Subangstrom Resolution by Underfocused Incoherent Transmission Electron Microscopy

P. D. Nellist[1,*] and S. J. Pennycook[2]

[1]*Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom*
[2]*Oak Ridge National Laboratory, Solid State Division, P.O. Box 2008, Oak Ridge, Tennessee 37831-6030*
(Received 24 June 1998)

It is quantitatively explained why incoherent transmission electron microscope imaging is extremely robust to the effects of chromatic aberration, which usually limits the resolution in the conventional coherent mode of imaging. Combining this robustness with using underfocus to counter the effects of spherical aberration, we demonstrate subangstrom lattice resolution and information transfer to 0.078 nm. [S0031-9007(98)07430-4]

PACS numbers: 07.78.+s, 41.85.−p, 41.85.Gy, 61.16.Bg

Incoherent imaging in transmission electron microscopy (TEM) is a powerful technique that has already been shown to provide direct, structure images at atomic resolution [1]. In this Letter we show how it can also overcome the conventional resolution limits to achieve subangstrom (1 Å = 0.1 nm) resolution. The pursuit of ultrahigh resolution TEM is motivated not only by the desire to resolve individual atomic column spacings in a variety of crystal orientations [2], but also by the development of quantitative image analysis methods where the precision of measurement is limited by the resolution [3]. Unlike light microscopes, where the resolution approaches the wavelength limit, the electron optics limit TEM resolution to approximately 100 times larger than the de Broglie wavelength of the electrons used. Inherent spherical aberration [4] corrupts high spatial frequency image information, and chromatic aberration leads to a coherence envelope in conventional TEM [5] that limits the resolution of information imaged by the microscope, corrupted or not. Here we develop a theoretical understanding of why incoherent imaging is much less susceptible to chromatic aberration than conventional TEM. Combining this robustness with the use of underfocus to counter the effects of spherical aberration, we achieve subangstrom resolution, resolving a projected atomic spacing of 0.093 nm and showing information transfer to better than 0.078 nm, in a microscope with a conventional point resolution for weak-phase object specimens of 0.19 nm.

Hitherto, high-resolution TEM imaging has been performed mostly using conventional TEM (which we refer to as "coherent TEM") in which a thin, electron transparent specimen is illuminated with almost perfectly plane-wave illumination. An objective lens after the specimen forms a magnified image. Deconvolution of the effects of spherical aberration from such an image requires that the phase of the image-plane wave function be retrieved, since only the intensity is detected. Of the various methods for phase retrieval and correction of spherical aberration (for example, [6–9]), none has achieved subangstrom imaging. Chromatic aberration means that instabilities in the objective lens current and the spread of incident electron energies

create a spread in the objective lens focal length. The resulting coherence envelope imposes an "information limit" on the resolution of information that may be passed in any one micrograph. Alternatively, high-voltage microscopes reduce the electron wavelength yet further. Their resolution is still restricted by the lens aberrations, but there is an improvement in resolution, and using a machine operating at 1250 kV, Ichinose has clearly resolved a spacing of 0.109 nm in a specimen of SiC (see Fig. 10 in Ref. [2]), and there is evidence that the atomic columns in Nb along the ⟨112⟩ orientation have been resolved, with a smallest spacing of 0.095 nm [10].

If instead of plane-wave illumination the specimen is illuminated by a large incoherent electron source, an incoherent image will be formed. This situation for light optics was first analyzed by Lord Rayleigh [11], who showed that interference effects between the radiation emitted from spatially separated parts of the specimen were suppressed, and that the resolving power of the microscope was improved by a factor of 2 relative to the coherent case. Efficient TEM implementation of this mode uses a scanning transmission electron microscope (STEM), in which an electron lens before the specimen focuses a beam of electrons to form a small probe. The probe is scanned over the specimen, and a variety of signals can be detected as a function of probe position, forming an image. By the principle of reciprocity [12], the detector in the far field of a STEM is equivalent to the electron source in a conventional TEM, so incoherent imaging conditions in a STEM simply require a large detector. Because TEM specimens are mainly phase objects to the high-energy electron beam, with little absorption, a detector that collects all of the electrons transmitted by the specimen would show very little contrast. Instead, a detector is used that has a hole at its center forming an annulus leading to annular dark-field (ADF) imaging. As long as the detector geometry is large relative to the convergence angle of the incident focused electron beam, incoherent imaging will result [13].

