Ultramicroscopy 166 (2016) 48–60

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic

Full length article

Generation and application of bessel beams in electron microscopy

Vincenzo Grillo [a,b,*], Jérémie Harris [c], Gian Carlo Gazzadi [a], Roberto Balboni [d], Erfan Mafakheri [e], Mark R. Dennis [f], Stefano Frabboni [a,e], Robert W. Boyd [c], Ebrahim Karimi [c]

[a] CNR-Istituto Nanoscienze, Centro S3, Via G Campi 213/a, I-41125 Modena, Italy
[b] CNR-IMEM, Parco Area delle Scienze 37/A, I-43124 Parma, Italy
[c] Department of Physics, University of Ottawa, 25 Templeton St., Ottawa, Ontario, Canada K1N 6N5
[d] CNR-IMM Bologna, Via P. Gobetti 101, 40129 Bologna, Italy
[e] Dipartimento di Fisica Informatica e Matematica, Università di Modena e Reggio Emilia, via G Campi 213/a, I-41125 Modena, Italy
[f] H.H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form
15 March 2016
Accepted 23 March 2016
Available online 26 March 2016

Keywords:
Bessel beams
Holographic beam shaping
Vortex beams
STEM
Fresnel diffraction

ABSTRACT

We report a systematic treatment of the holographic generation of electron Bessel beams, with a view to applications in electron microscopy. We describe in detail the theory underlying hologram patterning, as well as the actual electron-optical configuration used experimentally. We show that by optimizing our nanofabrication recipe, electron Bessel beams can be generated with relative efficiencies reaching $37 \pm 3\%$. We also demonstrate by tuning various hologram parameters that electron Bessel beams can be produced with many visible rings, making them ideal for interferometric applications, or in more highly localized forms with fewer rings, more suitable for imaging. We describe the settings required to tune beam localization in this way, and explore beam and hologram configurations that allow the convergences and topological charges of electron Bessel beams to be controlled. We also characterize the phase structure of the Bessel beams generated with our technique, using a simulation procedure that accounts for imperfections in the hologram manufacturing process.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Electron vortex beams have recently drawn significant attention within the electron microscopy community, and have shown great potential for a host of applications [1–3]. The OAM-carrying capacity of free electron beams was highlighted in a seminal theoretical paper by Bliokh [4], which precipitated considerable experimental efforts directed toward the generation of structured electron beams [5]. For example, electron vortex beams have recently been produced with orbital angular momenta as large as $200\hbar$ per electron; such beams show promise for potential applications in magnetic measurement [6]. For this reason, a great deal of effort has been expended in attempts to optimize the efficiency of vortex beam generation. In particular, holographic elements have emerged as promising candidates for high efficiency structured electron beam generation [7–12].

Holographic optical elements can allow electron beams to be shaped by modulating the transverse phase and amplitude profiles of incident electron waves with high precision. Amplitude modulation of incident electron beams can be achieved by alternating thick fringes made from opaque material with regions of high transparency. By contrast, phase modulation is carried out by varying the transverse thickness profile of a nearly transparent material, so as to produce disparities in the electron-optical path lengths experienced by different transverse components of the incident beam [8,9].

Phase-modulating elements have already found a range of applications in electron microscopy [13–15]. Specifically, phase plates can be used in transmission electron microscopy (TEM) to improve the contrast of weak phase objects, or to compensate for spherical aberration effects [16]. Attempts have also been made to produce phase plates for scanning transmission electron microscopy (STEM), in one case resulting in a Fresnel lens analogous to zone plate lenses for X-rays [17]. However, these types of lenses pose a significant nanofabrication challenge.

Beyond the examples mainly focused on vortex beams, relatively little work has been done with a view to shaping electron beams using holographic elements [8,11,12], and still less with reference to specific practical applications. This is not to suggest that this area is entirely unexplored; studies have previously investigated silicon nitride (Si_3N_4) as a candidate holographic material for electron beam shaping, for its low electron-optical density, and its ability to modify the beam phase directly on axis [18].

* Corresponding author at: CNR-Istituto Nanoscienze, Centro S3, Via G Campi 213/a, I-41125 Modena, Italy.
E-mail address: vincenzo.grillo@cnr.it (V. Grillo).

http://dx.doi.org/10.1016/j.ultramic.2016.03.009
0304-3991/© 2016 Elsevier B.V. All rights reserved.

However, no medium, no matter how transparent, can ever act as a perfect phase plate, since atoms in the material always produce inelastic or high-angle scattering that can, in essence, be treated as absorption and/or as loss of coherence, especially in the forward direction. This scattering, along with the limited control that can be exerted over the phase induced in an oncoming beam, can represent a significant hindrance to the use of on-axis phase holograms, producing a "frosted glass" effect, and blurring of the transmitted beam, and a reduction in its quality [19]. The use of Si_3N_4 holograms for on-axis electron beam shaping faces another drawback, in that it requires that thickness modulations be applied with precisions on the nanometer scale, a significant challenge even for state-of-the-art nanofabrication techniques.

In this sense, the introduction of off-axis amplitude holograms can be considered a significant development. These holograms, which consist of a modulated diffraction grating, benefit from the absence of unwanted scattering from their transparent regions by alternating fully absorbing and fully transparent fringes. A second advantage to this approach is that the phase imprinted on the incident wavefront is encoded in the transverse grating profile, and is therefore readily controlled, even when imperfect manufacturing techniques are employed. This technique does suffer from an important drawback, however, in that it typically results in low-efficiency generation of the desired output beam. Recently, we introduced off-axis phase holograms that allow this limitation to be overcome, theoretically reaching efficiencies as large as 100% [8,9]. Here, we report a detailed study of electron Bessel beam generation using this technique.

Bessel beams are widely used in photonics, and have recently been discussed theoretically in the context of a number of electron microscopy applications. In the ideal case, Bessel beams possess a propagation-invariant profile, and are therefore referred to as diffraction-free modes (see the discussion in Section 3). These beams hold great promise for their ability to reduce channeling [20], to control aberrations and their potential applicability to new imaging modes, as well as for the generation of optical tractor beams, and other exotic applications. Apart from their wide range of potential applications, Bessel beams have also drawn considerable interest on theoretical grounds, for their unusual properties [21].

Notably, electron beams of approximately Bessel form have been generated using on-axis techniques such as hollow cone illumination [22,23]. However, electron beams generated in this way suffer from large intensity losses due to the partial blocking of the beam required by the technique. Still more critically, this strategy does not allow for the modification or control of key beam parameters, such as topological charge and convergence.

Here, we report a detailed study of the first off-axis Fresnel phase hologram to generate electron Bessel beams [8], and examine: 1) the conditions under which Bessel beams can be generated and applied to microscopy and imaging; 2) techniques by which key beam and hologram parameters, including topological charge, transverse wavenumber, and hologram aperture radius can be adjusted; and 3) the main practical limitations of electron Bessel beam generation.

2. Holographic generation of structured electron beams

Holographic plates can be used to confer spatial structure upon arbitrary electron beams with high efficiency. These devices are fabricated by inducing spatially varying changes in the optical thickness and transmittance of a material, and therefore amount to optical phase and amplitude masks. When an incident plane wave is transmitted through such a mask, it gains a position-dependent phase $\Delta\varphi(\rho, \phi)$ relative to a reference wave having traveled an identical distance in vacuum, and experiences a spatial amplitude modulation $A(\rho, \phi)$, such that the mask may be described by a transmittance

$$T(\rho, \phi)=A(\rho, \phi)e^{i\Delta\varphi(\rho,\phi)} \tag{1}$$

where ρ, ϕ, z are the standard cylindrical coordinates. The transverse wavefunctions $\psi_{in}(\rho, \phi)$ and $\psi_{tr}(\rho, \phi)$, respectively corresponding to the incident and transmitted beams, are then related by $\psi_{tr}(\rho, \phi)=T(\rho, \phi)\psi_{in}(\rho, \phi)$. Three nontrivial classes of holograms may be distinguished, with reference to Eq. (1). First, *phase holograms* are those for which $\Delta\varphi(\rho, \phi)$ exhibits a spatial dependence, while the hologram's amplitude modulation function is spatially constant, i.e. $A(\rho, \phi)=A_0$. By contrast, *amplitude holograms* induce a spatially varying amplitude modulation, but produce a spatially constant phase in the incident beam, so that $\Delta\varphi(\rho, \phi)=\Delta\varphi_0$. Finally, *mixed holograms* are characterized by spatially varying phase and amplitude modulations, so that neither $A(\rho, \phi)$ nor $\Delta\varphi(\rho, \phi)$ is spatially constant for these masks.

In what follows, we shall restrict our attention to phase holograms, which may in general be associated with a transmittance $T(\rho, \phi)=A_0e^{i\Delta\varphi(\rho,\phi)}$. Physically, the phase modulation $\Delta\varphi(\rho, \phi)$ is induced in the incident beam due to the mean inner potential $V(\rho, \phi, z)$ of the material from which the holographic mask is constructed. This potential results in the addition of an energy term $eV(\rho, \phi, z)$ to the total Hamiltonian governing the time evolution of the electron beam in the material, resulting in a phase shift of the transmitted beam, relative to a reference wave having traveled the same distance in vacuum. From the general solution to the relativistically corrected Schrödinger equation, this phase shift is found to be

$$\Delta\varphi(\rho, \phi) = C_E \int_0^{t(\rho,\phi)} V(\rho, \phi, z)dz, \tag{2}$$

where $t(\rho, \phi)$ is the variation in the thickness of the hologram as a function of position in the transverse plane, and $C_E=\frac{2\pi e}{\lambda}\frac{E+E_0}{E(E+2E_0)}$ is a constant for a particular electron kinetic energy E, rest energy E_0, and λ de Broglie wavelength. In our case, the inner potential of the phase mask may be approximated by its mean value, V_0, such that [24,25]

$$\Delta\varphi(\rho, \phi) \approx C_E V_0 \int_0^{t(\rho,\phi)} dz = C_E V_0 t(\rho, \phi). \tag{3}$$

Hence, an arbitrary transverse phase profile can be imprinted on the incident beam, provided that variations in the local phase mask thickness $t(\rho, \phi)$ can be controlled with sufficient precision.

3. Generation and propagation of Bessel beams

We shall now focus our attention specifically on the generation of electron Bessel beams, which are described by scalar wavefunctions of the form

$$\Psi(\rho, \phi, z; t) = J_n\left(k_\rho\rho\right)e^{in\phi}e^{-i\left(E/\hbar t-k_z z\right)}, \tag{4}$$

where J_n represents an nth order Bessel function of the first kind, n is an integer, k_ρ and k_z are respectively the wavefunction's transverse and longitudinal wave vector components; $\hbar$ is the reduced Plank constant. These beams carry an amount of orbital angular momentum (OAM) along their propagation direction given by $L_z=n\hbar$ per electron, as indicated by the presence of a phase term $e^{in\phi}$ in the expression (4).

