

Scanning Microscopy

Volume 1992
Number 6 *Signal and Image Processing in
Microscopy and Microanalysis*

Article 21

1992

A Robust Solution to the Super-Resolution Phase Problem in Scanning Transmission Electron Microscopy

J. M. Rodenburg
Cambridge

B. C. McCallum
Cambridge

Follow this and additional works at: <https://digitalcommons.usu.edu/microscopy>



Part of the [Biology Commons](#)

Recommended Citation

Rodenburg, J. M. and McCallum, B. C. (1992) "A Robust Solution to the Super-Resolution Phase Problem in Scanning Transmission Electron Microscopy," *Scanning Microscopy*. Vol. 1992 : No. 6 , Article 21. Available at: <https://digitalcommons.usu.edu/microscopy/vol1992/iss6/21>

This Article is brought to you for free and open access by the Western Dairy Center at DigitalCommons@USU. It has been accepted for inclusion in Scanning Microscopy by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



A ROBUST SOLUTION TO THE SUPER-RESOLUTION PHASE PROBLEM IN SCANNING TRANSMISSION ELECTRON MICROSCOPY

J. M. Rodenburg* and B. C. McCallum

Cavendish Laboratory, Madingley Road,
Cambridge CB3 0HE, U.K.

Abstract

From a set of images, each of poor resolution, recorded in a transmission microscope under many different incident angles of coherent illumination, it is possible to obtain wavelength-limited resolution even if there is a narrow aperture lying in the back-focal plane of the imaging lens. This is achieved by a deconvolution algorithm which retrieves the phase of the Fourier transform of the specimen. The method accounts for complex components in the transfer function of the lens, is not very sensitive to defocus and is remarkably resilient to noise. It may have important applications in overcoming the resolution limit in the scanning transmission electron microscope (STEM), where such data are readily available.

Key Words: Super-resolution, phase-retrieval, scanning transmission electron microscopy, microdiffraction, deconvolution.

*Address for correspondence:
John M. Rodenburg,
Cavendish Laboratory,
Madingley Road,
Cambridge, CB3 0HE, U.K.

Telephone number: 44-233-337332
FAX number: 63263

Introduction

We are concerned here with overcoming the conventional coherent Rayleigh resolution limit in any circumstance where it is possible only to manufacture a poor quality lens of small useful numerical aperture. The limitations on the lens may either be physical (as in the case of an objective aperture inserted in the back-focal plane of an image-forming lens) or may be associated with instability or incoherence of one form or another which restricts the useful region of the back-focal plane - for example the instrument function which attenuates the contrast transfer function in the transmission electron microscope (TEM) (Frank, 1973). The Abbé theory of light suggests that under these circumstances, resolution is unavoidably limited because a low-pass filtering process occurs in the reciprocal space of the image. However, this view neglects a large number of other experiments we could perform on the same apparatus. We also have the opportunity to record bright and dark-field images from all possible angles of incidence of the illuminating radiation. What we wish to show here is that this four-dimensional set of data, even when it can only be recorded in intensity, provides unlimited access to super-resolution information in a way which is relatively easy to implement and which is surprisingly robust. In practice, the necessary data set is most easily recorded in the scanning transmission electron microscope (STEM), although the technique could also be used in a conventional TEM.

Theory

Let us first consider the simplest example of the Fourier phase-retrieval method proposed by Bates and Rodenburg (1989), and described in detail by Rodenburg and Bates (1992). Suppose we have an unknown, one-dimensional function, $f(r)$, which we wish to image by scanning across it an aperture, $a(r)$, filled with constant-phase coherent illumination. Rather than simply measuring the total transmitted intensity, suppose we can record the intensity of the far-field Fraunhofer diffraction pattern which arises from $a(r-\rho)f(r)$, where ρ is the displacement of the aperture relative to $f(r)$. That is to say we can measure $|M(r',\rho)|^2$, where:

$$M(r',\rho) = \int a(r-\rho)f(r) \exp(i2\pi r.r') \, dr, \quad (1)$$

and where r' is a reciprocal space coordinate in the Fraunhofer diffraction plane. This could be very easily realized on the optical bench by employing the configuration shown in Figure 1.

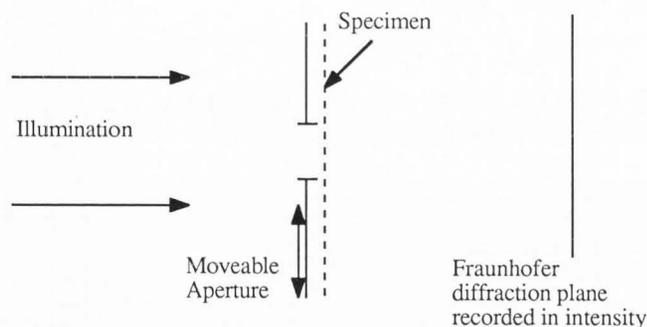


Figure 1. The simplest optical configuration which would allow measurement of the $|M(r,\rho)|^2$ data set. An aperture scans across the specimen while the intensity of the Fraunhofer diffraction plane is recorded in the far-field.

Now consider the nature of the function $|M(r,\rho)|^2$ in the very simplest case of a square ('top hat') aperture function and a specimen consisting of two real delta functions (see Figure 2). Although this example is somewhat artificial, it is useful for understanding how direct phase-retrieval is possible from an intensity function such as $|M(r,\rho)|^2$ composed of both real-space and reciprocal-space coordinates. At the values of ρ where the aperture function overlaps both delta functions, we will measure in the far-field a set of Young's slits interference fringes. At positions either side of this, we see either constant intensity (as a function of r') or no intensity at all, depending on whether the aperture overlaps either one or none of the delta functions. Taking the back Fourier transform of $|M(r,\rho)|^2$ with respect to r' will yield a function of real-space coordinates which we will call $L(r,\rho)$, and which is illustrated in Fig 2c. For each value of ρ , a one-dimensional horizontal strip across the function (as drawn in Fig 2c) is simply the autocorrelation of the region of specimen illuminated by the aperture when it was positioned at ρ . We have obtained the autocorrelation (or 'Patterson function', Patterson, 1934) because only intensity was measured in the far-field and so any one such strip of data presents us with the usual ambiguities of the classic phase problem (see, for example, Bates and McDonnell, 1986).