The electron wave function of a focused STEM probe can be considered as a superposition of coherent partial plane waves in the illuminating cone of electrons,

$$P(\mathbf{R}) = \int H(\mathbf{K}_i) \exp[i\chi(\mathbf{K}_i)] \exp(-i2\pi\mathbf{K}_i \cdot \mathbf{R})\, d\mathbf{K}_i,$$

$$(1)$$

where $\mathbf{K}_i$ is the transverse component of the incident partial plane wave, H is a circular top-hat function representing the position and size of the objective aperture which blocks the higher angle more spherically aberrated beams, and

$$\chi(\mathbf{K}_i) = \pi z\lambda|\mathbf{K}_i|^2 + \tfrac{1}{2}\pi C_S\lambda^3|\mathbf{K}_i|^4 \qquad (2)$$

is the phase shift of the incident plane wave due to spherical aberration, C_S, and defocus of the lens, z. An incoherent image can be written as a convolution between the intensity of the illuminating probe $|P(\mathbf{R})|^2$ and an object function that consists of localized peaks at the atomic-column positions [3,13]. The transfer function for incoherent imaging, $T(\mathbf{Q})$, is thus the Fourier transform of the probe intensity. Since the probe's complex amplitude in (1) is the inverse Fourier transform of $H(\mathbf{K}_i)\exp(i\chi)$, $T(\mathbf{Q})$ can be written as an autocorrelation [14],

$$T(\mathbf{Q}) = \int H(\mathbf{K}_i)\exp[i\chi(\mathbf{K}_i)]H(\mathbf{K}_i - \mathbf{Q})$$

$$\times \exp[-i\chi(\mathbf{K}_i - \mathbf{Q})]\, d\mathbf{K}_i, \qquad (3)$$

so the contribution to an image at a spatial frequency, $\mathbf{Q}$, comes from the integration over all pairs of available partial plane waves in the incident cone separated by the vector, $\mathbf{Q}$, which constrains $\mathbf{K}_i$ to lie in the overlap region between the two discs of the H functions (Fig. 1).

The phase of the integrand of (3) is the difference between two quartics of the form given in (2), one centered at the origin and one at $\mathbf{Q}$ (Fig. 1). To maximize the transfer, negative defocus is used to counteract the effects of the spherical aberration so that this phase difference varies as little as possible within the disc overlap region. A VG Microscopes HB603U STEM with an electron energy of 300 keV and a coefficient of spherical aberration
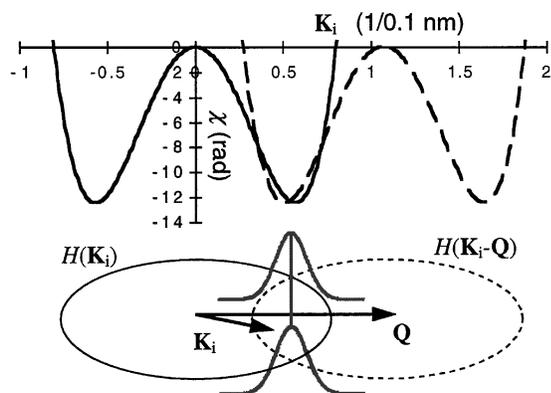
of 1 mm was used for these experiments. For these values the optimum "Scherzer" [15] transfer function is shown in Fig. 2(a). The maximum frequency transferred in (3) will be twice the objective aperture radius, which contrasts with coherent imaging where spatial frequencies only up to the aperture radius will be imaged.