The generation of a Bessel beam necessarily entails imprinting a phase of the form $\Delta\varphi = \beta = k_\rho\rho + n\phi$ onto the incident wavefunction (see Appendix I). This is equivalent to imposing a conical wavefront on the electron beam [26], and can be achieved by

choosing a phase hologram with transmittance $T(\rho, \phi)=A_0 e^{i\beta}$. An additional grating term $k_x x = k_x \rho \cos\phi$, where $k_x=\frac{2\pi}{\Lambda}$ and Λ is a grating constant, can also be introduced to β for later convenience, so that

$$\beta = k_\rho \rho + n\phi + k_x x. \tag{5}$$

A functionally identical hologram, for which the imprinted phase becomes $\Delta\varphi = \mathrm{Mod}(\beta, 2\pi)$, where $\mathrm{Mod}(a, b)$ represents the remainder obtained when dividing a by b, would be equally well-suited to generating Bessel beams. We refer to this latter phase mask, in which $T(\rho, \phi)=A_0 e^{i\,\mathrm{Mod}(\beta,2\pi)}$, as a *blazed hologram*. Although blazed holograms are optimal from the standpoint of maximizing the efficiency of Bessel beam generation, they are difficult to produce in practice due to the finite resolution of existing fabrication techniques, which make use of a limited number of imprinted pixels to produce phase masks. As a result, the ideal blazed holograms must often be approximated by alternative configurations. In particular, by choosing the experimentally achievable phase imprint function $\Delta\varphi = \varphi_0 \cos(\beta)$, Bessel beams may be generated without prohibitively low efficiency. Phase masks of this form are referred to as *sinusoidal (or cosinusoidal) holograms*. From Eq. (3), we note that in this case $\varphi_0 \cos(\beta)=C_E V_0 t(\beta)$, so that in practice, these holograms can be produced by inducing (co-) sinusoidal variations $t(\beta)=\frac{1}{2}t_0 \cos\beta$ in the mask thickness, where t_0 is the peak-to-valley thickness of the holographic material. Sinusoidal holograms are characterized by transmittance functions of the form

$$T(\rho, \phi) = A_0 e^{i\varphi_0 \cos(\beta)}. \tag{6}$$

Hence, the wavefunctions associated with the incident and transmitted electron beams are related by $\psi_{tr}(\rho, \phi)=e^{i\varphi_0 \cos(\beta)}\psi_{in}(\rho, \phi)$. The Jacobi–Anger expansion may be applied to the exponential term to obtain $e^{i\varphi_0 \cos(\beta)}=\sum_{m=-\infty}^{\infty} i^m J_m(\varphi_0) e^{im\beta}$, where m is an integer, so that upon substitution of Eq. (5),

$$\psi_{tr}(\rho, \phi) = \psi_{in}(\rho, \phi) \sum_{m=-\infty}^{\infty} i^m J_m(\varphi_0) e^{im(k_\rho \rho + n\phi + k_x x)}. \tag{7}$$

For the case of a planar incident electron wavefunction of the form $\Psi_{in}(\rho, \phi, z;t)=e^{-i(E/\hbar t - k_z z)}$, we have $\psi_{in}(\rho, \phi)=1$, and therefore one obtains for the total transmitted wavefunction

$$\Psi_{tr}(\rho, \phi, z; t) = e^{-i(E/\hbar t - k_z z)} \sum_{m=-\infty}^{\infty} i^m J_m(\varphi_0) e^{im(k_\rho \rho + n\phi + k_x x)}. \tag{8}$$

Each term in the above expansion contains a component $e^{im\,k_x x}=e^{i\frac{2\pi m}{\Lambda}x}$, so that the transmitted wavefunction consists of an infinite number of diffracted beams, spaced apart at angles $\theta_m=m\frac{k_x}{k}$ where k is electron wavenumber. We refer to m as the order of diffraction, and note that the mth-order diffracted beam will carry an OAM of $mn\hbar$, and will be characterized by a conical phase front $\propto e^{im\,k_\rho \rho}$. Further, the transmitted electron beam will be split among the various diffraction orders, with the mth order receiving a fraction $|J_m(\varphi_0)|^2$ of the total transmitted intensity.

The most relevant example is that of the first diffracted order, for which the intensity is given by

$$I_1 = |J_1(\varphi_0)|^2. \tag{9}$$

This intensity is maximized for $\varphi_0 \approx 1.84$, at the first maximum of J_1. The corresponding peak to valley phase difference is 3.68, close to the value of π for which rectangular gratings are optimal.

In general, maximally efficient generation of the beam associated with the mth diffraction order would therefore require that a value of φ_0 be chosen such that $|J_m(\varphi_0)|^2$ be maximized. Immediately after the hologram (in the assumption of plane wave illumination and an unbounded hologram), the wavefunction associated with the mth diffracted beam takes the form $\Psi_{tr}^m(\rho, \phi, z;t) \propto e^{i\left(m(k_\rho \rho + n\phi) + k_z z - E/\hbar t\right)}$, where the z axis is now taken to lie along the propagation direction of the particular diffraction order in question.

While the electron wavefunction $\Psi_{tr}^m(\rho, \phi, z;t)$ does not take the form of a Bessel function immediately after the holographic mask, it can be shown (see Appendix I and Ref. [26]) to take on Bessel character within a range of propagation distances given by $z \le \frac{kR}{mk_\rho}$, where R is the radius of the aperture limiting the hologram. In this region, we have for the transverse wavefunction

$$\psi_{tr}^m(\rho, \phi) \approx N e^{i\left(k_z z - \frac{k_\rho^2}{2z} + mn\phi + \frac{m^2 k_\rho^2 z}{2k}\right)} J_{mn}(mk_\rho \rho) \tag{10}$$

where N is a normalization constant.

We may additionally consider the far-field electron wavefunction, which describes the beam after the hologram in the region $z \to \infty$. It can be shown (see Appendix II) that under these conditions, the probe intensity $I(K)$ assumes the form

$$I(K) \propto \left|\psi_{tr}^m(K)\right|^2 \propto \delta(K - mk_\rho) \tag{11}$$

where K is the spatial frequency.

Theoretical Fresnel (near-field) and Fraunhofer (far-field) intensities associated with a Bessel beam generated from a phase hologram are displayed in Fig. 1.

4. Properties of Bessel beams

Bessel beams of the form (4) are solutions to the scalar wave equation, which in vacuum is given by

$$-\frac{\hbar^2}{2M}\nabla^2 \psi(\rho, \phi, z; t) = -\frac{\hbar^2 k^2}{2M}\psi(\rho, \phi, z; t) \tag{12}$$

where M is the electron mass and $\hbar$ is the reduced Plank constant.

This can readily be observed by substituting a trial solution in cylindrical coordinates of the form $\Psi(\rho, \phi, z;t)=R_n(\rho)e^{i(n\phi + k_z z - Et/\hbar)}$ into Eq. (12), whence we find that

$$\rho^2 \frac{d^2 R_n(\rho)}{d\rho^2} + \rho \frac{dR_n(\rho)}{d\rho} + \rho^2\left(k^2 - k_z^2 - \frac{1}{\rho^2}n^2\right)R_n(\rho) = 0, \tag{13}$$

which has solution $R_n(\rho)=J_n(k_\rho \rho)$, where $k_\rho^2=k^2-k_z^2$ [27]. It then follows that $\Psi(\rho, \phi, z;t)=J_n(k_\rho \rho)e^{i(n\phi + k_z z - E/\hbar t)}$, in agreement with (4). We note also that the transverse amplitudes of Bessel beams, $R_n(\rho)$, are independent of the beam propagation distance z. For this reason, Bessel beams are referred to as *non-diffracting* beams [28,29]. Despite their attractive physical properties, Bessel beams of the form (4) are not normalizable, carry infinite energy, and are therefore unphysical. Nonetheless, they can be closely approximated in practice, as we shall see. In Fig. 2, we illustrate the non-diffractive propagation of an ideal Bessel beam, along with its propagation range, $z_{\max}$. It can be seen from the figure that the hologram convergence angle $\alpha \equiv k_\rho/k$ and size R jointly determine the length of the region over which the electron beam will take on Bessel character, due to the overlap between its component plane-waves.

5. Simulation of electron beam propagation

Beam propagation simulations were carried out numerically using STEM_CELL software [30], which allows electron beam

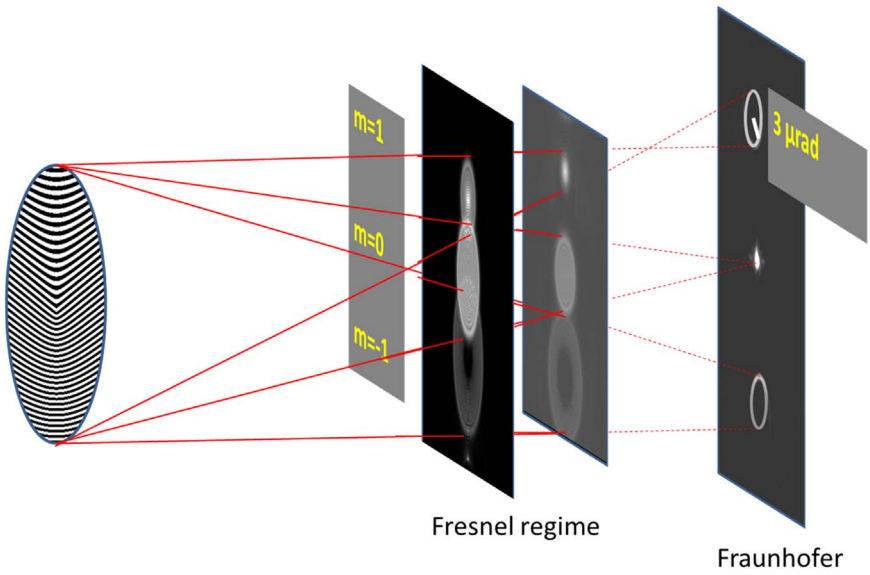

Fig. 1. Theoretical Fresnel (near-field) and Fraunhofer (far-field) intensity distributions associated with a Bessel beam generated from an off-axis phase hologram. The Bessel beam itself is formed at the first diffracted order ($m = 1$), and is found to take a ring-like shape in the far-field, in accordance with Eq. (11). The red lines indicate a schematic (i.e non-rigorous) wavefront evolution. In a free space propagation scheme the Fraunhofer plane should be located at $+\infty$. If the images are produced by a lens, the Fraunhofer plane is located at a finite distance, namely the focal length. The Fresnel diffraction regime, in which the beam takes on its smallest size, is the region in which the formation of the Bessel beam can be observed. The illustration also demonstrates that for the upper/lower beam the Bessel condition is reached before/after the focal plane.