Now let us consider the information in $L(r,\rho)$ resolved as a function of ρ . Along $r=0$, this is simply the intensity transmission of the specimen convolved with the intensity of the aperture function. Except for those frequencies where zeros occur in the Fourier transform of the aperture function, this line of data can in principle be deconvolved to arbitrarily good resolution (because in this example the aperture has sharp edges), giving an accurate representation of the intensity of the specimen function. Furthermore, we have the opportunity to deconvolve along other strips in $L(r,\rho)$ where $r \neq 0$. The deconvolution may be represented as a filtering process in $H(r,\rho')$, which we define as the Fourier transform of $L(r,\rho)$ with respect to ρ : in other words, $H(r,\rho')$ is the Fourier transform of $|M(r,\rho)|^2$ with respect to both r' and ρ . We may write this as:

$$H(r,\rho') = \iiint a(b-\rho)a^*(c-\rho) f(b)f^*(c) \exp(i2\pi[b.r'-c.r'-r.r'+\rho.\rho']) db dc dr d\rho, \quad (2)$$

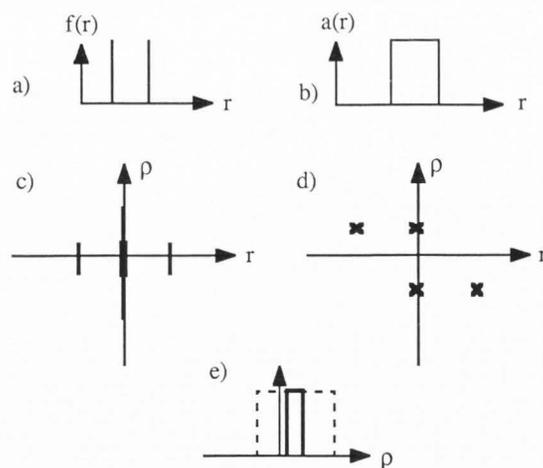


Figure 2. An illustration of how the deconvolution process retrieves phase. a) The function $f(r)$, consisting of two real-valued delta functions. b) The aperture function $a(r)$. c) The two-dimensional function $L(r,\rho)$. Heavy lines represent areas of large magnitude. d) The two-dimensional function $L_{decon}(r,\rho)$. Crosses represent the position of delta functions. These may have complex amplitude if $f(r)$ is complex. e) The function $a(\rho)a(\rho+r)$ plotted as a function of ρ along $r=r_0$, where r_0 is the separation of the delta functions. The nature of the offset narrow aperture breaks the symmetry about $r=0$ in $L_{decon}(r,\rho)$.

where b and c have been introduced as dummy variables and $*$ denotes the complex conjugate. After some elementary manipulation, it can be shown that:

$$H(r,\rho') = \chi_a(r,-\rho') \chi_f(r,\rho') \quad (3)$$

where for any general function q , we define:

$$\chi_q(r,\rho') = \int q^*(c)q(c+r) \exp(i2\pi c.\rho') dc. \quad (4)$$

For the purely real aperture function illustrated in Fig 2b, equations 3 and 4 indicate that each one-dimensional strip in $L(r,\rho)$, taken as a function of ρ along any value of r , is simply $f^*(\rho)f(\rho+r)$ convolved with an aperture function $a^*(\rho)a(\rho+r)$. Of course along $r=0$, this reduces to the convolution of intensity described above. By performing a deconvolution in the ρ -direction, a little thought will show that horizontal strips across the resulting deconvolved data set, which we will denote as $L_{decon}(r,\rho)$ and which is illustrated in Fig 2d, are no longer symmetric autocorrelation functions. This is because the function $a(\rho)a(\rho+r)$ along ρ at points $r=\pm r_0$, where r_0 is the separation of the delta functions, has the form of a narrow aperture function displaced with respect to the origin (Figure 2e). In other words we have resolved the usual ambiguity in the diffraction phase problem by utilizing, via the deconvolution, our knowledge of the absolute position of the aperture. This process is very closely related to Hoppe's idea (1969a,1969b; Hoppe and Strube,1969) of generalized diffraction (later referred to as "ptychography", Hoppe and Hegerl, 1980), though here all possible aperture positions are processed simultaneously (instead of the two positions used in the Hoppe construction).

Robust solution to the super-resolution phase problem

In summary, if we record $|M(r',\rho)|^2$, Fourier transform with respect to its two coordinates (which lie in both diffraction space and real space), divide (or preferable Wiener filter) by a function $\chi_a(r,-\rho')$ appropriate to the aperture function described by equation 4, and finally transform back with respect to the aperture position coordinate, we obtain:

$$L_{\text{decon}}(r,\rho) = f^*(\rho)f(\rho+r). \quad (5)$$

$L_{\text{decon}}(r,\rho)$ is complex and can be used to phase the whole specimen function relative to some arbitrary point. For example, if we were to decide that $f(0)$ had zero phase and a modulus of the square root of $L_{\text{decon}}(0,0)$, then we could write:

$$f(r) = L_{\text{decon}}(r,0)/f(0) \quad (6)$$

Unfortunately, this equation is only useful up to values of $|r|$ which are less than the width of the aperture function. (In r , $L(r,\rho)$ is only as wide as the autocorrelation function of the aperture). However, having determined the phase of some point in $f(r)$ within this distance, it is possible to re-use equation 5 along a value of ρ corresponding to the position of this new point, thus allowing us to phase a more extensive domain of $f(r)$. Indeed, provided there are no regions of zero value wider than the aperture, the phase of the whole of $f(r)$ can be determined by a series of similar steps. We refer to this process of phase assignment as "stepping out in ρ ". In fact, in the reciprocal space version of the formulation applicable to microscopy (next section), we step out in r' , but the principle is the same.

The origin of the phase information can be thought of as follows. If the delta functions in the above example had complex amplitudes of different phase, then the Young's slit diffraction fringes would be shifted laterally in the far-field, and hence their autocorrelation functions would also be complex. Indeed, the above analysis holds true even if both the aperture and the specimen functions are complex, implying that it represents a comprehensive direct solution to the phase problem. The only qualifications are that the zeros encountered in the deconvolution do not introduce too much error (see section on "Test Calculations") and that the specimen function does not have large zero regions wider than the aperture function.

The above analysis amounts to a more general statement that if we can record the intensity of spatially-resolved frequency distribution at appropriate sampling, such as that represented by, for example, a Gabor lattice (1950), then we can recover directly the complex amplitude of the original function.