If we now underfocus the objective lens by a much greater amount, the turning points of χ will occur for a much larger $\mathbf{K}_i$. For values of $\mathbf{Q}$ around twice that of the turning point, there will be a region of relatively flat phase in the integrand of (3) (Fig. 1) and significant transfer at higher spatial frequencies, as long as a large enough objective aperture is used, as has been demonstrated in a lower resolution instrument [16]. Figure 2(b) shows how significant transfer can be achieved at a spatial frequency of $(0.093\ \text{nm})^{-1}$. The corresponding probe intensity [Fig. 2(c)] has a narrower central maximum but also long oscillating side lobes. The lack of a phase problem means that this more complicated probe could now be deconvolved from the experimental data [3]. The gaps in the transfer could then be filled by combining data from images taken at other focus settings.

Using coherent TEM images at different focus values was initially proposed by Schiske [17], and a resolution improvement has been demonstrated [8]. Unfortunately, this method is still limited by the chromatic defocus spread, Δ, which gives rise to a damping envelope in frequency space [5],

$$E_{\text{chr}}(\mathbf{Q}) = \exp[-\tfrac{1}{2}\pi^2\lambda^2\Delta^2|\mathbf{Q}|^4], \qquad (4)$$
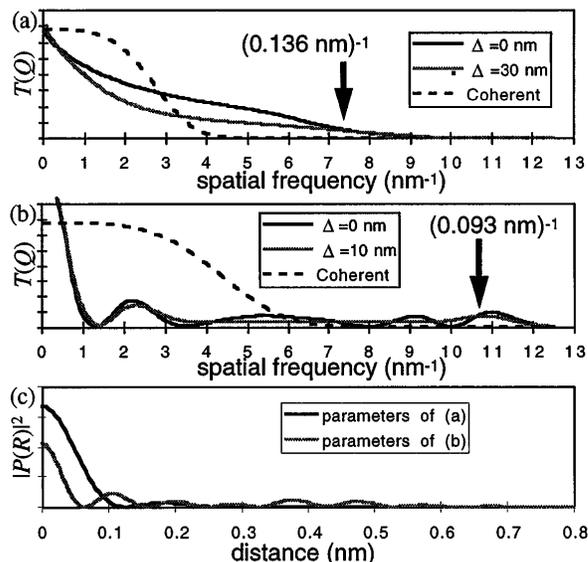

FIG. 2. Incoherent transfer functions with: (a) 44 nm and 9.4 mrad, (b) 125 nm and 17 mrad underfocus and objective aperture radius respectively, plotted on the same scale, and (c) the radial intensity profile of the probes corresponding to (a) and (b). Incoherent transfer functions are also shown for a defocus spread of Δ, along with the corresponding envelope for coherent imaging. The value of the transfer function at zero spatial frequency is proportional to the square of the objective aperture radius [21].

FIG. 1. A schematic diagram of the integrand of Eq. (3) for $|\mathbf{Q}| = (0.093\ \text{nm})^{-1}$. The quartic form of the phase variation across each disc is also shown for a defocus of -125 nm. The chromatic defocus spread gives rise to a modulating envelope in Eq. (5), shown schematically in grey.

for a Gaussian defocus spread. To examine the effect of defocus spread on incoherent imaging, we must integrate (3) over the Gaussian defocus distribution. We can use an analysis developed for overlapping discs in coherent

microdiffraction patterns to derive the effects of the defocus spread on the integrand of (3) [Eq. (11) of Ref. [18]], then integrate over the disk overlap to give the incoherent transfer function,

$$T_{\mathrm{chr}}(\mathbf{Q}) = \int H(\mathbf{K}_i)H(\mathbf{K}_i - \mathbf{Q}) \exp[i\chi(\mathbf{K}_i) - i\chi(\mathbf{K}_i - \mathbf{Q})] \exp[-\tfrac{1}{2}\pi^2\lambda^2\Delta^2(2\mathbf{K}_i \cdot \mathbf{Q} - |\mathbf{Q}|^2)^2]\,d\mathbf{K}_i. \quad (5)$$