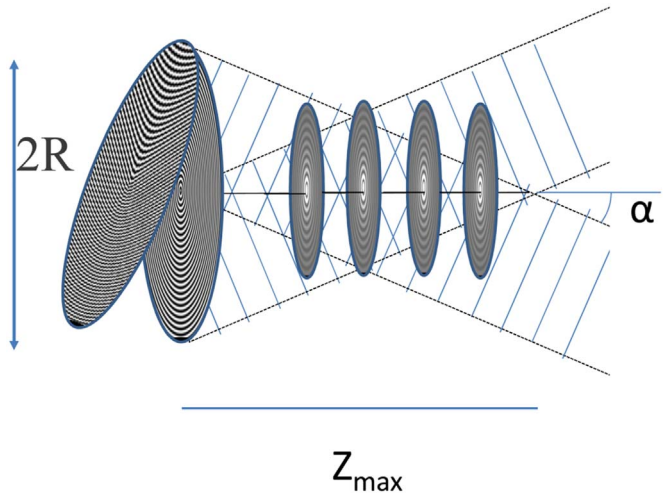

Fig. 2. Theoretical depiction of diffraction-free propagation of an idealized electron Bessel probe. Both the off-axis and on-axis hologram are displayed: they differ only by an in-plane carrier frequency (i.e. a tilt). For the on-axis hologram, one can readily discern a similarity to a parabolic Fresnel lens. Due to the conical phase imparted by the hologram (see Eq. 5), wavefronts at different azimuthal angles converge with an angle $\alpha \equiv k_\rho/k \approx k_\rho/k_z$. Since the hologram is laterally bounded by the hologram radius R, by neglecting diffraction effects at the aperture, we can imagine each wavefront as being "laterally bound plane waves" whose limits are the beam center and the radius of the aperture. These waves overlap only in a limited region, $z \leq \frac{kR}{k_\rho}$. The beam retains its Bessel form and diffraction-free characteristics only in this region of overlap. Finally, note that different diffraction orders are characterized by different radial phase (see Eq. (8)) gradients, so that in general we can write $z_{max}(m) = \frac{kR}{mk_\rho}$.

wavefunctions to be deduced based on our experimentally constructed hologram thickness maps. The electron wavefunction could then be calculated at different propagation distances by making use of the relation [31]

$$\psi(x, y; z + \Delta z) = P(x, y; \Delta z) \otimes \psi(x, y; z), \tag{14}$$

where $P(x, y; \Delta z)$ is the Fresnel propagator, which describes the beam's evolution over a distance Δz, and $\psi(z)$ is the electron wavefunction at position z, which serves as a pupil function in the context of the Fresnel propagation integral, and $\otimes$ represents the x, y convolution operation (see Appendix I). In the paraxial approximation, the propagator takes the form $P(\rho; \Delta z) = -\frac{i}{\lambda \Delta z} e^{\frac{i\pi\rho^2}{\lambda \Delta z}}$.

In practice, the electron wavefunction incident on the hologram is not perfectly collimated. For this reason, the aperture function $\psi(z)$ describing the incident beam is characterized by a slightly convergent wavefront. This requires that numerical simulations be carried out with a pixel size significantly smaller than the electron beam diameter.

We note that much of the blurring observed in the Bessel beams generated experimentally was due to the limited transverse coherence length of the source, brought upon by the finite size of the FEG Schottky emitter. This coherence length depends on the demagnification of the source at the the focal plane of the objective lens prefield. We accounted for limitations in transverse beam coherence by considering the Fresnel diffraction zone to be described by many mutually incoherent beams, each of which is characterized by a slightly different incidence angle [31]. Losses due to inelastic scattering, especially those due to the excitation of plasmons, have not been considered. Inelastically scattered waves have a spatial distribution resembling that of the grating in the near field, but lack coherence over the length scale of our grating. At a sufficient propagation distance from the holograms, these effects merely contribute a diffuse, incoherent background extended over several milliradiants, an angular size much wider than that of the diffraction grating.

Our experiments did not reveal any feature in the energy filtered grating diffraction except those coming from zero loss region. In general, however, one cannot ignore the role of inelastic scattering in modulating the amplitude of the transmitted wave: strictly speaking, our holograms are amplitude and phase holograms.

From this work, it is therefore clear that the generation of truly propagation-invariant Bessel beams is limited in efficiency by three considerations. First, Bessel beams generated in the laboratory are characterized by intensity oscillations at beam center throughout propagation, due to diffraction from the hologram aperture. Second, the limited range of applicability of the approximation scheme used to derive the near-field electron

wavefunction Eq. (10) predicts the breakdown of Bessel-like behavior at some maximal propagation distance, z_{max}. Indeed, well beyond this point, the wavefunction takes its far-field form Eq. (11), and loses all Bessel character. Finally, imperfections in hologram patterning can result in non-ideal, pseudo-Bessel beams. Great care must therefore be taken to ensure that an optimal hologram design is chosen, so as to produce high-quality beams.

6. Hologram patterning

TEM experiments were primarily performed using a JEOL 2200FS microscope, equipped with a Schottky field emission gun (FEG), operated at 200 keV. The hologram was inserted in the microscope's sample position, and beam images were obtained under low magnification, using the objective minilens as a Lorentz lens. This allowed for a large camera length and focal range, permitting imaging from the Fresnel to the Fraunhofer planes. This working mode, and the Fresnel mode in particular, are not calibrated in our microscope. As a result, we implemented a manual calibration scheme. The microscope was equipped with an Omega filter for energy loss imaging, and used to map hologram thickness profiles.

For STEM experiments, the hologram was mounted in the second condenser aperture of a FEI Tecnai TEM equipped with a Schottky FEG, and operated at 200 keV. A Dual-Beam instrument (FEI Strata DB235M), combining a focused gallium-ion beam (FIB) and a scanning electron microscope (SEM), was used to pattern the holograms by FIB milling 200 nm-thick silicon-nitride membranes coated with a 120 nm-thick gold film. The membranes were coated with the gold film in order to prevent electron transmission in all but the patterned areas.

We note that, at an accelerating voltage of 200 keV, a 120 nm gold layer is not sufficient to completely stop the electron beam used. However, the presence of the gold film does suffice to induce elastic scattering of electrons. The mean free path for scattering absorption associated with diffuse scattering is on the order of 60 nm while the extinction distance of the main diffractions occurs on the order of some tens of nanometers. Fortunately, the diffraction angles we explore are on the order of μrad (the Bragg angle for a grating with 100 nm step spacing), so that almost any scattering event produces a deviation from the angular range of interest. We found that, in practice, some detectable intensity was transmitted in the forward direction only when the beam was completely concentrated in one point. We experimentally determined the undesired forward transmittance to be well below 1%.

The procedure for hologram nanofabrication is implemented by starting with a bitmap picture of a computer-generated hologram, which is converted into a FIB pattern file containing three key pieces of information. These are respectively the pixel coordinates at which the FIB is switched on, the beam dwell time on each pixel, and the repetition number of the whole coordinate set, adjusted in such a way as to obtain the desired milling depth [9].

The second step is to adjust the FIB magnification according to the desired dimensions of the hologram. We selected a 50 nm width, and 100 nm periodicity for the stripes composing the hologram, resulting in a typical full hologram size on the order of 10 μm × 10 μm.

Once the computer-generated hologram has been designed, the holograms are patterned in two stages: first, the gold layer is uniformly removed from a circular region, 10 μm in diameter. To this end, the transmitted intensity from the secondary electron beam is monitored during milling until the signal is lost, indicating that the gold is no longer present. Next, the hologram pattern is superimposed on the uncovered region, and milled into the silicon nitride.

For reasons related to the finite pixel resolution accommodated by our software, we imprinted the ideal, blazed profile only onto holograms with large grating periods, and nearly sinusoidal profiles onto those with grating periods under ~300 nm. In order to control the experimental hologram thickness profile, we performed TEM energy loss analyses. Through imaging, and by comparing beam transmission spectra, we generated quantitative maps of sample thickness.

The result of this procedure is shown in Fig. 3, where we aimed to generate a sawtooth hologram profile. The inset shows that the thickness profile indeed corresponds closely to that of a blazed hologram. We can define the relative (or exit) efficiency η of the hologram as follows:

$$\eta = \frac{I_{m=1}}{\sum_m I_m}, \tag{15}$$

where I_m represents the intensity associated with the mth diffraction order. We note that this definition of efficiency differs from more canonical definitions, in that it explicitly considers beam intensities I_m *after* transmission through the hologram, rather than providing the ratio of desired beam intensity to the intensity of the beam incident on the hologram aperture [11]. While these two definitions coincide in the limit of a strictly non-absorbing hologram, they will not agree in general, and from the known absorption of Si_3N_4 were estimated to differ roughly by a factor of two to three in our experiment. This disagreement may be understood to arise from loss of the coherent fraction of the beam intensity due to the inelastic scattering of electrons by the hologram.

Using this groove thickness profile, we can plot the hologram's efficiency as a function of the peak-to-valley thickness of the holographic material, t_0, from which we can see (Fig. 3-c) that this profile allows a maximum efficiency of 38%. We obtained an efficiency of 37%, which is presently the best performance achieved by such a device, given that an uncertainty of about 3% must be allowed in order to account for the unknown intensity of the beams outside the field of view.

This also indicates that it is not possible to further increase the efficiency of this nanofabrication recipe; greater control of the groove profile is therefore necessary, but lies outside the scope of this work. Decreasing absorption effects will certainly be important to further progress, but we note also that the relative efficiency is an important factor for STEM applications, as a relative efficiency approaching 100% would mean a perfect suppression of all unwanted diffraction orders.

7. Results and discussion

In presenting the data, we distinguish between two classes of hologram, based upon their respective aperture radii R. This parameter determines the extent to which the electron probe will resemble an ideal Bessel beam. Large aperture radii allow for the generation of highly Bessel-like beams in the Fresnel region, whereas reductions in R lead to a decrease in the number of visible rings associated with the electron beam, all else being equal. It can also be shown, based on the uncertainity principle, or on the maximum propagation range, that the aperture radius is inversely proportional to the width of the transverse momentum distribution ΔK, such that $\Delta K(2R) \approx 1$. It is useful to compare with hollow cone illumination.

Thus, holograms with large apertures tend to produce ideal, delocalized Bessel beams suitable for interferometry, while smaller aperture (more precisely, apertures with small values of $k_\rho R$) holograms generate highly localized beams that are best suited to STEM imaging.