Application to Transmission Microscopy

An immediate problem we encounter when applying the above method to transmission microscopy is that it is hard to manufacture a small, sharp aperture function which can be made to run across the specimen. However, we do normally have a sharp aperture lying in the back-focal plane of the objective lens. The trick, therefore, is to aim to solve for the Fourier transform of the specimen function, $F(r')$. Writing $|M(r',\rho)|^2$ as:

$$|M(r',\rho)|^2 = \iint a(b-\rho)a^*(c-\rho) f(b)f^*(c) \exp(i2\pi[b.r'-c.r']) db dc, \quad (7)$$

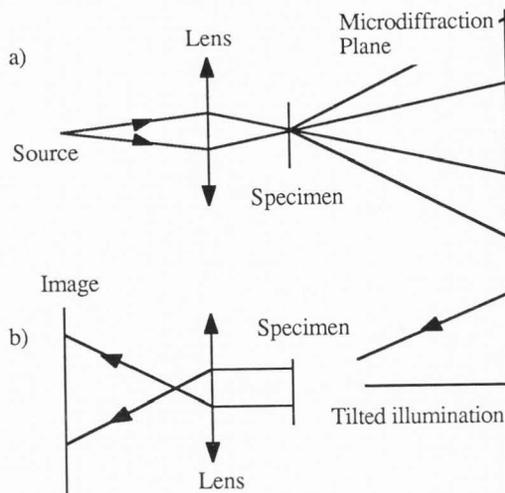


Figure 3. Two possible optical geometries for recording $|M(r',\rho)|^2$ in an aperture-limited microscope. a) By recording the microdiffraction pattern in a scanning transmission microscope for all beam-cross-over position. b) By recording images in a conventional microscope for each possible angle of illumination.

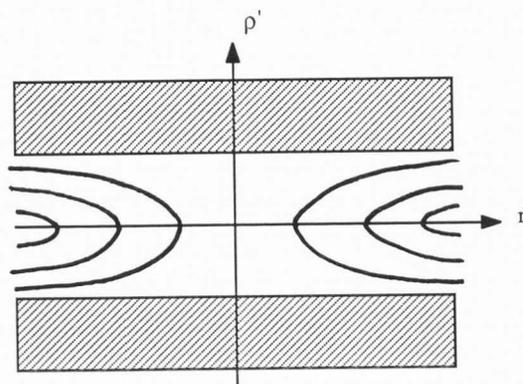


Figure 4. Schematic diagram of zeros in $\chi_A^*(\rho', r)$. Shaded regions lie beyond the autocorrelation of the aperture function. For a real-valued aperture function, zeros also occur on the lines shown.

where b and c are dummy variables, and substituting for $A(r')$ and $F(r')$, the Fourier transforms of $a(r)$ and $f(r)$ respectively, gives:

$$|M(r',\rho)|^2 = \iint A(b')A^*(c') F(r'-b')F^*(r'-c') \exp[i2\pi\rho.(b'-c')] db' dc', \quad (8)$$

Except for the exchange of ρ and r' , and the fact that we now scan our unknown function $F(r')$ relative to a fixed sharp aperture $A(r')$, equations 7 and 8 are of identical form. We may realize this arrangement optically by Figures 3a or 3b. Figure 3a shows a scanning transmission microscope in which a sharp aperture lying in the back focal plane of a lens is used to focus a beam cross-over through a specimen. In the far-field, or 'microdiffraction' plane (Cowley, 1978) lying in r' , we record the intensity of the complex convolution of the aperture function with the reciprocal space of the specimen

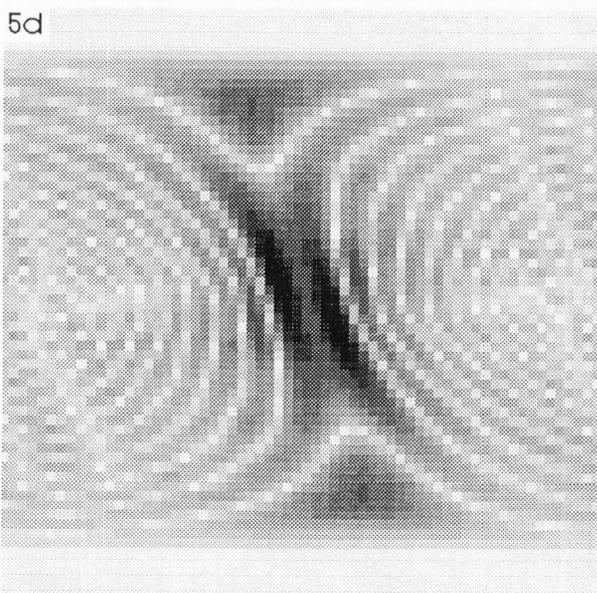
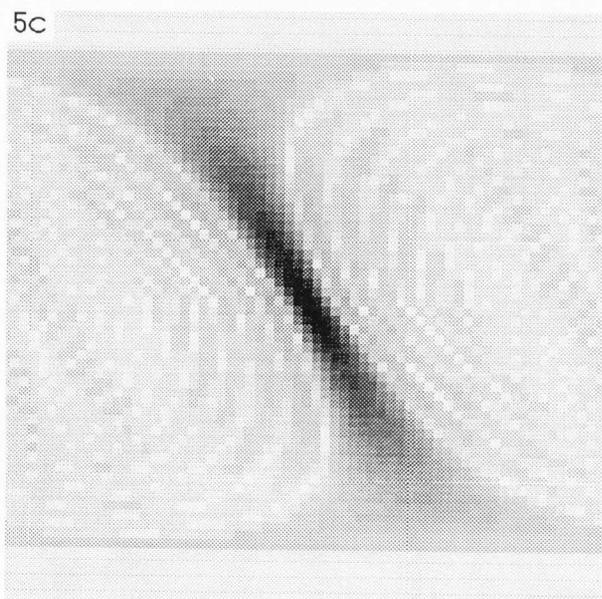
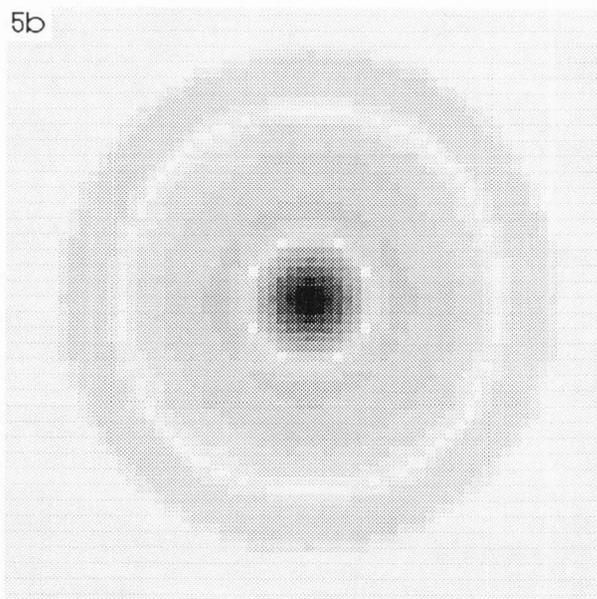
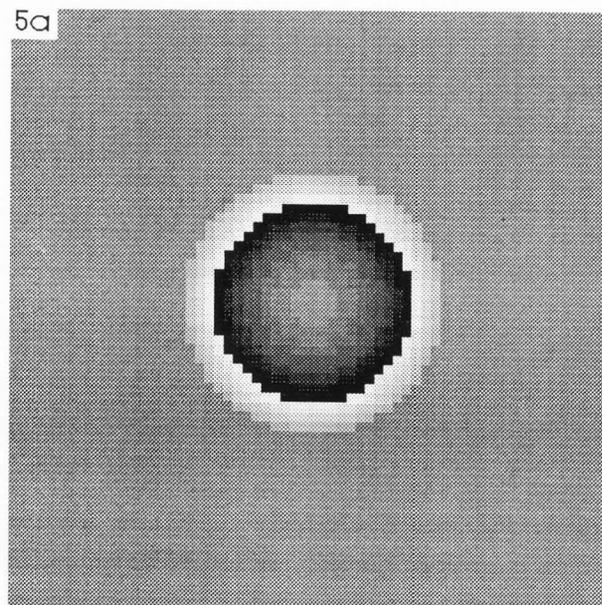