So for incoherent imaging, the chromatic defocus spread does not lead to a simple envelope modulating the transfer function. For any value of $\mathbf{K}_i$ along the perpendicular bisector of $\mathbf{Q}$ (the condition for which is $\mathbf{K}_i \cdot \mathbf{Q} = |\mathbf{Q}|^2/2$) there is no attenuation of the integrand (Fig. 1). This is interference between "achromatic" beams in the incident cone of illumination that make the same angle to the optic axis, and therefore suffer the same phase shift due to defocus. Parallel to $\mathbf{Q}$, there is a Gaussian attenuation of the integrand with a width proportional to $1/|\mathbf{Q}|$. Either the aperture size, H, or the coherence envelope can limit the range of the integral in (5). If Δ is sufficiently large for coherence to be the limiting factor, then the transfer function will approximately follow a $1/|\mathbf{Q}|$ dependence, and this can be seen in the midrange frequencies in Fig. 2(a). The $1/|\mathbf{Q}|$ dependence gives a much less severe truncation of the information transfer in reciprocal space than the equivalent envelope for coherent imaging, (4). Counterintuitively, chromatic spread has a minimal effect near the resolution limit, because the optimum transfer function follows an approximately linear dependence on $|\mathbf{Q}|$, and will eventually drop below the $1/|\mathbf{Q}|$ chromatic coherence dependence. The physical interpretation is that near the resolution limit, the interfering beams in the incident cone must be almost diametrically opposite, and therefore achromatic, with respect to each other. The lack of chromatic attenuation in the center of the disc overlaps also applies to ptychography [18] and by reciprocity to tilted beam [7] and hollow cone illumination [19] in conventional TEM, which all show robustness to chromatic defocus spread for certain spatial frequencies. These modes are not incoherent so, except for extremely thin specimens, their complicated transfer functions cannot be directly deconvolved and their chromatic envelopes show a sharp cutoff at twice the illumination tilt angle, unlike the $1/|\mathbf{Q}|$ behavior shown here. These modes have never achieved subangstrom imaging.

Under Scherzer conditions using ADF imaging, the HB603U microscope is capable of resolving the so-called "dumbbell" 0.136 nm projected column spacing of Si⟨110⟩. To reduce the transfer at the $(0.136\ \mathrm{nm})^{-1}$ spatial frequency to 0.75 of its original value an extremely large value of $\Delta = 30$ nm is required. At these values, the envelope for coherent imaging (4) has a value of 4×10^{-17} and to achieve an attenuation by 0.75 in the coherent case a value of $\Delta = 2.3$ nm is required, which is much harder to achieve experimentally. Thus underfocused ADF imaging has great potential for resolving subangstrom spacings. When viewed in the ⟨112⟩ projec-

tion, CdTe has a rectangular lattice of pairs of Cd and Te columns separated by 0.093 nm, which are not resolvable by the HB603U under Scherzer conditions (Fig. 3). Underfocusing by 125 nm, however, puts a passband at the $(0.093\ \mathrm{nm})^{-1}$ spatial frequency, and even a chromatic defocus spread of $\Delta = 10$ nm reduces the transfer at that frequency by only a factor of 0.9 [Fig. 2(b)], compared with $\Delta = 0.6$ nm that would be required for the same reduction in transfer in coherent imaging. Performing such an experiment does indeed resolve the pairs of atomic columns, and the Fourier transform of the image shows the transfer out as far as the {444} plane spacing, which corresponds to the 0.093 nm column separation. It is remarkable that a 300 kV microscope can achieve such a resolution, which has previously required costly and complex high-voltage microscopes [10].

