Journal of Electron Microscopy **53(3)**: 257–266 (2004)

Full-length paper

HAADF–STEM imaging with sub-angstrom probes: a full Bloch wave analysis

Yiping Peng[1,2], Peter D. Nellist[3] and Stephen J. Pennycook[1,*]

[1]Condensed Matter Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6030, USA, [2]Nanoscale Physics and Devices Laboratory, Institute of Physics, CAS, Beijing 100080, China and [3]Nion Co., 1102 8th St, Kirkland, WA 98033, USA

[*]To whom correspondence should be addressed. E-mail: pennycooksj@ornl.gov

Abstract A full coherent Bloch wave calculation is presented to investigate high-angle annular dark-field image formation for sub-angstrom probes in scanning transmission electron microscopy (STEM). With increasing illumination angle, the contribution of the 1s bound state increases to a maximum at an optimum probe angle, after which we find increasing contributions from high-angle plane wave states around the periphery of the objective aperture. Examination of image contributions from different depths within a crystal shows an oscillatory behavior due to the beating between 1s and non-1s states. The oscillation period reduces with decreasing probe size, while the relative contribution from a specific depth increases. This signifies a changeover from a projection mode of imaging to a depth-slicing mode of imaging. This new mode appears capable of resolving three-dimensional atomic structures in future generation aberration-corrected STEM.

Keywords scanning transmission electron microscopy, coherent scattering, elastic scattering, incoherent imaging, Z-contrast, depth slicing

Received 23 October 2003, accepted 1 April 2004

Introduction

Scanning transmission electron microscopy (STEM) has become a powerful tool for quantitative image analysis at atomic resolution. Especially, with the introduction of high-angle annular dark-field (HAADF) imaging, incoherent Z (atomic number)-contrast STEM has been shown to provide direct, compositionally sensitive structure images without the phase problem [1,2]. With this advantage over conventional coherent high-resolution transmission electron microscopy (HRTEM), HAADF-STEM has enabled imaging not only of individual atomic columns in crystals [1], but single dopant atoms on the surface [3] and within the bulk [4,5]. Historically, there has been much confusion regarding the origin of incoherence in annular dark-field STEM imaging. The original claim for incoherent imaging in STEM [2] assumed that the high-angle scattering would be dominated by incoherent thermal diffuse scattering (TDS), thus, rendering the image incoherent. Later analysis of image formation for thin specimens [6] showed that phonon scattering was not a prerequisite to form an incoherent image. More recently, Nellist and Pennycook [7] used a Bloch wave approach to demonstrate that incoherent imaging with dynamically scattered coherent electrons was achieved through a high-pass filtering effect provided by the annular dark-field detector geometry, with the result that the image contrast is dominated by the 1s-type bound Bloch states.

Recently, of particular interest is the pursuit of sub-angstrom resolution by use of a spherical aberration corrector (for example, [8,9]). Apart from the spectacular increase in available resolution and image contrast, sub-angstrom probes require much larger probe-forming apertures. In an uncorrected microscope, the probe-forming aperture is typically much smaller than the angular width of the 1s Bloch state, but with sub-angstrom beams this is no longer the case, especially with anticipated future generations of aberration correctors. We might expect, therefore, the dominance of the 1s Bloch state to disappear. The purpose of this paper is to examine this behavior and determine the primary contribution to image contrast in future generation STEM.

Our aim is not to perform a detailed comparison between simulated and experimental images. Many other studies have already been undertaken towards this goal [10–18]. Our intent is to analyze the contributions to the image from the various excited Bloch states as a function of depth to determine the image contrast mechanism. The strength of the Bloch wave approach is the insight it provides into the physics of STEM image formation. Bloch waves in high-energy electron diffraction can be described well as 'atomic' or 'molecular' orbitals denoted by quantum numbers 1s, 2s, 2p, etc. [19], and have provided a powerful method for quantitative understanding of imaging and diffraction in high-energy electron microscopy. We go beyond the analysis presented in [7] by including all cross terms between different Bloch states, which allows us to investigate the image contributions from any Bloch state or combination of states as a function of depth within a crystal.

It is well known that the intensity falling on the HAADF detector is predominantly TDS and it is also widely accepted that this is effectively generated incoherently. However, we do not want to make this assumption at the outset, since that will guarantee an incoherent image. Rather, we use a coherent formulation to show how this incoherence arises. A coherent imaging formulation ensures that the coherent bright-field and HAADF images are treated on the same theoretical basis and, therefore, highlights the origins of the very different nature of the two images. We stress again that our intent is not an accurate simulation, but a detailed understanding of the reasons for incoherent imaging and, in particular, the role of the 1s Bloch state with increasing probe-forming aperture. The coherent formulation does not assume incoherent imaging will apply simply because phonons are involved. These results can be viewed as a transition from the coherent phase object approach valid for thin crystals to the incoherent scattering assumption that is a good approximation for thick crystals. We examine the phase object contributions from thin slices at specific depths within a crystal, but we do not integrate those contributions into the final image from the whole crystal. In this way, we see the important contributing Bloch states at any depth.

