Phase retrieval from images in the presence of first-order vortices

L. J. Allen, H. M. L. Faulkner, K. A. Nugent, M. P. Oxley, and D. Paganin
School of Physics, University of Melbourne, Victoria 3010, Australia
(Received 22 September 2000; published 27 February 2001)

We discuss retrieval of the phase of quantum-mechanical and classical wave fields in the presence of first-order vortices. A practical method of phase retrieval is demonstrated which is robust in the presence of noise. Conditions for the uniqueness of the retrieved phase are discussed and we show that determination of the phase in a given plane requires a series of at least three two-dimensional intensity images at different propagation distances. The method is applicable to a wide range of scenarios such as the imaging of imperfect crystals, quantitative determination of the strength of vortex filaments in high-temperature superconductors, and x-ray and electron holography.

DOI: 10.1103/PhysRevE.63.037602 PACS number(s): 42.30.Rx

The study of wave-field singularities, both quantized and classical, is a rich and diverse topic for fundamental and applied research. Noninterferometric determination of the phase of quantum-mechanical and classical wave fields (i.e., phase retrieval) is a topic of current interest in a number of areas where either phase imaging or structure retrieval is an issue. For example, phase measurement is topical for optical [1], x-ray [2,3], neutron [4], electron [5–8], and atom [9] wave fields. In all of these areas it is possible that vortices and other phase singularities may arise [10–12]. This means that the phase is not necessarily continuous and single valued and a method of phase retrieval is required that can accommodate this. In this Brief Report we demonstrate a practical approach to phase retrieval in the presence of first-order vortices which is robust in the presence of noise. Conditions for the uniqueness of the retrieved phase are discussed and we show that determination of the phase in a given plane requires a series of at least three two-dimensional intensity images at different propagation distances.

Consideration of vortices in quantum mechanics appears in many places, for example, Dirac's pioneering article on quantized singularities in the electromagnetic field [10], quantized vortices around wave-function nodes [13], the Aharonov-Bohm wave function [14], angular-momentum eigenstates of the hydrogen atom [15], superconductors [5], superfluids [16], and Bose-Einstein condensates [17]. The study of vortices in the visible-light optics community was largely spurred on by the work of Nye and Berry [18] in the mid 1970s. In this context, vortices are often associated with phenomena like diffraction-free beams [19], vortex solitons [20], or holographically produced screw dislocations [21]. We emphasize that the presence of wave-field vortices is ubiquitous rather than exotic. Indeed, the physics community is becoming increasingly aware of the fact that vortex-free propagating wave fields will almost always develop vortices after evolving through a certain critical length of time or space [12,18,22,23].

We restrict our discussion to monoenergetic propagating wave functions $\psi(\mathbf{r})$ that satisfy the time-independent free-space Schrödinger equation

$$(\nabla^2 + k^2)\,\psi(\mathbf{r}) = 0, \tag{1}$$

where $k = 2\pi/\lambda$ and λ is the free-space de Broglie wavelength. As is well known, an equation of this form is also satisfied by monochromatic scalar electromagnetic waves in free space, allowing us to treat simultaneously the cases of monoenergetic electrons and monochromatic electromagnetic waves. We write the wave function in terms of its modulus and phase in the form $\psi(\mathbf{r}) = |\psi(\mathbf{r})| \exp[i\phi(\mathbf{r})]$. From the continuity and single valuedness of the wave function, it follows that the squared modulus (intensity) of the wave function will also be continuous and single valued. However, the fact that the phase is undefined wherever $|\psi(\mathbf{r})| = 0$ implies that $\phi(\mathbf{r})$ may be a discontinuous and multivalued function of the coordinates $\mathbf{r}$. Since the phase is only defined modulo 2π, its integral around a closed loop need not vanish and will in general equal an integer multiple m of 2π (m is called the topological charge):

$$\oint_\Gamma \nabla\phi(\mathbf{r}) \cdot d\hat{\mathbf{n}} = m(2\pi). \tag{2}$$

Here, $\hat{\mathbf{n}}$ is the unit vector tangential to the closed path Γ around which the loop integral is taken; Γ is any closed loop in the space of coordinates $\mathbf{r}$ over which $|\psi(\mathbf{r})| > 0$. A nonzero value for the circulation integral above heralds the presence of a phase vortex.

