Movable Aperture Lensless Transmission Microscopy: A Novel Phase Retrieval Algorithm

H. M. L. Faulkner and J. M. Rodenburg

Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom

(Received 12 November 2003; published 9 July 2004)

We propose an iterative phase retrieval method that uses a series of diffraction patterns, measured only in intensity, to solve for both amplitude and phase of the image wave function over a wide field of view and at wavelength-limited resolution. The new technique requires an aperture that is scanned to two or more positions over the object wave function. A simple implementation of the method is modeled and demonstrated, showing how the algorithm uses overlapping data in real space to resolve ambiguities in the solution. The technique opens up the possibility of practical transmission lensless microscopy at subatomic resolution using electrons, x rays, or nuclear particles.

DOI: 10.1103/PhysRevLett.93.023903 PACS numbers: 42.30.Rx, 07.78.+s, 42.30.Kq, 42.30.Va

Introduction.—We propose a new principle of transmission microscopy, suitable for all forms of radiation, which does not rely on the use of a lens, a holographic reference wave, or any other form of far-field interferometry. The technique could provide wavelength-limited resolution of transparent objects over a wide field of view, with potential applications largely with those radiations where the manufacture of high quality lenses with large numerical apertures is difficult, in particular, electron microscopy and x-ray microscopy. Lens imaging with radiation of very short wavelength encounters two grave difficulties. First, the lens must have very low aberrations; otherwise, phase errors are introduced to the diffracted wave before it is reinterfered into an image. This is extremely difficult to achieve in the case of both electron lenses and x-ray zone plates. Second, the experimental setup must be stable enough so that high-angle beams, lying at the extremes of the diffraction plane, still interfere coherently at the image plane. This constraint, predicated by the stability of the microscope and the chromatic spread in the illuminating beam, is particularly debilitating in the case of electrons. Therefore, even with sophisticated aberration correction, the usable angular range of a typical electron lens is roughly 1° or 2°, as shown schematically in Fig. 1(a). The effect of this limit in momentum space (i.e., the range of lateral scatter of the incident radiation that can be processed by the lens) is that spatial resolution is severely compromised.

An alternative strategy is to discard the lens and simply record the diffraction pattern intensity directly via a photographic film or charge-coupled device (CCD) detector, as illustrated schematically in Fig. 1(b). The great experimental advantage of diffraction is that the interference condition is determined only by scattering within the specimen itself: we do not require the reinterference of beams that have traveled large distances through optical apparatus. This means that the effective spatial resolution of x-ray or electron crystallography can be very high (sub-Å), even though the stability required in the experimental setup is modest. The fact that we can measure intensity only in the far field may at first seem to pose an intractable problem. However, if the object is physically small and its size is accurately known, unique solution of the phase problem is usually possible, provided the diffraction pattern is sampled on a sufficiently fine angular resolution [1,2]. A way to find this solution, which has gained considerable interest recently, is the iterative method first suggested by Gerchberg and Saxton [3], and later developed by Fienup [4]. These have recently been applied to the geometry shown in Fig. 1(b) for both electrons and x rays [5–7].

There are, however, certain difficulties with this arrangement. Experimentally, it is exceedingly difficult to isolate a sufficiently small object in order to undertake the diffraction experiment, since in most microscopic situations it is more useful to collect a wide field of view and then magnify only a particular area of interest. At the theoretical level, there can also be problems with convergence if the object is complex (it introduces both amplitude and phase changes to the illuminating beam), which is generally the case, especially in electron imaging [7]. The presence of noise can also undermine the convergence of such algorithms.