So far, we have assumed an infinitesimal electron source in the STEM. A finite source can be treated as an ensemble of incoherent point sources, which leads to the probe intensity being convolved with the source distribution after taking into account the demagnification effects
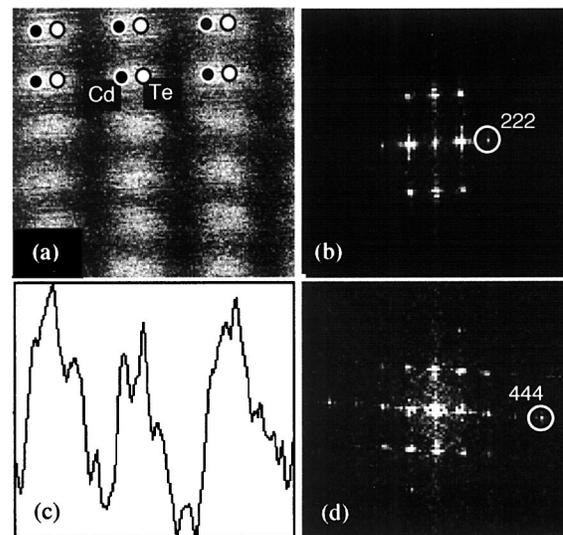

FIG. 3. (a) An image of CdTe⟨112⟩ taken under Scherzer imaging conditions and (b) the magnitude of its Fourier transform. (c) The profile plot, by summing vertically over 200 pixels, of an image of CdTe⟨112⟩ recorded using a 17 mrad objective aperture and 125 nm underfocus and (d) the magnitude of its Fourier transform. The {444} fringes can be seen in (c) and the corresponding spot in (d). Although the probe now has large sidelobes [Fig. 2(c)], it main maximum is still central, and so the split main peaks in (c) correspond to the Cd and Te columns being resolved.

of the objective and condenser lenses. This convolution leads to the transfer function being modulated by an envelope which is the Fourier transform of the source intensity distribution. At higher spatial frequencies this envelope will tend to zero, leading to another information limit. In principle this information limit can be moved out in reciprocal space by further demagnifying the electron source, but at the expense of the signal-to-noise ratio which will itself provide an information limit. To determine just how far out in reciprocal-space underfocusing can retrieve information, we performed an experiment using Si in the $\langle 110 \rangle$ orientation. Under Scherzer defocus, all the atomic columns are resolved showing the 0.136 dumbbell spacing (Fig. 4). By underfocusing the probe, we found that the {444} spatial frequency of the object function could be transferred, corresponding to a real-space distance of 0.078 nm.

In conclusion, we believe that the route to sub-angstrom resolution TEM is through combining spherical aberration compensation or correction with incoherent imaging, which is extremely robust to the loss of coherence due to chromatic aberrations, usually the final limiting factor for coherent TEM. Recently there has been activity in building a spherical aberration corrector for STEM [20], but there is currently no correction of chromatic aberration. For such a corrector, we suggest that its major application for resolution improvement will b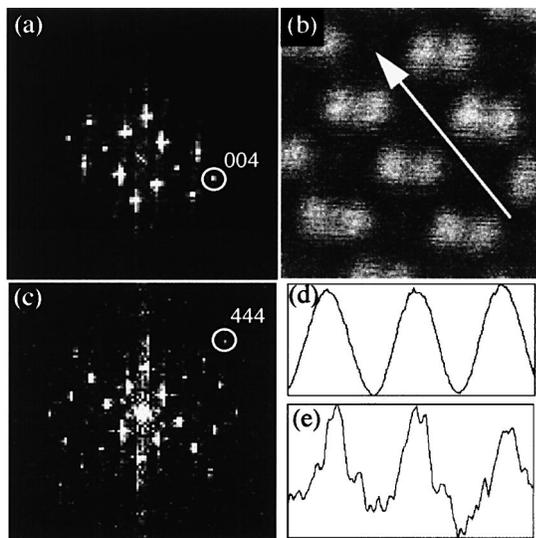e in incoherent imaging. It is yet to be shown whether such correctors can produce ultrahigh resolution images. Here we have overcome the resolution limit due to spherical aberration by using underfocus to compensate, resolving an atomic-columns spacing of 0.093 nm, and showing information transfer down to 0.078 nm.