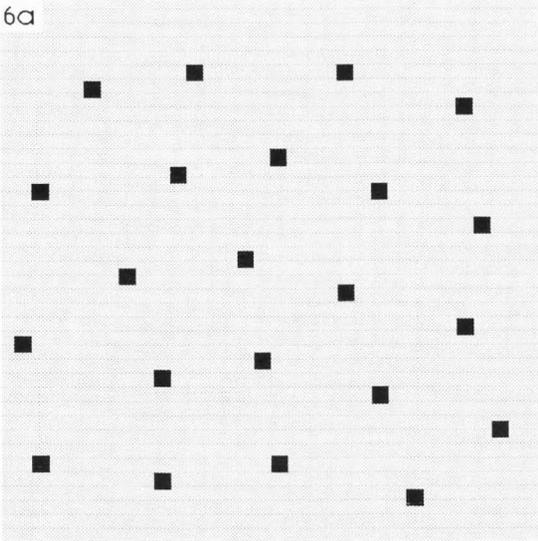
Figure 5. The function $\chi_A^*(\rho', r)$ calculated on a 64 pixel side hypercube. To accommodate the dynamic range of the function, in b, c and d we plot the square root of magnitude. a) Shows the two-dimensional aperture function plotted in phase (darker pixels are for positive phase, mid-grey is zero) calculated for the instrument characteristics described in the text. b) A 2D slice through $\chi_A^*(\rho', r)$ at $r=0$. c) A 2D slice through $\chi_A^*(\rho', r)$ with one component of r and ρ' held at zero. r is plotted horizontally. Note that the ρ' cut-off is visible. d) A slice parallel to that shown in (c) but with the constant component of ρ' at 8 pixels from the origin. (c) and (d) both intersect (b) along vertical lines.

Figure 6. Reconstruction of a real-valued specimen of delta functions with noise. Sampling in real space (i.e. the pixel size) is 0.5 Å. a) The specimen function (dark represents high value data). b) Magnitude of the probe function (calculated according the instrument parameters in text). c) Bright-field image (auto-scaled to full scale) with 10% noise. Light pixels represent high values of intensity. d) As (c) with 100% noise. e) Reconstruction with 100% noise, using only $r'=0$ data (resolution doubled). f) Reconstruction with 10% noise up to twice the ρ' cut-off. →

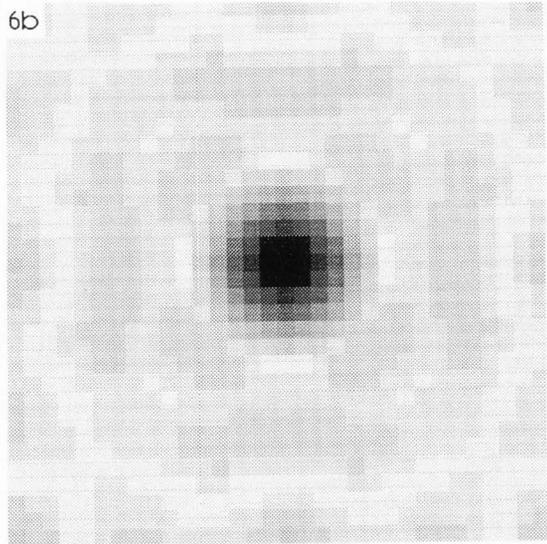
function, for all cross-over positions ρ . Alternatively, by invoking the principle of reciprocity (Cowley, 1969), this is the same as illuminating a specimen in a conventional microscope from a range of different angles r' (Figure 3b) and