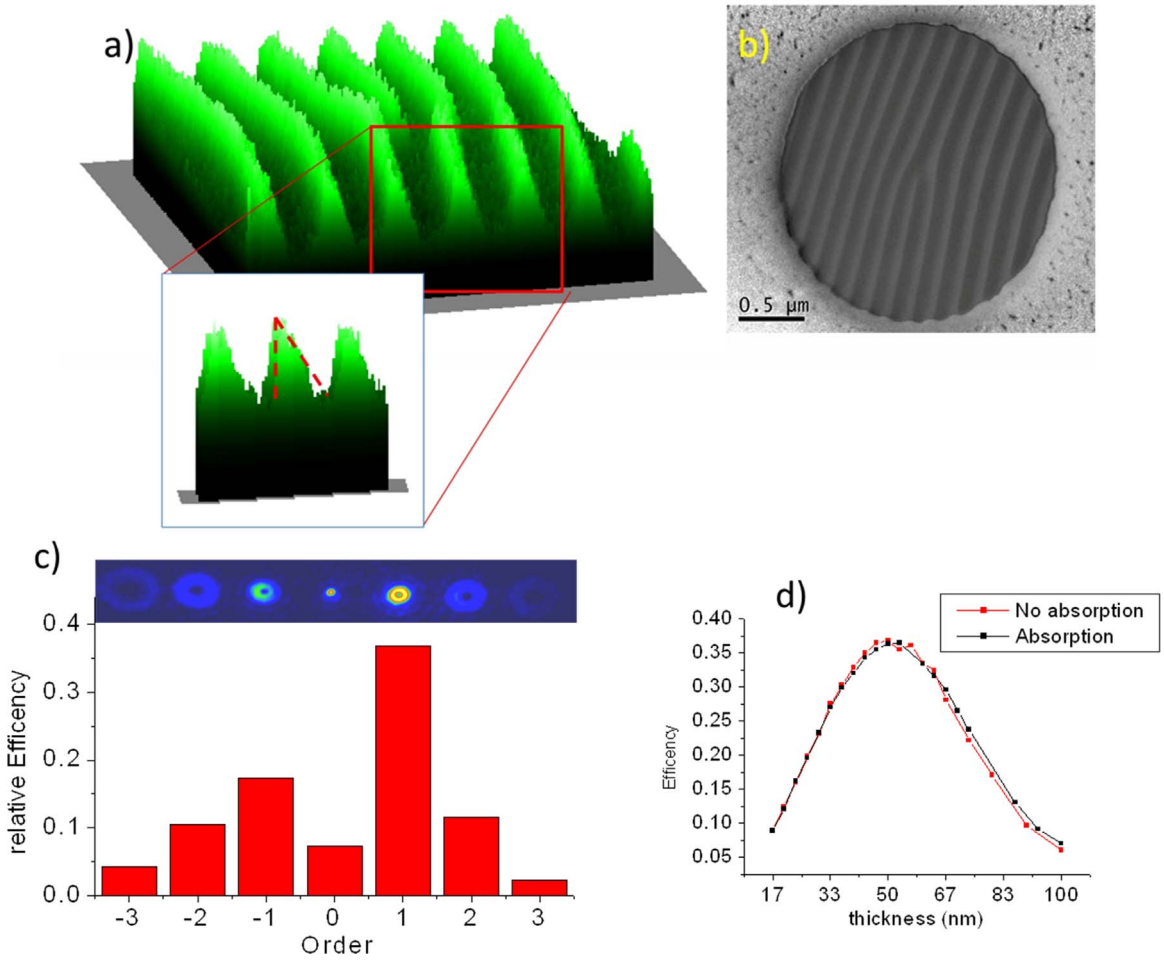

Fig. 3. Experimental hologram patterning. (a) Three-dimensional rendering of an energy filtered TEM-based thickness map of the center of a hologram with parameters $n = 1$, k_ρ=3.2 × 10^{-5} Å^{-1}, and $R = 1.22$ μm. (b) SEM image of the same hologram. (c) Experimental diffraction pattern obtained from the hologram displayed in (b). Below it an Histogram shows the intensities of each of the diffraction orders, (d) Simulation of hologram relative efficiency as a function of thickness scaling factor t_0. Simulations were performed with and without considering the effect of amplitude modulation: the amplitude (absorption) modulation is shown to have a negligible effect. The figure shows the expected trend, the efficiency exhibiting a dependence roughly of the form $|J_1(\varphi_0)|^2$ (see Eqs. (8) and (9)) on the thickness, where $\varphi_0 \propto t_0$ (see the discussion preceeding Eq. (6)). For the thickness profile considered, a peak-to-valley thickness of 50nm is found to result in a maximal efficiency of 38%. This limit is better than the maximum of $|J_1(\varphi_0)|^2$ due to the triangular thickness profile. We determined the efficiency of the hologram experimentally to be 37 ± 3%. The uncertainty on this efficiency accounts for the estimated contribution of beams at higher diffracted orders.

Of course, in STEM, the need for optimal probes must be balanced with the need for sufficient intensity. Therefore, one must optimize the optical configuration, including the hologram convergence, and efficiencies must be strongly enhanced to be competitive with normal probes. This will be the object of forthcoming work.

We note that, for a fixed k_ρ and large aperture R, the first-order diffracted beam will closely approximate a Bessel beam, whereas for smaller R the hologram will predominantly act as a pinhole, resulting in significant overlap between the zeroth and first-order diffracted beams. We note also that, in the Fresnel regime, increases in aperture size do not increase the convergence of the generated beam.

7.1. Bessel beams with large aperture radii

Fig. 4-a and b shows two holograms, characterized by respective hologram convergence k_ρ/k of 6 μrad and 15 μrad, and large, identical aperture sizes. Fig. 4-c and d shows the corresponding Bessel-like beams generated from these holograms in the Fresnel region, when they are illuminated by approximately collimated incident electron beams. The holograms were prepared with $n=0$, and therefore impart no OAM to the transmitted electron beams. Both holograms were 10 μm in diameter and contained 100 grating lines.

The Bessel beams shown in Fig. 4-c and d reveal the critical role played by the radial wavenumber k_ρ in defining the spread and number of visible fringes in the transmitted beams. For holograms with smaller values of k_ρ, the first-order diffracted beams are subject to relatively insignificant spreading during propagation, and the Bessel beams generated from these masks are therefore readily isolated from the zeroth diffracted order. By contrast, holograms manufactured with larger k_ρ produce strongly divergent transmitted beams, resulting in significant overlap between the zeroth and first orders of diffraction, though this overlap can be reduced by increasing the main separation k_x. Indeed, the extent of this overlap can be so significant that the isolation of the first diffracted order from the zeroth order becomes challenging (Fig. 4-d). This overlap also results in the apparent deformation of the first-order diffracted beam at its center. Holograms manufactured with small k_ρ are also found to produce Bessel beams with fewer rings than would be the case for those manufactured with larger transverse wavenumbers, as expected theoretically. Hence, for a given aperture size, an

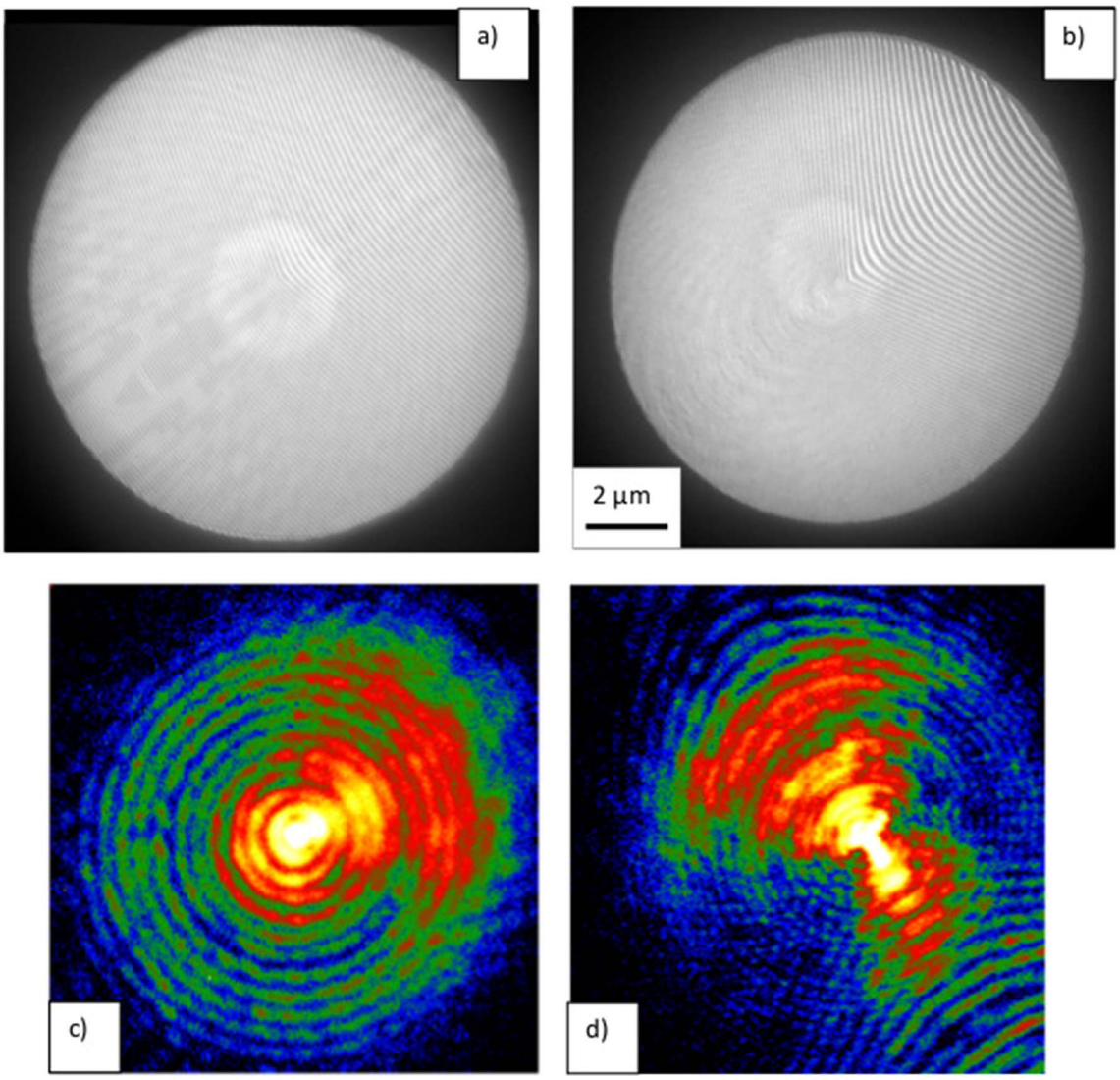

Fig. 4. Bessel beam generation by large aperture phase holograms. (a) Scanning electron microscope (SEM) image of a phase hologram with aperture radius $R=5$ μm and convergence angle $\alpha=6$ μrad. (b) SEM image of a phase hologram with aperture radius $R=5$ μm, and large convergence angle $\alpha=15$ μrad. In both images a and b, a small contamination area is visible about the center. (c) Near-field intensity pattern obtained experimentally from the hologram depicted in part (a). (d) Near-field intensity pattern obtained from the hologram depicted in part (b). Notice that up to 20 rings of intensity are visible.

increase in k_ρ will result in a more Bessel-like electron beam in the Fresnel near-field, with a greater number of visible fringes.