We show here that we can very usefully extend the simplicity and elegance of iterative methods to the experimental arrangement shown in Fig. 1(c). A movable aperture is positioned downstream of an extended object. We collect two or more diffraction patterns, as the aperture is shifted to different positions. As we move the aperture, a different region of object exit wave function is transmitted, so the variables we are solving for progressively change. The algorithm we describe below surmounts this difficulty, with the result that we can solve for a large field of view of an extended object. The method also solves some of the convergence problems that can arise with any one single diffraction pattern, especially when the object is complex. We note that the ability to move an aperture in this way allows an alternative solution to the phase problem [8–12], and so it is not surprising that collecting multiple sets of data improves the

 0031-9007/04/93(2)/023903(4)$22.50

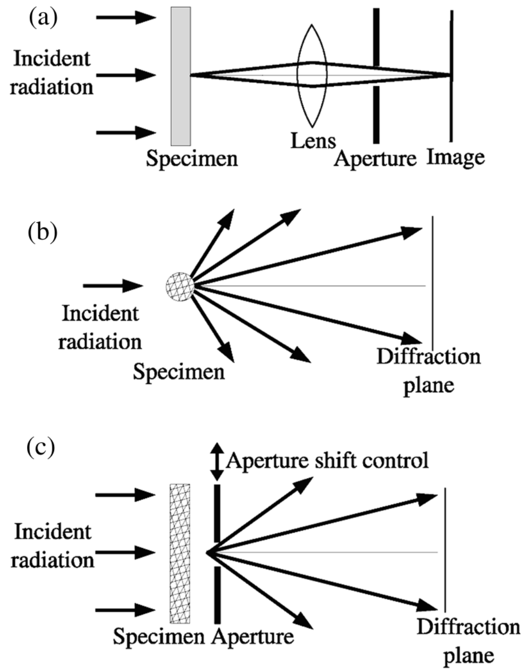

FIG. 1. Comparison of experimental setups. (a) The use of a conventional lens, e.g., in electron or x-ray microscopy, seriously limits the angular range of scattered waves that can be processed simultaneously. A limiting aperture exists in the lens back focal plane. (b) Diffraction does not limit the angular range, but the phase of the scattered wave field is lost, although if the object is small, methods exist to recover the phase. (c) A moving aperture arrangement allows a large field of view to be measured at wavelength resolution.

iterative method. The great advantage of our method is that the aperture can be moved relatively large distances (half the aperture width or more) before each diffraction pattern is recorded. This is in contrast to other diffraction pattern processing methods [8,13], in which the sampling periodicity of the aperture (or, in the case of scanning transmission electron microscopy, the electron probe) position must be at the same scale as the resolution of the final reconstruction. This suggests that the new method could scan very large fields of view and obtain very high resolution images. In this way, all the complications of conventional lens imaging can be disposed of, while maintaining the convenience and speed of parallel imaging.

Conventional iterative algorithms.—Most iterative phase retrieval algorithms have a structure similar to that illustrated in Fig. 2. They require images in two or more different planes, related by some kind of transformation. Starting with a guessed version of the wave function in one plane, an iterative phase retrieval algorithm then computes the result of transforming that guessed wave to the second plane. At the second plane a constraint is imposed on the resulting wave function, before transforming it back to the first plane, imposing the first constraint, and repeating this process until convergence is achieved. The constraints are known information about the wave function in that plane. In practice the most commonly used constraints involve the wave intensity, since that is the observable quantity in a complex wave.

The original iterative phase retrieval algorithm was invented by Gerchberg and Saxton in 1972 [3]. In the Gerchberg-Saxton method, intensity in the image and diffraction planes is used as input information for the algorithm, and the transform relating these is the Fourier transform. A variation on this idea is the Fienup method [4], which requires only the support (i.e., the area where the function is nonzero) in the real plane to be known, rather than the entire intensity. This means an entire real space image need not be measured. Methods of discovering the support, based on the autocorrelation of the wave function have been explored by Weierstall *et al.* [14], and are very successful in some situations. However, this algorithm is limited by the fact that in order to satisfy the Nyquist criterion the support must cover less than half of the image area and must often be much smaller to allow practical recovery of the wave function. The Fienup method is also subject to problems with nonuniqueness in the phase retrieval, for example, retrieving the complex conjugate of the actual wave function.