Robust solution to the super-resolution phase problem

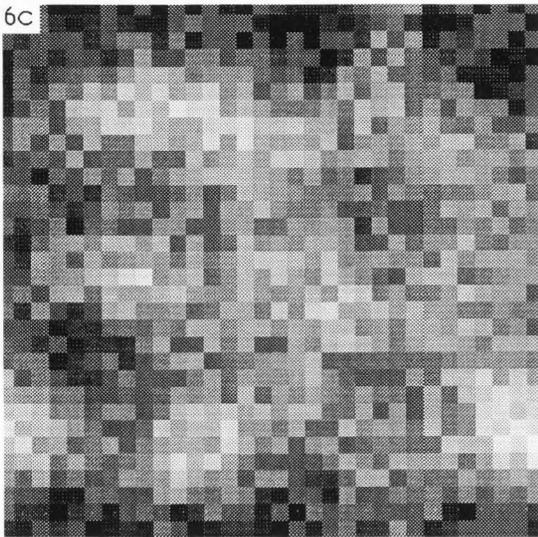
6a



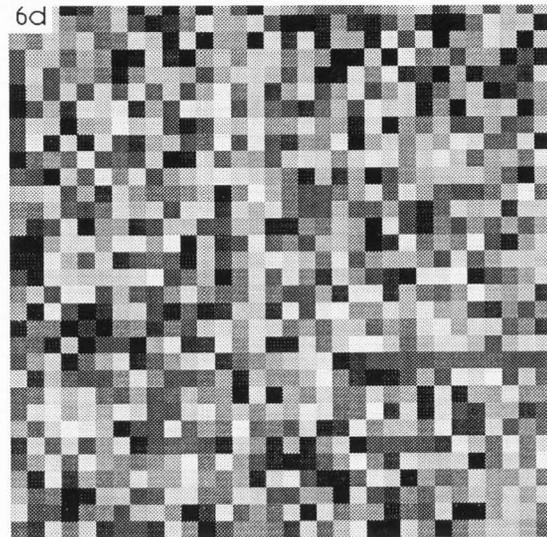
6b



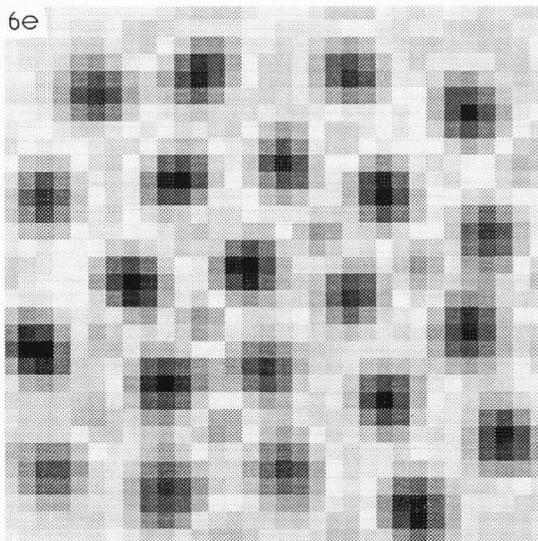
6c



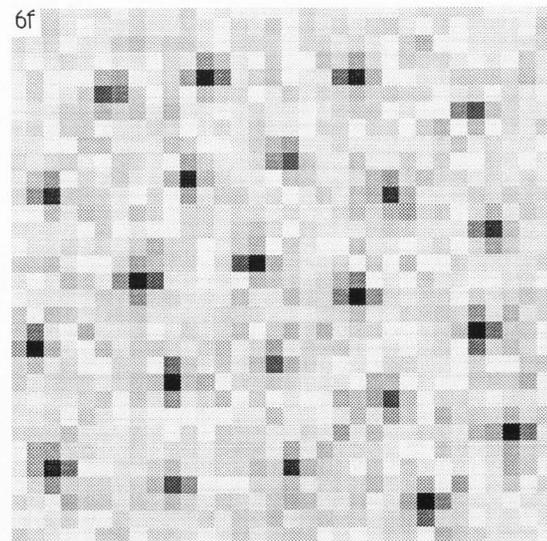
6d



6e



6f



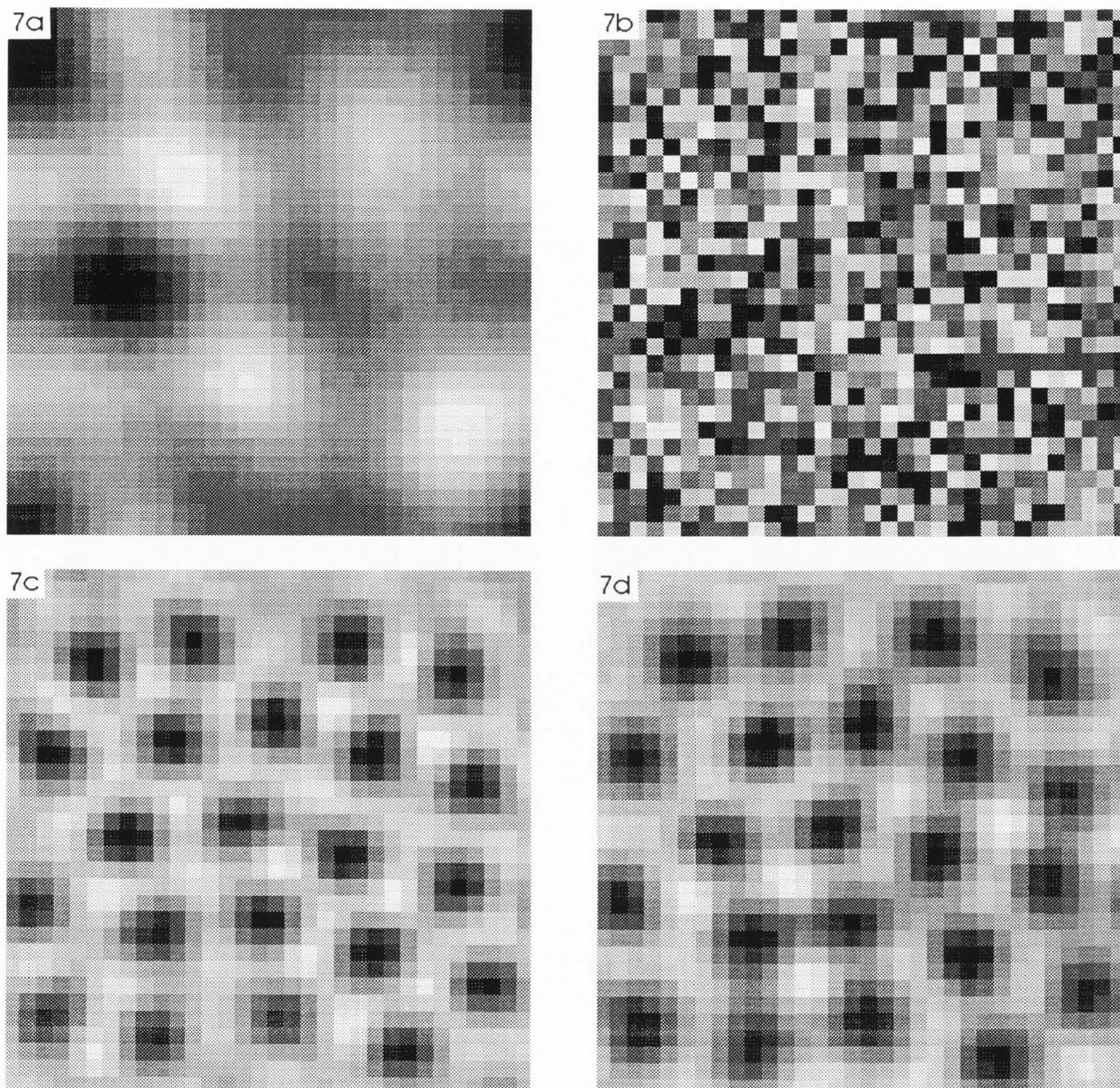
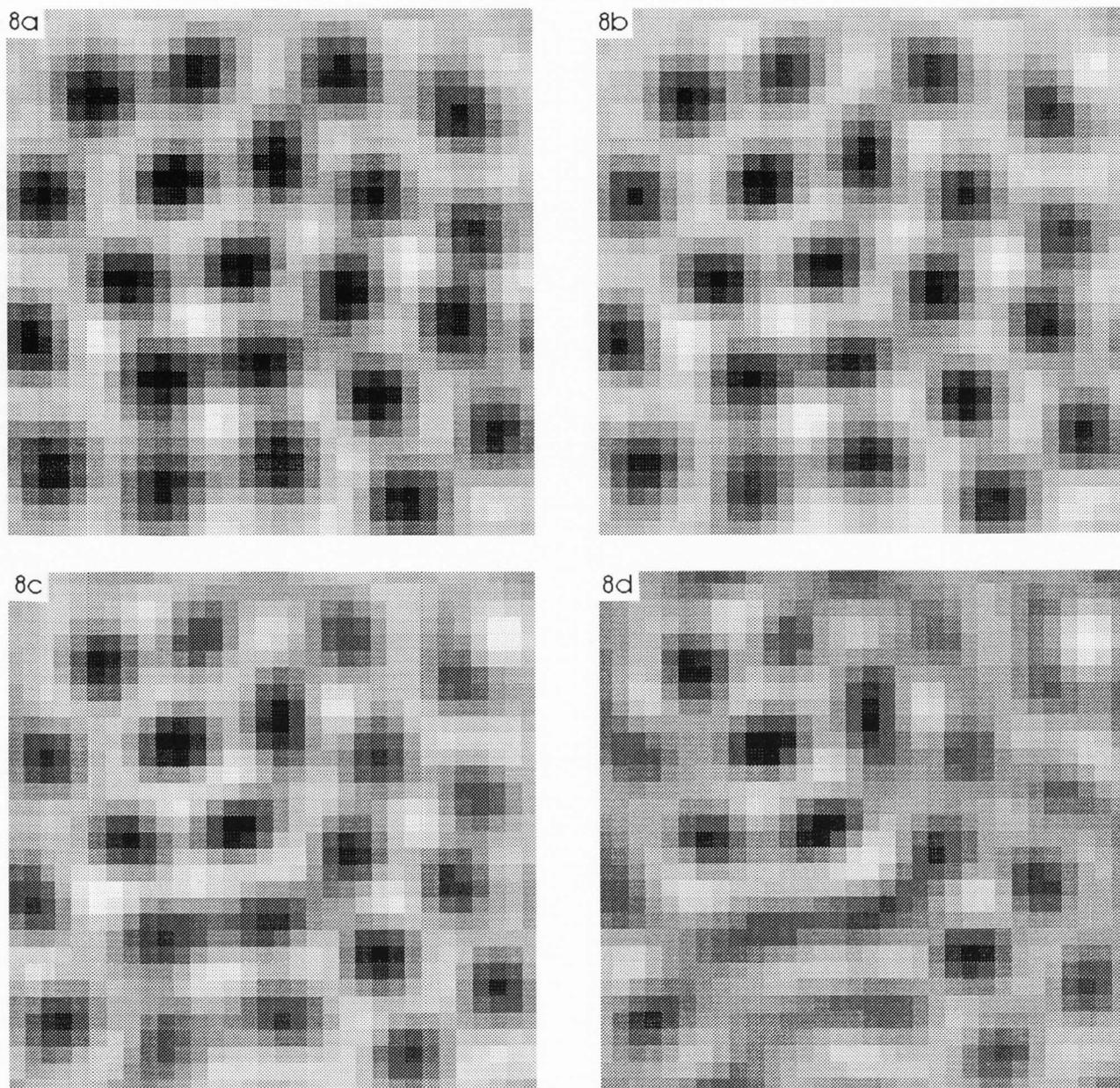