7.2. Bessel beams with small aperture radii

For comparison, we show in Fig. 5a series of holograms manufactured with smaller aperture radii, along with corresponding intensity profiles for the first diffracted orders of the transmitted electron beam. In the figure, we compare the cases $n = 0,1,2$. In each case, the holograms were manufactured with a hologram convergence $\alpha = k_\rho/k$ of 1 μrad. Notably, in the case of $n = 2$, we reach a relative efficiency of almost $37 \pm 3\%$, which is by far the largest value ever achieved for off axis holograms. Large relative efficiencies are particularly important in order to remove the effect of overlap with other beams.

Under these conditions, the beam consists only of a very faint ring about the beam center, and its shape depends strongly on propagation distance. This can be understood to occur as a consequence of the small hologram aperture, which does not allow higher order fringes to manifest themselves upon propagation, resulting in a beam with almost no Bessel character. Such beams produce novel probes well suited to STEM imaging, owing to the small size of their central intensity maximum (the 0th-order Bessel beam is characterized by the smallest central spot size among all beams with a given numerical aperture [28,29]), which results in a beam localization that is only slightly inferior to that of an aperture-limited probe.

8. Propagation

In order to characterize the effective propagation range of the Bessel beams generated using our technique, we examined the intensity at beam center for the first diffracted order, in the case $n = 0$, i.e. for an electron beam carrying zero OAM. The holograms used in this experiment featured large aperture radii, and resembled the holographic mask shown in Fig. 3-a. The intensity values thus obtained are shown as a function of propagation distance in Fig. 6, along with theoretical plots obtained from simulations.

Our results indicate that, apart from some oscillations, the beam intensity rises to a maximum value at $z_{max}=0.7$ m. This is

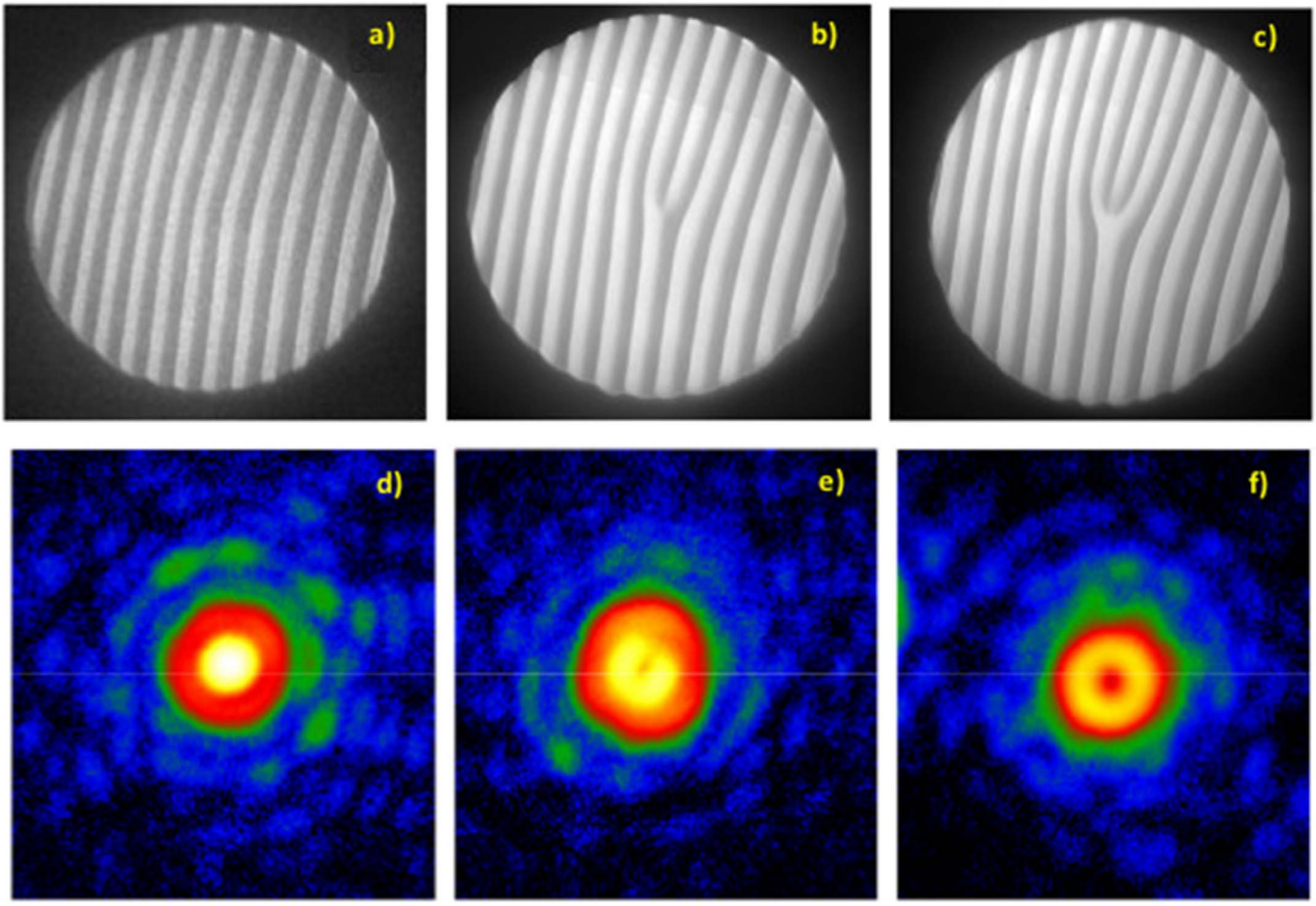

Fig. 5. Beam generation by small-aperture phase holograms. a, b, and c show in-focus bright-field images of phase holograms with small aperture radii $R=1.22$ μrad, convergence angles $\alpha=1$ μrad and respective topological charges $n=0$, 1 and 2. d, e and f show the experimental intensity patterns obtained from these respective holograms. Notably, the hologram with $n=0$ gives rise to a single, well-defined point of maximum beam intensity, whereas higher topological charges lead to doughnut-shaped intensity patterns.

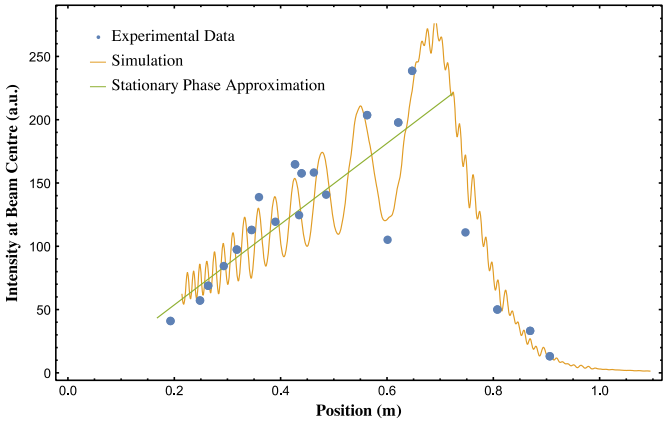

Fig. 6. Intensity at beam center as a function of propagation distance. Simulated (orange curve) and experimental (blue points) intensities at beam center as a function of position for an electron Bessel probe with parameters $n=0$, $\frac{k_\rho}{k}=6\mu rad$, and $R=5\mu$m. Some of the experimental points correspond to images in [8] and Fig. 4c.

consistent with the leading linear factor in stationary phase approximation (SPA), (described in Appendix I) In fact, the intensity's z-dependence is shown to be dominated by an initial, linear increase in average intensity with z, and a smooth decrease after z_{max} (see Appendix I), corresponding closely to the theory and experimental plots displayed in Fig. 6.

Notice that the cutoff at z_{max} should be abrupt according to the simple geometrical scheme in Fig. 2, but diffractive effects blur the wavefronts, producing the smooth decrease observed.

We note also that it can be difficult to identify the plane at which the Fraunhofer condition is satisfied when carrying out experiments involving small aperture radii R. By definition, the Fraunhofer plane is the position at which the 0th diffraction order of the transmitted beam is most tightly focused. However, when R is small, it is in practice difficult to clearly identify the zeroth-diffracted order in beam cross-section images obtained experimentally. Further complicating matters, each diffraction order is focused at a different position, so that an unambiguous identification of the Fraunhofer plane is challenging to achieve. Notwithstanding these limitations, techniques have been developed that allow the zeroth diffracted order to be identified, by deliberately introducing a condenser astigmatism to the beam, as reported in reference [32].

9. Phase description

Since vortex beams are most completely described by referring to their transverse phase structure, a great deal of emphasis has been placed on the development of techniques that might allow for the retrieval of phase information from such beams [33]. For our purposes, a realistic reconstruction of the phase of the electron