It is not essential to use the Fourier transform to relate the data from the two planes in an iterative algorithm. One approach that uses a different transform has been known of for some time [15,16], and more recently improved by Allen *et al.* [17]. This is the through focal series algorithm approach, which uses the free space propagator to relate data at different defocii. Our algorithm is similar mathematically to these methods because it increases convergence and reduces ambiguity by using multiple data sets. However, our method does not rely on the use of a lens to focus on the Fresnel image, and so it is not limited by the usual constraint of partial coherence and instability in the beam. It therefore promises a route to wavelength-limited resolution.

A new algorithm.—In devising a new algorithm for phase retrieval we wish to retain the useful features of the known iterative algorithms such as convergence, noise tolerance, and general stability. We also wish to avoid the experimental difficulties posed by currently available iterative techniques. This is achieved by creating an algorithm that uses measured information in the diffraction plane only, negating the need for focused images or a known, fixed support function. Other desirable features

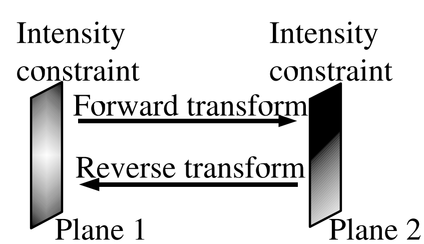

FIG. 2. Schematic of iterative algorithm structure.

to include are the ability to examine any part of the specimen that is of interest and the ability to recover large areas of the specimen.

We assume that the intensity of the diffraction pattern is measured at one or more known aperture positions and that the aperture shape is also known. The algorithm, illustrated in Fig. 3 for two aperture positions (it may be trivially extended to many more positions), works as follows: (1) Start with a guess at the object function. (2) Multiply the current guess at the object function by the aperture at the current position, producing the exit wave function for that aperture position. (3) Fourier transform to obtain the guessed diffraction pattern for that aperture position. (4) Correct the intensities of the guessed diffraction pattern to the known values. (5) Inverse Fourier transform back to real space to obtain a new and improved guess at the exit wave function. (6) Update the guessed wave function in the area covered by the aperture. (7) Move the aperture to a new position, which in part overlaps the previous aperture position. (8) Repeat (2)–(7) until the sum squared error (SSE) as measured in the diffraction plane is sufficiently small. When the SSE is small, it means that the object function has been correctly found in the region covered by the different apertures. The partial overlapping of each aperture with one or more of the other apertures is a crucial aspect of the algorithm's success. Using multiple overlapping apertures allows examination of a wide field of view. It also requires the algorithm to find a solution which is consistent with all parts of the object function that are allowed through one or more of the apertures. This requirement for consistency between different measurements eliminates possible uniqueness problems, by breaking the symmetry of the situation. This is comparable with the way in which the Fienup algorithm overcomes uniqueness problems by using a nonsymmetric support.

If only one aperture position is used, the algorithm reduces to the Fienup algorithm. If more than one aperture position is used, but they do not overlap, the result is multiple concurrent Fienup algorithms. In both cases the resulting solution gives less information than if the apertures overlap.

Using this algorithm, assuming an electron wavelength of 0.037 Å (100 keV), a CCD detector array of 2000×2000 pixels, each 20 μm in width, and an aperture size of 100 nm, the required camera length would be roughly 0.540 m, and the final resolution would be about 0.5 Å. A larger detector array could be used to increase the final resolution.

Note that this algorithm does not require any sort of focused illumination. All that is required is a source of coherent radiation, an aperture, and a method of moving the aperture a known distance across the sample. This last may be achieved with the use of suitable piezoelectric devices. As a result, the equipment requirements of this algorithm are significantly less than those of the conventional iterative algorithms. This method has the potential to revolutionize microscopy and allow lensless imaging over a large field of view.