We will generate wave fronts with vortices in the phase from an input model wave function without vortices that has the intensity and phase shown in Fig. 1. The arbitrary nature of these input data serves to highlight the fact, mentioned earlier, that almost any choice of images for the probability density and phase of the wave function will lead to phase vortices after the wave field has been allowed to propagate a sufficient distance through free space.

Our example can be considered to apply to two cases. First assume that the intensity and phase maps in Fig. 1 (each 160×160 pixels) have dimensions $75\,\text{Å} \times 75\,\text{Å}$ and are formed by a beam of electrons of energy 100 keV (wavelength 0.037 Å). Alternatively we can consider the image in Fig. 1 to have dimensions $1.28\,\text{mm} \times 1.28\,\text{mm}$ and the incident radiation to be laser light of wavelength 6328 Å (HeNe laser). Then the image and phase maps in Fig. 2 correspond to the defocus values indicated in the square brackets. Spherical aberration could be included in constructing the

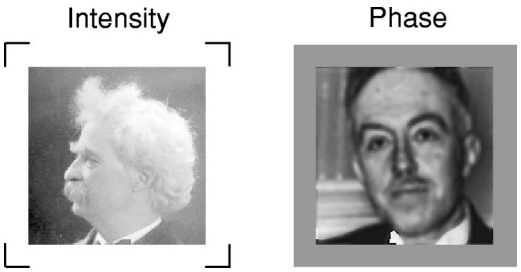

FIG. 1. Input intensity (Mark Twain) and phase map (Louis-Victor de Broglie) which form our initial wave function. The intensity varies between 0 and 1 and the phase varies between $-\pi$ and π. Pictures originally containing 128×128 pixels have been padded out to 160×160 pixels—by 1s in the intensity and zeros in the phase.

focal series [23] but would be more pertinent for electron optics. Propagation of the wave function in free space yields the series of phase maps shown in Fig. 2. Vortices are present in all these phase maps. (The phase maps are wrapped—unwrapping in the presence of vortices has been discussed by Carter [24]). We indicate one such vortex by an arrow in each phase map. The vortices we have indicated are each connected to a counter-rotating partner by branch lines.

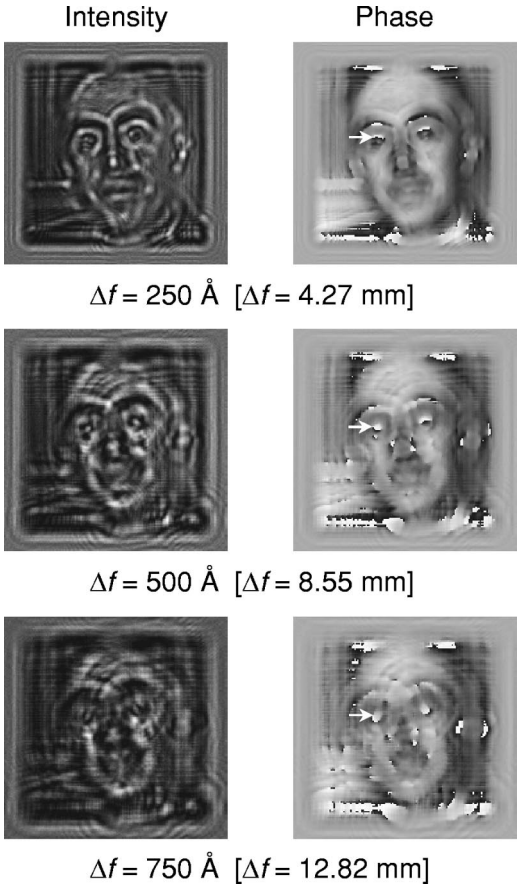

FIG. 2. Series of defocused images and phase maps generated from the wave function shown in Fig. 1. The defocus values in Å correspond to the electron scattering case discussed in the text and those in square brackets in mm to the laser optics case.

The images obtained after propagation were first modified by the addition of noise before display in Fig. 2. We assigned a number of counts to each pixel in the images assuming that an intensity of unity contained noise at the 5% level. The statistical errors at each pixel were assigned using a random deviate drawn from a Poisson distribution with mean corresponding to the noise-free number of counts for a given pixel [25].