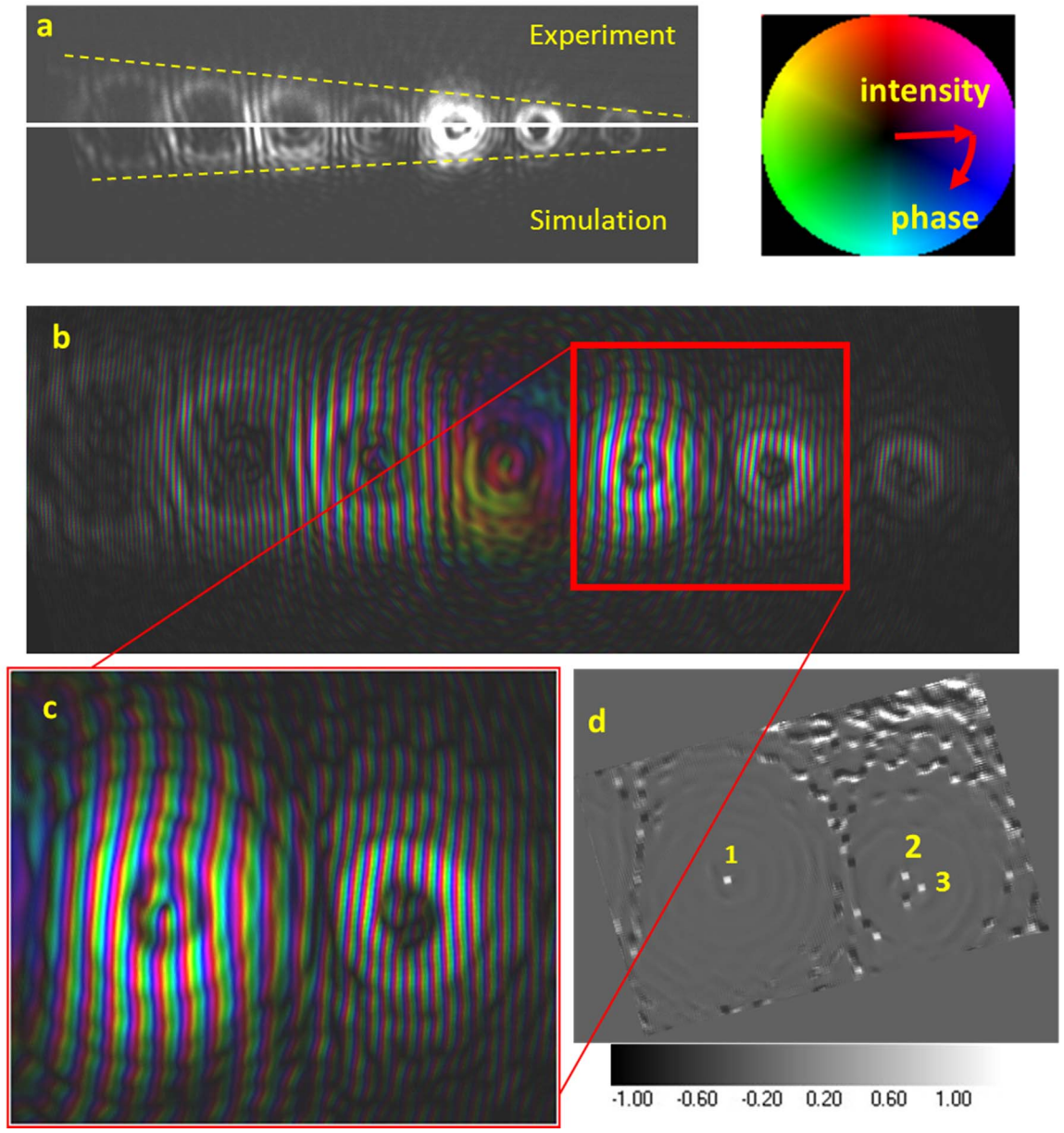

Fig. 7. Correspondence between experimental and theoretically calculated beam phase structures. (a) Simulated and experimental beam propagation, showing agreement at $z=0.1$ m. From this propagation distance and the known hologram profile, it is possible to reconstruct the phase structure of the beam. (b) Orders of diffraction obtained from a phase hologram with parameters $n = 1$, $\frac{k_p}{k} = 1$ μrad, and $R=1.22$ μrad, with superimposed phase structures. In the figure, the beam phase is indicated by color hue, and intensity by brightness. (c) Enlarged view of the first and second diffracted orders shown in part (b). (d) Reconstruction of the positions of phase vortices in the original beam (see Appendix III). The area about the indicated singularities is typically 30–100 times less intense than the rest of the map. Indeed, the reconstruction is based on a hologram thickness map with low noise, and is therefore quite robust even where the diffracted beam intensity is low near the singularity.

beam can be achieved from calculations based on experimentally measured hologram profiles. Since the wavefunction of a transmitted electron beam can be determined from the hologram thickness profile, the beam structure can be calculated at any propagation plane, using the techniques discussed earlier.

In a previous study [9], we demonstrated that when beam coherence effects are accounted for, a very good agreement exists between the modeled electron wavefunction, and the beam's experimentally observed intensity distribution. Thus, this technique provides an initial, indirect means by which to characterize the transmitted electron beams. Transverse intensity and phase profiles calculated for an electron beam carrying an OAM of $n = 1$, generated by a small aperture, are shown in Fig. 7. Given that the intensity pattern calculated for the beam corresponds closely to those obtained experimentally, we assume that the calculated phase distribution represents an accurate picture of the beam phase structure as well.

We also carried out a simple phase analysis, analogous to that reported in [34], to locate beam phase vortices. Our results show that, in the case $n = 2$, in Fig. 5c the second-order vortex decomposes into two separate vortices of first order, as predicted in [35]. This observation cannot be ascribed to lensing effects, owing to the fact that this phenomenon is not accounted for by our simulation technique. Rather, we believe this decomposition to arise from imperfections in the grating [36,37].

If only the OAM content of the first-order beam is of interest, a more direct characterization of the first diffracted order can be achieved by interfering the first-order diffracted beam with the

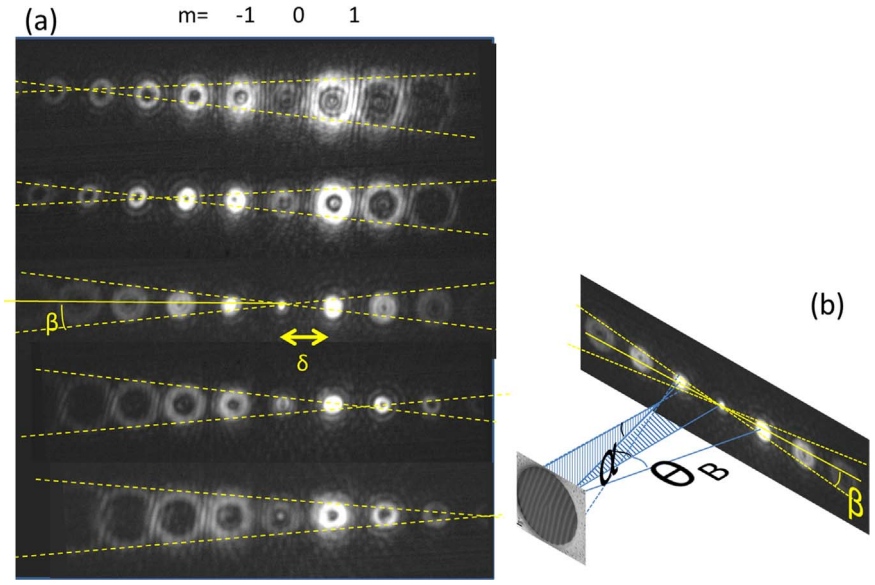

Fig. 8. (a) Propagation and focal characteristics of various diffracted orders. Experimentally obtained intensity profiles associated with various orders of diffraction, which are visibly focused at different propagation distances. The anticipated linear dependence of beam size on diffraction order is verified by pasting together transverse the intensity profiles for each order at identical propagation distances. Notice that the lateral cutoff on the beam size does not enter into play when the beam is very small. (b) Illustration of the scattering geometry for different diffraction orders.

zeroth-order as a reference. The resulting pitchfork-shaped interference pattern produces a vortex dislocation that indicates the OAM content of the first-order beam.

In Fig. 8, we show beam cross-sectional images obtained for several diffracted orders at various effective propagation distances about the $m = 0$ order focal point. As can be gathered from the figure, every diffracted order is found to focus at a different location. Further, the sizes of the diffracted beams are found to depend linearly on the indices m of the respective diffracted orders (a consequence of the conical shape of the beam), in agreement with the anticipated range of validity of Eq. (10) (See Appendix I), $\rho < \pm R - mk_\rho z/k$.

It can be readily be seen for the Fraunhofer condition that the angle β is related to the beam convergence $\alpha = k_\rho/k$ through the angular separation of the order, θ_B (proportional to δ in Fig. 8), so that $\beta\theta_B=\alpha$.

10. Conclusion

We have explored the theory of Bessel beam holographic generation in detail, examining the impact and importance of hologram parameters such as the groove shape and depth, aperture size, fringe spacing and modulation. By optimizing these parameters, we have experimentally achieved Bessel beam generation with efficiencies as high as $37 \pm 3\%$. Moreover, we have demonstrated experimentally the successful generation of Bessel beams characterized by variable transverse wavenumbers, topological charges and ranges of non-diffractive propagation through direct measurement and observation of beam structure. We believe that this systematic study will greatly facilitate the application of Bessel beams to imaging and electron microscopy.

Appendix I. Fresnel propagation of diffracted electron beams

See Fig. A1.

The wavefunctions associated with each diffraction order will evolve through free space propagation, beyond the holographic plate. The effect of this propagation can be calculated by evaluating the Fresnel diffraction integral

$$\Psi_{tr}^{m}(x, y, z) = \frac{e^{ikz}}{i\lambda z} \iint dx' dy' A_m(x', y') e^{-i\frac{k}{2z}\left((x-x')^2+(y-y')^2\right)}, \quad \text{(A1)}$$

where $A_m(x', y')$ is the aperture function, which describes the phase and amplitude modulation induced in an incident beam by the holographic phase mask. With respect to the mth diffraction order, the aperture function will take the form $A_m(x', y') \rightarrow A_m(\rho', \phi') = e^{im(k_\rho \rho' + n\phi')}$. Hence, the diffraction integral becomes

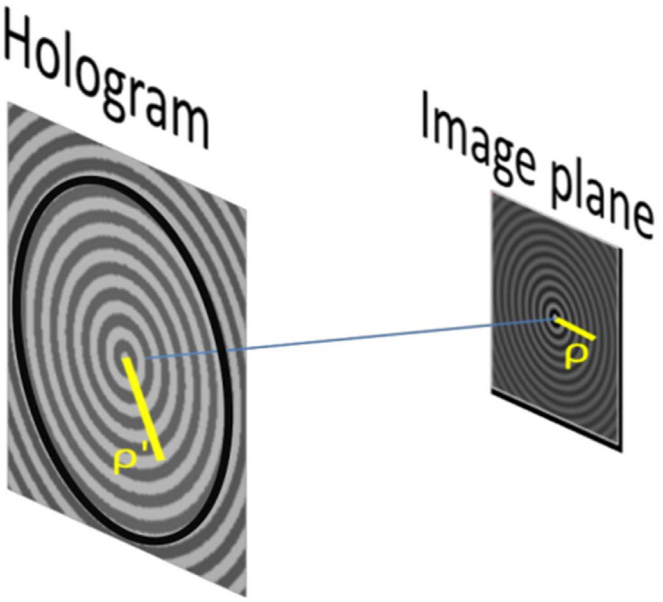

Fig. A1. Schematic representation of a Bessel phase hologram and beam image plane for the purpose of determining near-field and far-field beam profiles. Cylindrical coordinates are used to indicate points in both planes. Primed coordinates refer to the hologram plane, while unprimed coordinates refer to the image plane.

$$\Psi_{tr}^{m}(\rho,\phi,z) = \frac{e^{ik(z-\rho^2/2z)}}{i\lambda z}\int_0^R d\rho'\rho' e^{i\left(mk_\rho\rho'-k(\rho')^2/2z\right)}\int_0^{2\pi} d\phi' e^{imn\phi'}e^{ik(\rho\rho'/z)\cos(\phi-\phi')},$$

where the limits of the outer integral reflect the finite aperture of the holographic mask. Applying the Jacobi–Anger expansion to the integral over ϕ yields

$$\int_0^{2\pi} d\phi' e^{imn\phi'}e^{ik(\rho\rho'/z)\cos(\phi-\phi')} = \int_0^{2\pi} d\phi' e^{imn\phi'}\sum_{p=-\infty}^{\infty} i^p J_p\left(k(\rho\rho'/z)\right)e^{ip(\phi-\phi')}$$

And with $\int_0^{2\pi} d\phi' e^{i(mn-p)\phi'} = 2\pi\delta_{mn,p}$, this gives

$$\int_0^{2\pi} d\phi' e^{imn\phi'}e^{ik(\rho\rho'/z)\cos(\phi-\phi')} = 2\pi i^{mn}e^{imn\phi}J_{mn}\left(k(\rho\rho'/z)\right)$$

So that we have

$$\Psi_{tr}^{m}(\rho,\phi,z) = -\frac{2\pi i^{mn+1}}{\lambda z}e^{ik\left(z-\frac{\rho^2}{2z}\right)}e^{imn\phi}\int_0^R d\rho'\rho' e^{i\left(mk_\rho\rho'-\frac{k(\rho')^2}{2z}\right)}J_{mn}\left(k\left(\frac{\rho\rho'}{z}\right)\right) \quad (A2)$$

We can try to find an approximate solution in a range where $k\left(\frac{\rho\rho'}{z}\right)$ is small, for example close to the axis. The Bessel function J_{mn} can also be assumed to vary slowly.