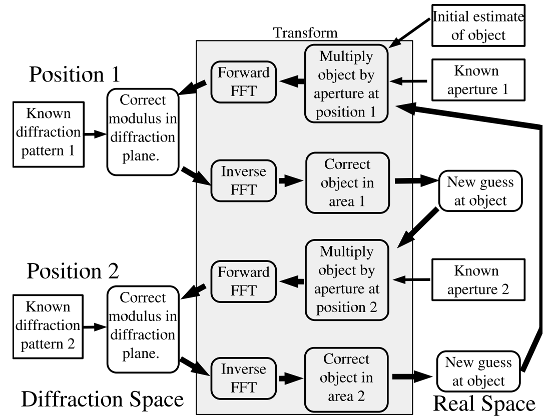

FIG. 3. New phase retrieval algorithm, shown for two aperture positions.

Example simulation.—Figures 4(a)–4(d) show the test data used for a demonstration of the new algorithm. The test object 4(a) and 4(b) consists of photos of a schnauzer (intensity, varying from 0 to 1), and of a cormorant (phase, varying from 0 to π). This is a complex data set, representing a difficult problem in phase retrieval. Random noise of up to $\pm 10\%$ has been added to the diffraction pattern values, in order to more accurately simulate the experimental situation. The simple top-hat function used to represent the aperture is shown in 4(c). A typical diffraction pattern for this data set is shown in 4(d), scaled for visibility so black $= 0$ and white $= 1$. There would usually be several of these, i.e., one for each aperture position. These diffraction patterns and the known aperture shape are used as input data for the algorithm.

Figures 4(e)–4(h) show the results of this simulation for different possible aperture positions. All results are scaled to the same scale as the original object intensity and phase, for ease of comparison. In 4(e) and 4(f) one aperture position was used, reducing our algorithm to the simple Fienup algorithm. The retrieved result has an error of 0.0500 after 200 iterations, and covers a small area of the object wave function. A wider area is covered in 4(g) and 4(h), where four different aperture positions have been used. These positions do not overlap, so the algorithm is effectively running four separate Fienup algorithms, with no overlapping information to connect the resulting wave functions together. It is of particular interest to note the secondary complex conjugate image of a cormorant recovered in the lower left corner of 4(h), and the complex conjugated cormorant head in the top right corner. These are examples of ambiguity in the data set allowing an incorrect solution to be found and stagnation at a wrong solution to occur. The final error in diffraction space is 0.0514.

Figures 4(i) and 4(j) show an example retrieval for which the aperture positions overlap. While the recovered

