate design.

Availability of Data and Materials

The authors have declared that no datasets apply to this piece.

Supplementary Material

To view [supplementary material](https://doi.org/10.1093/mam/ozae108) for this article, please visit <https://doi.org/10.1093/mam/ozae108>.

Financial Support

The authors acknowledge funding from the European Union under grant agreement no. 101094299 (IMPRESS), financial support under the NRRP funded by the European Union—NextGeneration EU—PRIN 2022 “AI-TEM”—CUP B53D23003780006—Grant Assignment Decree No.

- Tavabi AH, Rosi P, Ravelli RBG, Gijsbers A, Rotunno E, Guner T, Zhang Y, Rocaglia A, Belsito L, Pozzi G, Denneulin T, Gazzadi GC, Ghosh M, Nijland R, Frabboni S, Peters PJ, Karimi E, Tiemeijer P, Dunin-Borkowski RE & Grillo V (2024). Symmetry and planar chirality measured with a log-polar transformation in a transmission electron microscope. *Phys Rev Appl* 22, 014083. <https://doi.org/10.1103/PhysRevApplied.22.014083>
- Tavabi AH, Rosi P, Roncaglia A, Rotunno E, Beleggia M, Lu P-H, Belsito L, Pozzi G, Frabboni S, Tiemeijer P, Dunin-Borkowski RE & Grillo V (2022). Generation of electron vortex beams with over 1000 orbital angular momentum quanta using a tunable electrostatic spiral phase plate. *Appl Phys Lett* 121, 073506. <https://doi.org/10.1063/5.0093411>
- Tavabi AH, Rosi P, Rotunno E, Roncaglia A, Belsito L, Frabboni S, Pozzi G, Gazzadi GC, Lu P-H & Nijland R (2021). Experimental demonstration of an electrostatic orbital angular momentum sorter for electron beams. *Phys Rev Lett* 126, 094802. <https://doi.org/10.1103/PhysRevLett.126.094802>
- Troiani F, Rotunno E, Frabboni S, Ravelli RBG, Peters PJ, Karimi E & Grillo V (2020). Efficient molecule discrimination in electron microscopy through an optimized orbital angular momentum sorter. *Phys Rev A (Coll Park)* 102, 043510. <https://doi.org/10.1103/PhysRevA.102.043510>
- Verbeeck J, B     A, M           K, Guzzinati G, Luong MA & Den Hertog M (2018). Demonstration of a 2×2 programmable phase plate for electrons. *Ultramicroscopy* 190, 58–65. <https://doi.org/10.1016/j.ultramicro.2018.03.017>
- Yu C-P, Vega Iba     F, B     A & Verbeeck J (2023). Quantum wavefront shaping with a 48-element programmable phase plate for electrons. *SciPost Phys* 15, 223. doi:10.21468/SciPostPhys.15.6.223

APPENDIX A

Harmonic constraint

In the absence of free charges, $\nabla^2 V = 0$. Here, we investigate the extent to which it is then possible to assume that

$$\nabla_{xy}^2 \varphi = 0. \quad (\text{A1})$$

The condition $\nabla^2 V = 0$ is equivalent to $\nabla_{xy}^2 V = -\frac{\partial^2 V}{\partial z^2}$. Ignoring constants,

$$\nabla_{xy}^2 \int V dz = \int \nabla_{xy}^2 V dz = \int -\frac{\partial^2 V}{\partial z^2} dz = -\frac{\partial V}{\partial z} \Big|_{z \rightarrow \infty}. \quad (\text{A2})$$

In most cases, this quantity vanishes because the fields and their derivatives vanish at large distances.

Any charge distribution can be written as a sum of multipoles of order j , where $j=1$ corresponds to a monopole, $j=2$ to a dipole, *etc.* The potential at a large distance along z can be approximated as $V \sim \frac{1}{z^j}$. For all cases apart from $j=1$, $-\frac{\partial V}{\partial z} \Big|_{z \rightarrow \infty} = 0$.

If there is a net charge, then the integral of the potential $V \sim \frac{1}{z}$ does not converge because of a weak logarithmic divergence,

which can be regularized by assuming that a compensating charge of opposite sign is present in the microscope column.

It should be noted that a net charge is responsible for an Einzel-lens-like behaviour that produces focusing. In fact, whether or not this focusing effect is important depends on the geometry/aspect ratio of the lens.

The focusing effect is a non-harmonic phase landscape (with $\nabla_{xy}^2 \rho^2 = 4$).

The z dependence of the potential can be difficult to control, as it depends on the details of the electrical contacts and ground of the column and of the micro-electromechanical system (MEMS) device that is used to apply the field, but it can be used to produce a more complex set of lenses. It is theoretically possible to encode even a three-dimensional stacking in a single element. However, in this paper, we assume that $\nabla_{xy}^2 \varphi = 0$.

Finally, since there are no magnetic monopoles, a similar situation can never occur for magnetic phase elements, which $\nabla_{xy}^2 \varphi = 0$ always holds.

APPENDIX B

A coordinate transformation is conformal if and only if, when expressed in complex notation, the function is holomorphic or anti-holomorphic.

Holomorphic and anti-holomorphic functions, which are defined by the property of being complex differentiable, can be regarded as compositions of two harmonic functions $q(g) = u(g) + i v(g)$ (i.e., $u = \text{Re}\{q\}$ and $v = \text{Im}\{q\}$) (Markushevich & Silverman, 1965). Partial derivatives of the real and imaginary parts, i.e., $u = \text{Re}\{q\}$ and $v = \text{Im}\{q\}$, must satisfy the Cauchy–Riemann (CR) conditions $u_x = \mp v_y$ and $u_y = \pm v_x$, where $u_{x,y}$ and $v_{x,y}$ are the x and y derivatives of u and v . The upper and lower signs are for the anti-holomorphic and holomorphic function, respectively.

The harmonicity condition for the functions u and v separately can be expressed equivalently by having their respective Laplacian equal to 0, i.e., $\nabla^2 u = 0$ and $\nabla^2 v = 0$.

The coordinate transformation is controlled, in the geometric phase approximation, by the phase of the first phase element φ_1 through the relation introduced by Hossack:

$$u, v = \frac{2\pi}{f\lambda} \nabla \varphi_1,$$

where f is the focal length of the diffraction lens used for the transformation. In complex terms, this is more elegantly written in terms of Wirtinger derivatives as

$$q = u + iv = \frac{2\pi \partial \varphi_1}{f\lambda \partial \bar{g}}.$$