With these assumptions in hand, the integral (A2) becomes tractable. In order to evaluate it, we turn to address a related problem, namely that of calculating the integral

$$\int_0^R F(x)e^{ig(x)}dx, \quad (A3)$$

where $g(x)$ is taken to represent a function with a single extremum, at $x = x_c$, such that $\frac{d}{dx}g(x_c)=0$. We now assume that the exponential term $e^{ig(x)}$ oscillates rapidly with x, except at x_c, where the derivative of $g(x)$ vanishes. As a result, only the extremum x_c of $g(x)$ will contribute materially to the integral (A3). The integral may therefore be approximated by taking $g(x)\approx g(x_c)+\frac{1}{2!}\frac{d^2g(x)}{dx^2}\Big|_{x=x_c}(x-x_c)^2$, and $F(x)=F(x_c)$, in accord with the stationary phase approximation, so that

$$\begin{aligned}\int_0^R F(x)e^{ig(x)}\downarrow dx &\approx e^{ig(x_c)}F(x_c)\int_0^R e^{\frac{i}{2!}\frac{d^2g(x)}{dx^2}\Big|_{x=x_c}(x-x_c)^2}dx\\ &= e^{ig(x_c)}F(x_c)\int_{x_c}^{R+x_c} e^{\frac{i}{2!}\frac{d^2g(x)}{dx^2}\Big|_{x=x_c}x^2}dx\\ &= e^{ig(x_c)}F(x_c)\int_{x_c}^{R+x_c} e^{i\alpha x^2}dx,\end{aligned} \quad (A4)$$

where $\alpha\equiv\frac{1}{2}\frac{d^2g(x)}{dx^2}\Big|_{x=x_c}$.

We note that the integral (A4) is mathematically equivalent to the expression (A2) that we wish to evaluate, if we choose $F(x)=xJ_{mn}(\gamma x)$, with $\gamma\equiv\frac{k\rho}{z}$, and $g(x)=mk_\rho x-\frac{kx^2}{2z}$. In this case, we find that $x_c=mk_\rho z/k$ and $\alpha(z)=-\frac{k}{2z}$, whence

$$\begin{aligned}&\int_0^R d\rho'\rho' e^{i\left(mk_\rho\rho'-\frac{k(\rho')^2}{2z}\right)}J_{mn}\left(k\left(\frac{\rho\rho'}{z}\right)\right)\\ &\quad\approx\frac{mk_\rho z}{k}J_{mn}(mk_\rho\rho)e^{i\left(\frac{m^2k_\rho^2 z}{2k}\right)}I(R,z),\end{aligned} \quad (A5)$$

where we have defined $I(R,z)\equiv\int_{x_c}^{R+x_c}e^{i\alpha(z)x^2}dx$, R representing the radius of the holographic mask. The approximate equality (A5) can now be substituted into (A2) to yield

$$\Psi_{tr}^{m}(\rho,\phi,z)\approx -i^{mn+1}mk_\rho e^{i\left(kz-\frac{k_\rho^2}{2z}+mn\phi+\frac{m^2k_\rho^2 z}{2k}\right)}J_{mn}(mk_\rho\rho)I(R,z). \quad (A5')$$

The expression for the transmitted wavefunction provided above includes a contribution from the aperture limiting function $I(R,z)$, which accounts for the finite size of the phase hologram.

This expression is approximate, and no attempts will be made to quantitatively compare it with the numerical solution of the Fresnel integral A1. Nevertheless, this calculation shows a characteristic oscillation, superimposed on a mainly linear growth in the intensity, consistent with our observations Fig. 6.

The term $I(R,z)$ does not affect the transverse profile of the beam, so that

$$\psi_{tr}^{m}(\rho,\phi) = Ne^{i\left(kz-\frac{k_\rho^2}{2z}+mn\phi+\frac{m^2k_\rho^2 z}{2k}\right)}J_{mn}(mk_\rho\rho), \quad (A6)$$

where N is a normalization constant. We note also that (A5) is contingent upon the approximation (A4), and therefore holds true only to the extent that there is a point $\frac{d}{d\rho'}g(\rho')\approx 0$, inside the hologram aperture. Such a point is located at

$$\rho'\approx\frac{mk_\rho z}{k}. \quad (A7)$$

Hence, only a narrow ring of radius $\frac{mk_\rho z}{k}$ in the hologram plane actually makes a stationary contribution to the transmitted beam associated with the mth diffraction order. The condition (A7) can be satisfied for values of z such that $z\le\frac{kR}{mk_\rho}$. For other values of z, the approximation (A4) fails, so that the Bessel character of the transmitted beam is no longer maintained.

Far from the axis, we are in the opposite regime, in which the Bessel function's oscillatory behavior changes the stationary phase analysis. Indeed, for large x,

$$J_{nm}(x)\sim\sqrt{\frac{2}{\pi x}}\cos\left(\frac{nm\pi}{2}+\frac{\pi}{4}-x\right).$$

Using Euler's formula, the integral of interest can be written as a sum of 2 oscillatory components:

$$\Psi_{tr}^{m}(\rho,\phi,z)\propto\int_0^R d\rho'\rho'\left[e^{ig_+(\rho')}+e^{ig_-(\rho')}\right]$$

with phase

$$g_\pm(\rho') = \pm\left[k\left(\frac{\rho\rho'}{z}\right)-\frac{nm\pi}{2}-\frac{\pi}{4}\right]+mk_\rho\rho'-\frac{k(\rho')^2}{2z}.$$

In this case, we can still use stationary phase arguments to find the lateral extent of the region where the intensity of the probe is non-negligible (see appendix of Ref. [30]).

The condition for a stationary contribution is therefore $\frac{d}{d\rho'}g(\rho')=mk_\rho\pm\frac{k\rho}{z}+k\left(\frac{\rho'}{z}\right)=0$. Consequently, the condition for stationary phase becomes $\rho'=\pm\rho+\frac{mk_\rho z}{k}$. However, since the integral (A2) has an upper bound at $\rho'=R$, the stationary phase condition can be met only for $\rho<R-mk_\rho z/k$ or $\rho<-R+\frac{mk_\rho z}{k}=\frac{mk_\rho(z-z_{\max})}{k}$, so that the generated beam will take Bessel form only in this range. In the asymptotic limit discussed, (namely for large ρ) the probe is therefore characterized by a double conical cutoff (as in Fig. A2b) that can be observed by varying the propagation distance z or the diffraction order m. This is therefore relevant to the analysis of Fig. 8 in the main text.

Appendix II. Far-field propagation

We determine the wavefunction of the mth diffraction order in the far-field by evaluating the Fraunhofer diffraction integral

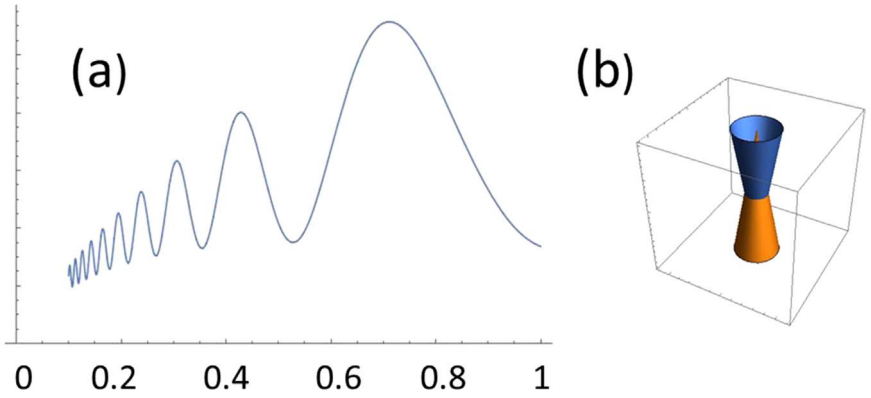

Fig. A2. (a) Evaluation of the function I for typical parameter values used in this paper. A mainly linear increase in intensity can be observed, with superimposed oscillations. (b) Geometrical illustration of the region of stationary phase as a function of z. This schematic demonstrates that the overall size of the probe reaches a minimum in what could be called the central central focus, before rising again.

$$\psi_{tr}^{m}(k_x, k_y) = N \iint dx' dy' A_m(x', y') e^{i(k_x x' + k_y y')}, \tag{A8}$$

where k_x and k_y are the far-field spatial frequency components in the x and y directions. Following a procedure analogous to that outlined in Appendix I, we choose $A_m(x', y') \rightarrow A_m(\rho', \phi') = e^{im(k_\rho \rho' + n\phi')}$, so that

$$\psi_{tr}^{m}(k_x, k_y) = N \int_0^R d\rho' \rho' e^{im\, k_\rho \rho'} \int_0^{2\pi} d\phi' e^{imn\phi'} e^{i\rho'(k_x \cos\phi' + k_y \sin\phi')}.$$

But since $k_x = K\cos\phi$ and $k_y = K\sin\phi$, this becomes

$$\psi_{tr}^{m}(K) = N \int_0^R d\rho' \rho' e^{im\, k_\rho \rho'} \int_0^{2\pi} d\phi' e^{imn\phi'} e^{i\rho' K\cos(\phi - \phi')}.$$

We address the inner integral first, once again resorting to the Jacobi–Anger identity to write

$$\int_0^{2\pi} d\phi' e^{imn\phi'} e^{i\rho' K\cos(\phi-\phi')} = \int_0^{2\pi} d\phi' e^{imn\phi'} \sum_{p=-\infty}^{\infty} i^p J_p(K\rho') e^{ip(\phi-\phi')}$$
$$= \sum_{p=-\infty}^{\infty} i^p J_p(K\rho') e^{ip\phi} \int_0^{2\pi} d\phi' e^{i(mn-p)\phi'},$$

so that, with $\int_0^{2\pi} d\phi' e^{i(mn-p)\phi'} = 2\pi\delta_{mn,p}$,

$$\int_0^{2\pi} d\phi' e^{imn\phi'} e^{i\rho' K\cos(\phi-\phi')} = 2\pi i^{mn} e^{imn\phi} J_{mn}(K\rho').$$

As a result, the transverse wavefunction is then given by

$$\psi_{tr}^{m}(K) = \tilde{N} e^{imn\phi} \int_0^R d\rho' \rho' e^{im\, k_\rho \rho'} J_{mn}(K\rho').$$

We now note that the above expression may be written equivalently in the form

$$\psi_{tr}^{m}(k_\rho) = -i\tilde{N} e^{imn\phi} \frac{\partial}{\partial(mk_\rho)} \int_0^R d\rho' e^{im\, k_\rho \rho'} J_{mn}(K\rho'). \tag{A9}$$