Figure 7. Reconstruction of a weak phase object. The specimen is the same as in Figure 6 except each delta function now only introduces a $\pi/20$ phase change on a unity magnitude transmission specimen. (a) and (b) are bright-field images with zero and 4% noise added respectively (light represents high intensity: autoscaled to full scale). (c) The phase (dark represents positive) of the specimen with its Fourier space limited to the ρ' cut-off (i.e. this is the best we could hope for with the reconstruction). (d) Calculated reconstruction with 4% noise added to the measured data set.

recording the image as a function of ρ . Those familiar with electron microscopy will know that in selected area diffraction mode, tilting the illumination has the effect of scanning the diffraction plane of the specimen with respect to the objective

aperture. Furthermore, for each such illumination angle, one can obtain the intensity of the Fourier transform of the resulting function: namely the conventional image. So for the simple case of a series of delta functions in $F(r')$ (ie for a crystalline specimen) we could apply the analysis developed in the previous section to obtain the complex value of all the reciprocal lattice points. For a given angle of tilt, the position of the interference fringes which occur in the image would indicate the relative phase of the beams falling within the objective aperture. It follows that any general specimen function may be solved in the same way, even when the aperture function is complex.

From equation 7 we progress, as before, by forming the quantity $H(r, \rho')$. We then divide by $\chi_a(r, -\rho')$, prescribed by



equation 4, which may be written in terms of the function $A(r')$ as:

$$\chi_a(r, -\rho') = \int A(c') A^*(c' + \rho') \exp(-i2\pi c' \cdot r) dc', \quad (9)$$

where it is useful to note that according to the definition in equation 4, the relationship between the real- and reciprocal-space versions of the χ function is:

$$\chi_a(r, -\rho') = \chi_A^*(\rho', r). \quad (10)$$

With reference to the schematic diagram shown in Figure 4, and by examining equation (9), it is clear that for a sharp aperture with no phase changes, zeros occur at all values of

Figure 8. Effect of a wrongly-estimated value of defocus in the probe/aperture function $\chi_A^*(\rho', r)$. The forward calculation to generate $|M(r, \rho)|^2$ was performed on a weak phase object (as in Figure 7) with defocus of -700\AA and 4% noise was added. Reconstructions were performed using only $r'=0$ with a) defocus = -700\AA , b) defocus = -400\AA , c) defocus = -200\AA d) defocus = zero.

$|\rho'|$ greater than the aperture width and on the loci of points represented by the curved lines. It is therefore advantageous to perform the final Fourier transform with respect to r , to avoid problems with the $|\rho'|$ cut-off, and thus form the quantity:

$$D_{\text{decon}}(r', \rho') = F^*(r' - \rho') F(r') \quad (11)$$

which is the reciprocal-space version of the analagous quantity $L(r,\rho)$. Phase retrieval may proceed as before, except now we "step out" in r' as far as we wish, giving us unlimited access to high-resolution information, provided $F(r')$ does not have large regions which are empty. A final Fourier transform gives the super-resolution reconstruction of $f(r)$.