Consider the image for $\Delta f=500$ Å in Fig. 2. Close examination, at higher magnification, shows that each zero in the intensity is a point and in the vicinity of a zero the intensity behaves as $I(x_i,y_i)=Ax_i^2+Bx_iy_i+Cy_i^2$ (coordinates x_i and y_i refer to a reference frame with the origin centered on the ith zero). This implies that in the vicinity of the zero the real and imaginary components of the wave function behave as $\psi_R=ax_i+by_i$ and $\psi_I=cx_i+dy_i$, respectively [26]. In other words we have simple first-order zeros in the wave function. This is likely to be the case for a large number of realistic physical situations. The phase in the vicinity of the zero is given by

$$\phi(x_i,y_i)=\arctan\left(\frac{c+d\tan\theta_i}{a+b\tan\theta_i}\right), \tag{3}$$

where θ_i is the polar coordinate in the x_iy_i plane. Since

$$\frac{d\phi}{d\theta_i}=\frac{(ad-bc)(1+\tan^2\theta_i)}{(a+b\tan\theta_i)^2+(c+d\tan\theta_i)^2}, \tag{4}$$

the phase is monotonically increasing or decreasing, depending on the sign of $(ad-bc)$, and exhibits a vortex behavior [26]. The change in the phase ϕ as the coordinate θ_i varies from 0 to 2π about the ith vortex is

$$\Delta\phi_i=\int_0^{2\pi}\frac{d\phi}{d\theta_i}d\theta_i=\pm2\pi. \tag{5}$$

Therefore a simple first-order zero in intensity is associated with a vortex in the phase with topological charge $m=\pm1$ and any correct solution to the phase problem must show this behavior. Therefore we have a situation where we know the positions of all vortices and that they have a topological charge $m_i=\pm1$.

In practice we use the following approach to retrieve the phase. Having made an initial guess (of a constant) for the phase at $\Delta f=500$ Å we propagate the wave in free space cyclically between the images in the series in Fig. 2, correcting the modulus to the known value after each propagation step until convergence is obtained. We discuss our approach to the phase retrieval in terms of the electron optics case but it is equally applicable to the laser optics case just mentioned (or any other equivalent situation for that matter). A sum squared error [27] of 6.48×10^{-4} was obtained after 939 iterations. The retrieved wave function at $\Delta f=500$ Å is shown in Fig. 3. Also shown is the wave function that is then obtained by back propagation to zero defocus, which can be compared with the wave function in Fig. 1. The effect of noise in the focal series in Fig. 2 is mainly evident in the intensity rather than the phase. For fast convergence of the

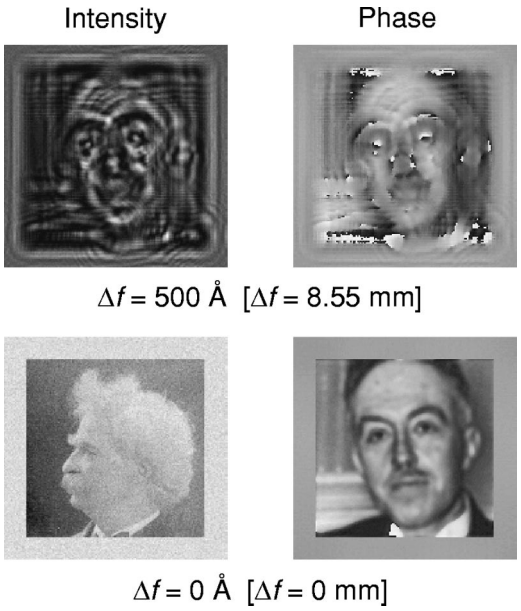

FIG. 3. The image and the recovered phase map for $\Delta f = 500$ Å for electron optics [$\Delta f = 8.55$ mm in the laser optics case] and the corresponding image and phase map for back propagation to $\Delta f = 0$ Å [mm].