The integral on the right hand side of this new equation can be further decomposed by making use of the Euler identity:

$$\int_0^R d\rho' e^{im\, k_\rho \rho'} J_{mn}(K\rho') = \int_0^R d\rho' \cos(mk_\rho \rho') J_{mn}(K\rho') + i \int_0^R d\rho' \sin(mk_\rho \rho') J_{mn}(K\rho').$$

The above expression takes on qualitatively different solutions, depending on the relative values of k_ρ and K. In particular, for $R \rightarrow \infty$, we have [38]:

$$\int_0^\infty d\rho' \cos(mk_\rho \rho') J_{mn}(K\rho') = \frac{\cos\left(mn \arcsin\left(\frac{mk_\rho}{K}\right)\right)}{\sqrt{K^2 - (mk_\rho)^2}}$$

and

$$\int_0^\infty d\rho' \sin(mk_\rho \rho') J_{mn}(K\rho') = \frac{\sin\left(mn \arcsin\left(\frac{mk_\rho}{K}\right)\right)}{\sqrt{K^2 - (mk_\rho)^2}}$$

when $mk_\rho < K$, and

$$\int_0^\infty d\rho' \cos(mk_\rho \rho') J_{mn}(K\rho') = -\frac{K^{mn} \sin\left(\frac{mn\pi}{2}\right)}{\sqrt{(mk_\rho)^2 - K^2}\left(mk_\rho + \sqrt{(mk_\rho)^2 - K^2}\right)^{mn}}$$

$$\int_0^\infty d\rho' \sin(mk_\rho \rho') J_{mn}(K\rho') = -\frac{K^{mn} \sin\left(\frac{mn\pi}{2}\right)}{\sqrt{(mk_\rho)^2 - K^2}\left(mk_\rho + \sqrt{(mk_\rho)^2 - K^2}\right)^{mn}}$$

when $mk_\rho > K$. Each of the four expressions above feature a singular point when $K = mk_\rho$, as do their derivatives with respect to mk_ρ. Hence, when they are substituted into Eq. (A9), we have that $\psi_t^m(\rho, \phi) \rightarrow \infty$ for $mk_\rho \rightarrow K$, so that, upon normalization, the absolute square of the far-field wavefunction becomes a delta function, centered at $mk_\rho = K$, i.e. $|\psi_t^m(K)|^2 \propto \delta(K - mk_\rho)$. For this reason, the electron beam intensity pattern forms a thin ring in the far-field, as shown in Fig. 1.

Appendix III. Locating phase singularities

Phase singularities occur for non-vanishing integer winding numbers n, as defined by closed contour integrals of the form

$$2\pi n = \oint \vec{\nabla}\varphi \bullet d\vec{s}, \tag{A10}$$

where φ is the phase of the wavefunction, and $d\vec{s}$ is an infinitesimal segment of the closed path around some region of interest. Here, $\vec{\nabla}\varphi$ is obtained from a special type of derivative, defined for example in [39], as

$$\frac{\delta\varphi}{\delta x} = -ie^{-i\varphi} \frac{\partial}{\partial x}\left(e^{-i\varphi}\right).$$

This definition removes the artificial 2π discontinuity that has no meaning for the phase.

In practice, the integral (A10) was evaluated over small rectangular paths, typically of dimension 5x5 pixels. The results were then entered into a map of the beam, and the integral was found to yield zero within floating point precision, and $2\pi n$ about the vortices [40].

References

[1] M. Uchida, A. Tonomura, Generation of electron beams carrying orbital angular momentum, Nature 464 (2010) 737.
[2] J. Verbeeck, H. Tian, P. Schattschneider, Production and application of electron vortex beams, Nature 467 (2010) 301.
[3] B.J. McMorran, et al., Electron vortex beams with high quanta of orbital angular momentum, Science 331 (2011) 192.
[4] K.Y. Bliokh, Y.P. Bliokh, S. Savel'ev, F. Nori, Semiclassical dynamics of electron wave packet states with phase vortices, Phys. Rev. Lett. 99 (2007) 190404.
[5] J. Harris, V. Grillo, E. Mafakheri, G.C. Gazzadi, S. Frabboni, R.W. Boyd, E. Karimi, Structured quantum waves, Nature Physics 11 (2015) 629.
[6] V. Grillo, G.C. Gazzadi, E. Mafakheri, S. Frabboni, E. Karimi, R.W. Boyd, Holographic generation of highly twisted electron beams, Phys. Rev. Lett. 114 (2015) 034801.
[7] J. Verbeeck, G. Guzzinati, L. Clark, R. Juchtmans, R. Van Boxem, H. Tian, A. Béché, A. Lubk, G. Van Tendeloo, Shaping electron beams for the generation of innovative measurements in the (S)TEM, C. R. Phys. 15 (2014) 190.
[8] V. Grillo, E. Karimi, G.C. Gazzadi, S. Frabboni, M.R. Dennis, R.W. Boyd, Generation of nondiffracting electron Bessel beams, Phys. Rev. X 4 (2014) 011013.
[9] V. Grillo, G.C. Gazzadi, E. Karimi, E. Mafakheri, R.W. Boyd, S. Frabboni, Highly efficient electron vortex beams generated by nanofabricated phase holograms, Appl. Phys. Lett. 104 (2014) 043109.
[10] J. Yuan, Bright electron twisters, Nature 509 (2014) 37–38.
[11] T.R. Harvey, J.S. Pierce, A.K. Agrawal, P. Ercius, M. Linck, B.J. McMorran, Efficient diffractive phase optics for electrons, New J. Phys. 16 (2014) 093039.
[12] N. Voloch-Bloch, Y. Lereah, Y. Lilach, A. Gover, A. Arie, Generation of electron Airy beams, Nature 494 (2013) 331.
[13] M. Malac, M. Beleggia, R. Egerton, Y. Zhu, Imaging of radiation-sensitive samples in transmission electron microscopes equipped with Zernike phase plates, Ultramicroscopy 108 (2008) 126–140.
[14] K. Murata, X. Liu, R. Danev, J. Jakana, M.F. Schmid, J. King, K. Nagayama, W. Chiu, Zernike phase contrast cryo-electron microscopy and tomography for structure determination at nanometer and subnanometer resolutions, Structure 18 (2010) 903–912.
[15] Y. Inayoshi, H. Minoda, Y. Arai, K. Nagayama, Direct observation of biological molecules in liquid by environmental phase-plate transmission electron microscopy, Micron 43 (2012) 1091–1098.
[16] F. Zernike, Das Phasenkontrastverfahren bei der mikroskopischen Beobachtung, Z. Tech. Phys. 16 (1935) 454.
[17] Y. Ito, A.L. Bleloch, L.M. Brown, Nanofabrication of solid-state Fresnel lenses for electron optics, Nature 394 (1998) 49–52.
[18] R. Shiloh, Y. Lereah, Y. Lilach, A. Arie, Sculpturing the electron wave function using nanoscale phase masks, Ultramicroscopy 144 (2014) 26.
[19] H. Rose, Nonstandard imaging methods in electron microscopy, Ultramicroscopy 2 (1977) 251–267.
[20] H.L. Xin, V. Intaraprasonk, D.A. Muller, Controlling channeling effects in aberration-corrected STEM tomography, Microsc. Microanal. 14 (2008) 926.
[21] K.Y. Bliokh, P. Schattschneider, J. Verbeeck, F. Nori, Electron vortex beams in a magnetic field: a new twist on landau levels and Aharonov–Bohm states, Phys. Rev. X 2 (2012) 041011.
[22] Z.L. Wang, Dynamical theories of dark-field imaging using diffusely scattered electrons in STEM and TEM, Acta Cryst. A 51 (1995) 569–585.
[23] L. Clark, A. Béché, G. Guzzinati, A. Lubk, M. Mazilu, R. Van Boxem, J. Verbeeck, Exploiting lens aberrations to create electron-vortex beams, Phys. Rev. Lett. 111 (2013) 064801.
[24] E. Völkl, L.F. Allard, D.C. Joy, Introduction to Electron Holography, Springer, Berlin, Heidelberg, 1999.
[25] H. Reimer, H. Kohl, Transmission Electron Microscopy, 3rd ed., Springer-Verlag, New York, 1993 (Springer Series in Optical Sciences).
[26] A. Vasara, J. Turunen, A.T. Friberg, Realization of general non diffracting beams with computer-generated holograms, J. Opt. Soc. Am. 6 (1989) 1748.
[27] M. Ornigotti, A. Aiello, Radially and azimuthally polarized nonparaxial Bessel beams made simple, Opt. Express 21 (2013) 15530–15537.
[28] J. Durnin, J.J. Miceli Jr., J.H. Eberly, Diffraction-free beams, Phys. Rev. Lett. 58 (1987) 1499–1501.
[29] J. Durnin, J.J. Miceli Jr., J.H. Eberly, Comparison of Bessel and Gaussian beams, Opt. Lett. 13 (1988) 79.
[30] V. Grillo, E. Rotunno, STEM_CELL: a software tool for electron microscopy: part I simulations, Ultramicroscopy 125 (2013) 97.
[31] E.J. Kirkland, Advanced Computing in Electron Microscopy, Plenum Press, New York, 1998.
[32] P. Schattschneider, M. Stöger-Pollach, J. Verbeeck, Novel vortex generator and mode converter for electron beams, Phys. Rev. Lett. 109 (2012) 084801.
[33] A. Lubk, G. Guzzinati, F. Börrnert, J. Verbeeck, Transport of intensity phase retrieval of arbitrary wave fields including vortices, Phys. Rev. Lett. 111 (2013) 173902.
[34] A. Lubk, L. Clark, G. Guzzinati, J. Verbeeck, Topological analysis of paraxially scattered electron vortex beams, Phys. Rev. A 87 (2013) 033834.
[35] F. Ricci, W. Löffler, M.P. van Exter, Instability of higher-order optical vortices analyzed with a multi-pinhole Interferometer, Opt. Express 20 (2013) 22962.
[36] M.R. Dennis, K. O'Holleran, M.J. Padgett, Singular optics: optical vortices and polarization singularities, Prog. Opt. 53 (2009) 293–363.
[37] M.R. Dennis, J.B. Götte, Topological aberration of optical vortex beams: determining dielectric interfaces by optical singularity shifts, Phys. Rev. Lett. 109 (2012) 183903.
[38] M. Abramoviz, I.A. Stegun, Handbook of mathematical functions, Applied Mathematics Series, Vol. 55 (10 ed.), New York, USA, 1972.
[39] M.J. Hytch, E. Snoeck, R. Kilaas, Quantitative measurement of displacement and strain fields from HREM micrographs, Ultramicroscopy 74 (1998) 131–146.
[40] K. O'Holleran, F. Flossmann, M.R. Dennis, M.J. Padgett, Methodology for imaging the 3D structure of singularities in scalar and vector optical fields, J. Opt. A: Pure Appl. Opt. 11 (2009) 094020.