The cut-off in ρ' can be thought of as the usual information limit of the microscope - remember that ρ is equivalent to a coordinate in the conventional image plane, which of course has a maximum frequency component dictated by the size of the objective aperture. Because the image is recorded in intensity, the maximum spatial frequency is twice that of the underlying complex component, which implies that if after performing the phase-retrieving deconvolution we do not step out at all in r' but merely use the $r'=0$ strip in $D(r',\rho')$ up to the ρ' cut-off (using an equation similar to equation 6 - see equation 13), then the final reconstruction will have twice the resolution of the conventional image. Although we use this effect in some of the examples below, it should be remembered that the extra information does not magically arise from nowhere. To perform the deconvolution in r' accurately, we must process at least an aperture's width of r' data from $|M(r',\rho)|^2$. In other words, we must process at least the whole undiffracted beam in the microdiffraction data set, and not merely the central pixel which is where we find the conventional bright-field image intensity.

Another significant property of $H(r,\rho')$ is that it can also be used to deconvolve directly partial coherence which exists in the illuminating beam. For the electron microscope, this is a crucially important result: it could be argued that the most severe difficulty impeding higher resolution is magnetic interference and power-supply instability which pose an absolute limit to resolution in reference-beam imaging (including deconvolution of the bright-field image and holographic techniques). In the case of a finite source in STEM (ie incoherence in the wavefield which illuminates the back-focal plane of the lens), we can employ the van Cittert-Zernike theorem (Born and Wolf, 1964) to calculate $\Gamma(r')$, the coherence function, from the normalized Fourier transform of the intensity distribution of the source. This may be accounted for in equation 8 by including a term $\Gamma(b'-c')$ under the integral signs, which serves to moderate the degree to which beams lying in the back-focal plane can interfere with one another in the microdiffraction plane. Surprisingly, however, the effect on $H(r,\rho')$ is to introduce only another multiplicative term as a function of ρ' , such that:

$$H(r,\rho') = \Gamma(-\rho') \chi_A^*(\rho',r) \chi_{F(-\rho',r)}^* \quad (12)$$

We can therefore also deconvolve Γ , and even if it does have a definite cut-off, it will not catastrophically compromise our reconstruction algorithm because, like the cut-off caused by the objective aperture, in $D_{\text{decon}}(r',\rho')$ we can simply take small steps in ρ' when stepping out to high-resolution data at large r' .

Test Calculations

The above method has very recently been shown to work on the optical bench in one-dimension (Friedman and Rodenburg, 1992). Here we will concentrate on presenting results from computer simulations we have performed on two-dimensional images. These have proved to be astonishingly

robust to noise on the measured data and errors in χ_A^* . One reason must be the fact that the measured data set is enormously redundant; remember that in order to solve for a two-dimensional (2D) image, we record a four dimensional data set (both ρ and r' become 2D vectors). What we gain in redundancy, though, results in a requirement for large computer resources. Essentially we have replaced a faulty parallel processor (an electron lens) with a digital computer. Simply keeping track of the data arrays is non-trivial. We have therefore implemented a four-dimensional version of the image-processing package IMPROC, developed by Dr. R. Lane at the University of Canterbury, New Zealand, which we run on a SUN 4/370 SPARC station with 40 Mbyte of on-line memory.

Figure 5 shows an amplitude plot of the χ_A^* function. Note that one can observe the lines of zeros represented in the schematic diagram Figure 4, though here they occupy asymmetric positions because we have included a phase changes across $A(r')$ determined by defocus and spherical aberration (Scherzer, 1949). All calculations have been performed assuming the optical characteristics of a VG Microscopes' HB501 STEM with standard resolution pole-piece: namely a spherical aberration constant of 3.1mm, an objective aperture semi-angle of convergence of 8 mrad, and a wavelength of 0.0037 nm. We were somewhat concerned that the zeros in χ_A^* would seriously compromise the fidelity of the reconstruction, but this does not seem to be the case in practice. The deconvolution is performed on $H(r,\rho')$ using a Wiener filter such that:

$$\chi_{f(r,\rho')} = \frac{\chi_A(\rho',r) H(r,\rho')}{|\chi_A(\rho',r)|^2 + \epsilon} \quad (13)$$

where ϵ is varied to obtain the best reconstruction, but is typically about 10^{-2} of the maximum modulus of $\chi_A(\rho',r)$. Note that in the $D(r',\rho')$, different spatial frequencies are absent at different values of ρ' , so presumably this offers enough redundancy to allow a good estimate of $F(r')$, though this will have to be the subject of further work.

Various examples of reconstructions are shown in Figures 6, 7 and 8. The calculations were performed over a hypercube of side 32 pixels, thus complex data sets such as $\chi_A^*(\rho',r)$ occupy 8 Mbyte double-precision arrays. Those labelled as "reconstructed from only $r'=0$ ", mean that we have not attempted any "stepping out" beyond the ρ' cut-off in $D(r',\rho')$. That is to say we have estimated $F(r')$ from:

$$F(r') = F(\rho') = \frac{D_{\text{decon}}^*(0,-\rho')}{\sqrt{D_{\text{decon}}(0,0)}} \quad (13)$$

In fact, exploiting all the phase information in $D_{\text{decon}}(r,\rho')$ is not straightforward, because for two-dimensional images there are many different routes through $D_{\text{decon}}(r',\rho')$ to reach any given point in $F(r')$ and choosing a measure for an optimal path is not obvious. Figure 6f shows reconstructions of a real delta function specimen where we have stepped out to four times the objective aperture radius. However, the phase-assignment routine we employ at present is rather elementary, only using a small number of r' -constant planes in $D_{\text{decon}}(r',\rho')$.

ρ'). In the case of a thin phase object, assigning phase beyond the ρ' cut-off becomes hazardous because any inaccuracies in the deconvolution corrupt significant regions of $D_{\text{decon}}(r', \rho')$. (Under these circumstances, the central disc in the microdiffraction plane is extremely bright.) However, using only $r'=0$ data (equation 13) still doubles the resolution relative to the conventional image (see section on "Applications to Transmission Microscopy"), and this appears to be exceptionally robust to noise. We have added random noise uniformly distributed with a range in proportion (measured as a percentage) of the largest intensity measurement in $|M(r', \rho)|^2$. The reconstruction is good for real delta functions (Figure 6) with up to 100% noise if we only phase over $r'=0$, or up to 10% when stepping up to twice the ρ' cut-off (four times the aperture radius). For a weak phase object (a unity modulus specimen with $\pi/20$ phase changes introduced at points corresponding to the delta functions in the earlier example), the reconstruction is good from $r'=0$ with up to 4% noise (see Figure 7). The reconstructions are also remarkably robust to having the wrong estimate of defocus in the aperture function used to generate $\chi_A^*(\rho', r)$ (up to 70 nm in the example shown in Figure 8). Of course, it should be remembered that because we are using a narrow objective aperture, the beam cross-over in the specimen plane does not change appreciably with this amount of defocus. The result is important, though, because defocus is hard to estimate and control accurately in the electron microscope.