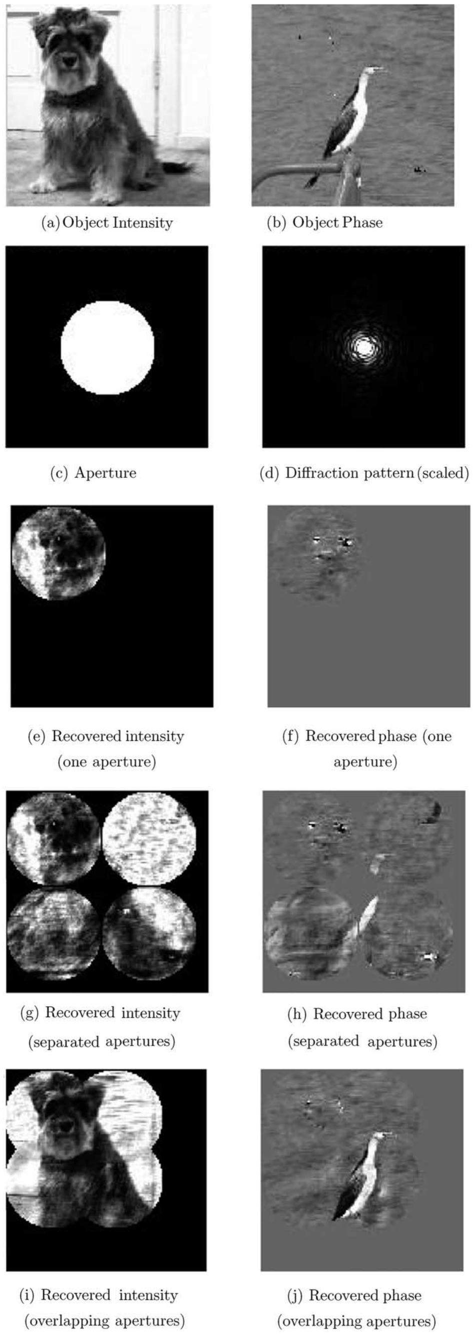

FIG. 4. Simulation of new phase retrieval method for varying arrangements of apertures.

area remains large, the ambiguity in the phase has been removed, and the error is the much improved value of 0.0056 after 200 iterations. It is clear that overlapping aperture positions greatly improve the success of the algorithm.

The algorithm has been tested with many different simulated objects. It works in every case where the aperture positions are chosen so that different diffraction information is present for each position. As long as this requirement is satisfied, the algorithm works for any complex object.

Conclusions.—We have demonstrated a new algorithm for retrieval of an object function, based on knowledge of a simple aperture, and diffraction pattern measurements at two or more different aperture positions. The proposed algorithm combines useful properties of iterative phase retrieval techniques with a new experimental arrangement that allows lensless microscopy, thus greatly simplifying the experimental requirements. The method is shown in simulation to retrieve the object structure with good levels of accuracy, comparing favorably with other iterative techniques. The use of overlapping aperture positions to eliminate ambiguities in the retrieved phase and the ability to retrieve a large area of the object wave function make this algorithm a powerful technique.

There are many potential future directions of this research. These include refining the algorithm by optimizing the number and location of the aperture positions used, using averaging techniques to implement a more parallel algorithm, and using a feedback loop to improve the effectiveness of the error reduction. The relatively simple apparatus required for these techniques has the potential to revolutionize electron microscopy and permit ‘‘lensless’’ imaging of structure at the atomic level.

The authors are grateful for financial support from EPSRC (Grant No. GR/R75076/02).

[1] D. M. Wrinch, Philos. Mag. **27**, 98 (1939).
[2] R. H. T. Bates, Optik **61**, 247 (1982).
[3] R. W. Gerchberg and W. O. Saxton, Optik **35**, 237 (1972).
[4] J. R. Fienup, Appl. Opt. **21**, 2758 (1982).
[5] J. M. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. A. Nagahara, Science **300**, 1419 (2003).
[6] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, Nature (London) **400**, 342 (1999).
[7] J. C. H Spence, U. Weierstall, and M. Howells, Phil. Trans. R. Soc. London **360**, 875 (2002).
[8] P. D. Nellist, B. C. McCallum, and J. M. Rodenburg, Nature (London) **374**, 630 (1995).
[9] P. D. Nellist and J. M. Rodenburg, Acta Crystallogr. Sect. A **54**, 49 (1998).
[10] W. Hoppe, Acta Crystallogr. A **25**, 495 (1969); **25**, 508 (1969).
[11] W. Hoppe and G. Strube, Acta Crystallogr. A **25**, 502 (1969).
[12] R. Hegerl and W. Hoppe, Ber. Bunsen-Ges. Phys. Chem. **74**, 1148 (1970).
[13] J. M. Rodenburg, B. C. McCallum, and P. D. Nellist, Ultramicroscopy **48**, 304 (1993).
[14] U. Weierstall, Q. Chen, J. C. H. Spence, M. R. Howells, M. Isaacson, and R. R. Panepucci, Ultramicroscopy **90**, 171 (2002).
[15] D. L. Misell, J. Phys. D **6**, L6 (1973).
[16] P. van Toorn, A. M. J. Huiser, and H. A. Ferwerda, Optik **51**, 309 (1978).
[17] L. J. Allen, H. M. L. Faulkner, K. A. Nugent, M. P. Oxley, and D. Paganin, Phys. Rev. E **63**, 037602 (2001).