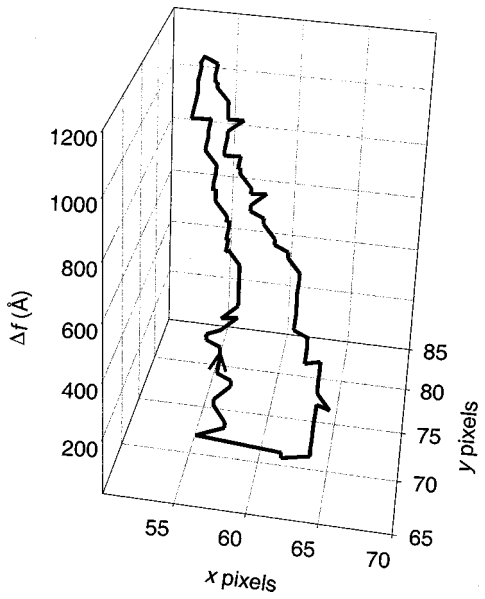

FIG. 4. Vortex trajectory on which the vortices indicated by the arrows in Fig. 2 lie. The counter-rotating partners evident in the phase maps also lie on the same trajectory. Moving along the trajectory in the direction of the arrow, the vortex is rotating counterclockwise.

algorithm the step between the images (250 Å in this case) must be sufficiently large for there to be significant differences between the images. At least three images are needed to obtain the correct phase map. Using five images 250 Å apart, convergence is obtained after only 230 iterations. Tested on other focal series of three or more images the method was always successful. A practical consideration for applications is that the images would need to be carefully aligned.

Having obtained the phase map at $\Delta f = 500$ Å from the intensity data in Fig. 2, we can use this phase as an initial estimate for a phase retrieval using the images at, say, $\Delta f = 500$ and 750 Å and a further image plane close to that at $\Delta f = 500$ Å, say, $\Delta f = 475$ Å. Assuming that the phase map at $\Delta f = 500$ Å is then essentially unchanged we can now consider the relationship of the retrieved phase to the so-called transport of intensity equation (TIE) [9]:

$$\nabla_\perp \cdot (I \nabla_\perp \phi_s) = -\sum_i \frac{m_i}{r_i} \frac{\partial I}{\partial \theta_i} - k \frac{\partial I}{\partial z}. \qquad (6)$$

In Eq. (6) the phase has been decomposed into scalar and vector components [9] via $\nabla_\perp \phi = \nabla_\perp \phi_s + (\nabla \times \boldsymbol{\phi}_v)_\perp$ (the symbol $\perp$ denotes operation in the image or xy plane) and the term containing the vector phase is determined by the first term on the right-hand side (RHS) of Eq. (6) (this term describes a differential rotation of probability density about the vortex core). The term on the LHS describes a differential translation of probability density transverse to the propagation direction z. Furthermore, r_i is the radial coordinate in the reference frame of the ith vortex and $\partial I / \partial z$ is the variation in intensity along the z direction, which can be estimated from the two closely spaced images. Our solution for the phase should satisfy Eq. (6) (within the limits of the paraxial approximation implicit within the TIE).

We ignore terms in Eq. (6) for which $\partial I / \partial \theta_i = 0$ [9]. Ambiguities corresponding to such ''unobservable'' phases can be identified using symmetry arguments. If the RHS is known then Eq. (6) may be uniquely solved for the phase. However, the RHS is not uniquely specified by the intensity measurements because the signs of the topological charges m_i are not determined. The only other possible solutions to the TIE correspond to different choices for the signs of the topological charges. There are $2^n - 1$ further sets of signs if there are n vortices. In general only the solution we have obtained by iteration will propagate correctly to a third plane and those obtained from the TIE corresponding to other sets of signs for the topological charges will not, and this can be checked.

The vortices lie on trajectories that may be tracked, as shown by the example in Fig. 4 (the vortices shown by the arrows in Fig. 2 are on this trajectory). Having chosen the sign of any vortex in the image plane on a given trajectory, the signs of all other vortices evident in the image plane on the same trajectory are determined. This can be used to reduce the number of possible sets of signs for the topological charges m_i (there is only one free parameter for charges on a closed vortex trajectory).

We now briefly discuss some anticipated applications of the method of phase retrieval we have demonstrated. Immediate applications may be made to the quantitative imaging of crystal defects using high-resolution transmission electron microscopy, where phase vortices are ubiquitous [23]. Importantly, the method discussed here is easily modified to correct explicitly for any aberrations that are present in the imaging system, provided that they are accurately known

[23]. The method may also be readily applied to the imaging of high-temperature superconductors containing a network of vortex filaments [5]. In the context of transmission electron microscopy on such samples, the correct phasing of the exit-surface wave function is easily related to the strength of the vortex filaments in the superconductor and thereby to their desirable technological properties. We stress that the evident ubiquity of the vortex implies the very limited domain of validity possessed in this context by phase-retrieval algorithms that assume the phase to be vortex free. Lastly, the methods presented can be applied to electron and x-ray holographic microscopy, obviating the need for any form of imaging optics.