Conclusion

The phase-retrieval deconvolution technique presented here is applicable to any form of transmission microscopy where a coherent source is available but where it is not possible to use a good quality lens of large numerical aperture. We have assumed that the specimen scatters multiplicatively and can be regarded as a 2D projection, which, for electron scattering from all but the thinnest specimens, will often not be the case. However, the robustness of the method, which arises because we collect such a comprehensive and redundant data set, suggests that this may prove to be a valuable processing technique. We are presently developing a suitable detector to collect the necessary data from the HB501 STEM in this laboratory, with the hope of achieving routinely sub-Ångstrom resolution. Undoubtedly, many problems will arise - for example, specimen drift and contamination, breakdown of the projection approximation and effects due to multiple scattering. The results shown here, though, give us some hope of success.

Acknowledgements

Both authors are grateful for financial support from the Royal Society and the SERC.

References

- Bates RHT, McDonnell MJ (1986). *Image Restoration and Reconstruction*. Clarendon Press, Oxford. pp 98-127.
- Bates RHT, Rodenburg JM (1989). Sub-Ångstrom transmission microscopy: a Fourier transform algorithm for microdiffraction plane intensity information. *Ultramicrosc.* **31**, 303-308.
- Born M, Wolf E (1964). *Principles of Applied Optics*. 2nd Edition. Pergamon, Oxford. p. 508.
- Cowley JM (1969). Image contrast in a transmission scanning electron microscope. *Appl. Phys. Lett.* **15**, 58-59.
- Cowley JM (1978). *Electron Microdiffraction*. *Adv. Elect. and Elect. Phys.* **46**, 1-53.
- Frank J (1973). The envelope of electron microscopic transfer functions for partially coherent illumination. *Optik* **38**, 519-536.
- Friedman SL, Rodenburg JM (1992). Optical demonstration of a new principle of far-field microscopy. *J. Phys. D* **25**, 147-154.
- Gabor D (1950). *Communication Theory and Physics*. *Phil. Mag.* **41**, 1161-1187.
- Hoppe W (1969a). Beugung im Inhomogenen Primärstrahlwellenfeld I. Prinzip einer Phasenuessung von Elektronenbeugungsinterferenzen (Diffraction with an inhomogeneous primary wave distribution. I. Principle of phase measurement from interfering diffracted waves). *Acta. Cryst.* **A25**, 495-501.
- Hoppe W (1969b). Beugung im Inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei imperiodischen Objekten (Diffraction with an inhomogeneous primary wave distribution. III. Amplitude and phase determination with non-periodic objects). *Acta Cryst.* **A25**, 508-514.
- Hoppe W, Hegerl R (1980). Three dimensional structure determination by electron microscopy (non-periodic specimens). In: *Computer Processing of Electron Microscope Images*. Hawkes PW (ed.). Springer, Berlin. *Topics in Current Physics* **13**, 127-185.
- Hoppe W, Strube G (1969). Beugung im Inhomogenen Primärstrahlwellenfeld. II. Lichtoptische Analogieveruche zur Phasenmessung von Gitterinterferenzen (Diffraction with an inhomogeneous primary wave distribution. II Light optical analogue studies on phase measurement from grating interference). *Acta. Cryst.* **A25**, 502-507.
- Patterson AL (1934). A Fourier series method for the determination of the components of the interatomic distances in crystals. *Phys. Rev.* **46**, 372-376.
- Rodenburg JM, Bates RHT (1992). The theory of super-resolution electron microscopy via Wigner-distribution deconvolution. *Phil. Trans. A* **339**, 521-553.
- Scherzer O (1949). The theoretical resolution limit of the electron microscope. *J. Appl. Phys.* **20**, 20-29.

Discussion with Reviewers

D. Van Dyke: The calculation time can increase prohibitively with increasing number of sampling points in real and reciprocal space. Is it not possible to find a shortcut or to restrict the operation to a small area?

Authors: Yes. Firstly, it is possible to tile together small areas to make a large image, which works out to be much cheaper computationally. Secondly, if one makes certain assumptions about the specimen or transfer function (e.g., the former is a weak phase object or that the latter is accurately known or has large regions of constant phase), then many numerical shortcuts become possible. In this respect, there is a pay-off between the generality of the calculation and its size.

D. Van Dyke: Can the robustness of the procedure be enhanced by changing the order of the operations?

Authors: The exact details of the most robust use of the data is certainly open to speculation. We have found that iterative solution methods can be more robust, but at the expense of longer calculation times.

D. Van Dyke: In principle, the resolution obtained after deconvolution is limited by the statistics of the probe (width versus signal to noise). For a STEM, these statistics are comparable to those of the point-spread function in TEM, in which the information limit is mainly determined by spatial and temporal coherence. Is it therefore too optimistic to speak about super-resolution?

Authors: We regard the crucial advantage of the method is that it is not limited by the coherence widths. We have modelled contributions from source size and defocus wobble and find their effects are small, or can be taken out by the deconvolution.

D. Van Dyke: The shape of the probe is determined by the focus. How can the focus itself be measured with sufficient accuracy to make the method work?

Authors: As the calculations demonstrate, the method is not very sensitive to errors in defocus. There are also many ways of estimating the defocus from the data set itself, such as the one proposed by Friedman and Rodenburg (1992) or by blind deconvolution methods.

D. Van Dyke: Does reciprocity lead to an equivalent procedure for TEM (e.g., beam rocking)?

Authors: Yes. The difficulty with TEM is that to perform the deconvolution satisfactorily for strong phase objects, many samples in r' (the equivalent of the beam-tilt coordinate) are required. This may cause the experiment to be prohibitively long with regard to specimen drift etc. The degree of beam deflection required (relative to the modest shifts needed in the STEM probe) may also create complications with hysteresis. Also, in processing the dark-field conditions, TEM would still be fully illuminating the specimen, thus leading to heavier specimen damage.