---

*Now at Nanoscale Physics Research Laboratory, School of Physics and Astronomy, The University of Birmingham, U.K.
Electronic address: P.D.Nellist@bham.ac.uk

[1] S. J. Pennycook and D. E. Jesson, Phys. Rev. Lett. **64**, 938 (1990).
[2] D. J. Smith, Rep. Prog. Phys. **60**, 1513 (1997).
[3] P. D. Nellist and S. J. Pennycook, J. Microsc. **190**, 159 (1998).
[4] O. Scherzer, Z. Phys. **101**, 593 (1936).
[5] R. H. Wade and J. Frank, Optik **49**, 81 (1977).
[6] P. D. Nellist, B. C. McCallum, and J. M. Rodenburg, Nature (London) **374**, 630 (1995).
[7] A. I. Kirkland *et al.,* Ultramicroscopy **57**, 355 (1995).
[8] W. Coene *et al.,* Phys. Rev. Lett. **69**, 3743 (1992).
[9] A. Orchowski, W. D. Rau, and H. Lichte, Phys. Rev. Lett. **74**, 399 (1995).
[10] G. Möbus *et al.,* J. Electron Microsc. **46**, 381 (1997).
[11] Lord Rayleigh, Philos. Mag. **42**, 167 (1896).
[12] J. M. Cowley, Appl. Phys. Lett. **15**, 58 (1969); E. Zeitler and M. G. R. Thomson, Optik **31**, 258 (1970); **31**, 359 (1970).
[13] D. E. Jesson and S. J. Pennycook, Proc. R. Soc. London A **441**, 261 (1993); R. F. Loane, P. Xu, and J. Silcox, Ultramicroscopy **40**, 121 (1992).
[14] G. Black and E. H. Linfoot, Proc. R. Soc. London A **239**, 522 (1957).
[15] O. Scherzer, J. Appl. Phys. **20**, 20 (1949).
[16] D. H. Shin, E. J. Kirkland, and J. Silcox, Appl. Phys. Lett. **55**, 2456 (1989).
[17] P. Schiske, in *Image Processing and Computer Aided Design in Electron Optics* (Academic, New York, 1973), p. 82.
[18] P. D. Nellist and J. M. Rodenburg, Ultramicroscopy **54**, 61 (1994).
[19] W. K. Jenkins and R. H. Wade, in *Proceedings of the Electron Microscopy and Analysis Group Conference, EMAG77, Glasgow, 1977,* edited by D. L. Misell, IOP Conf. Proc. No. 36 (Institute of Physics, Bristol, 1977).
[20] O. L. Krivanek *et al.,* in *Proceedings of the Electron Microscopy and Analysis Group Conference, EMAG97, Cambridge, 1997,* edited by J. M. Rodenburg, IOP Conf. Proc. No. 153 (Institute of Physics, Bristol, 1997).
[21] Use of a larger objective aperture does worsen the image signal-to-noise ratio because of the increased uniform background, though the noise scales only as the square root of the image intensity.

FIG. 4. (a) The magnitude of its Fourier transform showing transfer to the {004} spacing at 0.136 nm, and (b) an image of Si$\langle 110 \rangle$ showing the dumbbells resolved. (c) The magnitude of the Fourier transform of a highly underfocused image showing that information to the {444} spacing (0.078 nm) is being transferred. Projecting the image along 200 pixels in the direction indicated by the arrow gives rise to profile plots showing (d) just the sinusoidal {111} fringes for the Scherzer image, and (e) a significant sharpening of the intensity at the dumbbell pairs for the underfocused image showing higher resolution information being transferred.