In conclusion, there exist many scenarios where the presence of vortexlike phase singularities demands a phase-retrieval algorithm that is able to cope with their existence. The ideas discussed here comprise a simple, robust, and rapid method for the phase retrieval of both quantum-mechanical and classical wave fields in the presence of first-order vortices. Uniqueness of the retrieved phase can be checked *a posteriori*.

L.J.A., K.A.N., and D.P. acknowledge financial support from the Australian Research Council. We thank Professor Isaac Freund for helpful correspondence.

[1] A. Barty, K. A. Nugent, D. Paganin, and A. Roberts, Opt. Lett. **23**, 817 (1998).
[2] K. A. Nugent, T. E. Gureyev, D. Cookson, D. Paganin, and Z. Barnea, Phys. Rev. Lett. **77**, 2961 (1996).
[3] T. E. Gureyev, C. Raven, A. Snigirev, I. Snigireva, and S. W. Wilkins, J. Phys. D: Appl. Phys. **32**, 563 (1999).
[4] B. E. Allman, P. J. McMahon, K. A. Nugent, D. Paganin, D. L. Jacobson, M. Arif, and S. A. Werner, Nature (London) **408**, 158 (2000). .
[5] A. Tonomura, Physica B **280**, 227 (2000).
[6] S. Bajt, A. Barty, K. A. Nugent, M. McCartney, M. Wall, and D. Paganin, Ultramicroscopy **83**, 67 (2000).
[7] J. C. H. Spence, Acta Crystallogr., Sect. A: Found. Crystallogr. **54**, 7 (1998).
[8] L. J. Allen, H. M. L. Faulkner, and H. Leeb, Acta Crystallogr., Sect. A: Found. Crystallogr. **56**, 119 (2000).
[9] K. A. Nugent and D. Paganin, Phys. Rev. A **61**, 063614 (2000).
[10] P. A. M. Dirac, Proc. R. Soc. London, Ser. A **133**, 60 (1931).
[11] T. E. Gureyev, A. Roberts, and K. A. Nugent, J. Opt. Soc. Am. A **12**, 1942 (1995).
[12] D. L. Fried, J. Opt. Soc. Am. A **15**, 2759 (1998).
[13] J. O. Hirschfelder, C. J. Goebel, and L. W. Bruch, J. Chem. Phys. **61**, 5456 (1974).
[14] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).
[15] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961), Vol. 1.
[16] R. Blaauwgeers, V. B. Eltsov, M. Kruslus, J. J. Ruohio, R. Schanen, and G. E. Volovik, Nature (London) **404**, 471 (2000).
[17] M. R. Matthews, B. P. Anderson, P. C. Haljan, D. S. Hall, C. E. Wieman, and E. A. Cornell, Phys. Rev. Lett. **83**, 2498 (1999).
[18] J. F. Nye and M. V. Berry, Proc. R. Soc. London, Ser. A **336**, 165 (1974).
[19] A. Vasara, J. Turunen, and A. Friberg, J. Opt. Soc. Am. A **6**, 1748 (1989).
[20] P. Di Trapani, W. Chinaglia, S. Minardi, A. Piskarskas, and G. Valiulis, Phys. Rev. Lett. **84**, 3843 (2000).
[21] M. S. Soskin, V. N. Gorshkov, M. V. Vasnetsov, J. T. Malos, and N. R. Heckenberg, Phys. Rev. A **56**, 4064 (1997).
[22] M. V. Berry, *Physique des Défauts/Physics of Defects*, Proceedings of the Les Houches Summer School of Theoretical Physics, Session XXXV, 1980, edited by R. Balian *et al.* (North-Holland, Amsterdam, 1981) 453–543.
[23] L. J. Allen, H. M. L. Faulkner, M. P. Oxley, and D. Paganin Ultramicroscopy (to be published).
[24] W. H. Carter, Opt. Commun. **94**, 1 (1992).
[25] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
[26] I. Freund, Opt. Commun. **137**, 118 (1997).
[27] W. O. Saxton, *Computer Techniques for Image Processing in Electron Microscopy* (Academic Press, New York, 1978).