Methods: coherent Bloch wave formulation

Scattering to the HAADF detector occurs primarily by elastic scattering, as is assumed in a frozen phonon calculation, for example. This is because typical phonon momenta are much smaller than typical Bragg momenta and the phonon wave vector, therefore, represents a minor perturbation to the final scattered wave vector. Transverse phonon momentum components ensure that the pattern on the detector from each successive electron is shifted slightly and the end result is a diffuse pattern. It is well known that the annular dark-field detector destroys transverse coherence very effectively, just because it averages over many elastic scattering discs [6,7]. Vertical phonon momentum components can couple the zero layer reflections to the Ewald sphere. This destroys the coherence in the beam direction quite effectively and the thickness dependence of the image changes from the oscillatory dependence of perfectly coherent scattering to a dependence that is monotonically increasing, with the final image being the incoherent sum of partial images from all slices. We stress, however, that the contribution to the image from any one slice within the crystal is dominated by the elastic scattering to the detector. Phonons only ensure that the contributions from different slices add incoherently. Thus, a coherent Bloch state formulation is able to determine the important Bloch states contributing to the HAADF image from any depth within a crystal.

Recent studies of probe propagation have indicated a reduced depth of field for sub-angstrom probes [20,21]. In this paper, we wish to examine the role of the 1s state with increasing probe aperture angle and the contribution of the 1s state to slices at different depths in the crystal. Our results represent 'partial images' from specific Bloch states or combinations of states as a function of depth of the crystal slice within the crystal. The results show that with increasing probe angle the 1s state contribution first peaks and eventually fades. At the same time, the contribution from a specific depth increases. This represents a changeover from the dominance of channeling behavior, and projection imaging, to the dominance of non-channeling behavior, or depth slicing.

In the framework of the Bloch wave method developed by Nellist and Pennycook [7], the coherent interference between excited Bloch states can be split into two components: depth-independent $j = k$ diagonal terms (named 'self terms') and depth-dependent $j \neq k$ off-diagonal terms (called 'cross terms'). The latter have been shown to decay to zero at the limit of large depth [7], leaving the so-called residual object function as being the sum over the self terms only. In the present paper, we include all self and cross terms, but as a simplification we do not include any phenomenological absorption, although a thermally smeared potential is used. The central idea in setting up a Bloch wave method for STEM lies in that the incident electron wave is expressed in terms of the eigenstates of the system, which should be stationary Bloch states in the case of a periodic crystal. For mathematical convenience, the electron probe can be described as a coherent superposition of plane waves and an aberration-free probe is assumed throughout. To put emphasis on the primary objective of our study, we neglect the higher-order Laue zone reflections, which have only a small contribution to the annular dark-field image [12,22].

Following the deductive derivations in [7] and the notations therein, the Fourier transform of the real-space STEM image intensity $I(R_0, z)$ (probe position R_0 and specimen depth z) for a zone-axis-aligned crystal can be given by

$$\tilde{I}(Q, z) = \int A(K_\mathrm{i})A^*(K_\mathrm{i} + Q) \sum_g D(K_\mathrm{i} + g)F(g, K_\mathrm{i}, Q, z)\mathrm{d}K_\mathrm{i} \qquad (1)$$

Here, Q is the image spatial frequency vector in the same units as the reciprocal lattice vector $\boldsymbol{g}$; $A(K_i)$, known as the aperture function, is the complex amplitude of the incident partial plane wave with transverse wave vector K_i and in STEM usually has a magnitude given by a circular top-hat function with radius of the objective aperture; $D(K_i + \boldsymbol{g})$ is the detector function taking the value of unity if the wave vector of a Bragg beam, $K_i + \boldsymbol{g}$, falls on the detector range and zero otherwise; and the term $F(\boldsymbol{g}, K_i, Q, z)$ represents the sum of coherent interference in the following form

$$F(\boldsymbol{g}, K_i, Q, z) = \sum_{j,k} \phi_0^{(j)*}(K_i)\phi_g^{(j)}(K_i)\phi_Q^{(k)}(K_i)\phi_g^{(k)*}(K_i) \times \exp\{-i2\pi z[k_z^{(j)}(K_i) - k_z^{(k)}(K_i)]\} \quad (2)$$

where $\phi_g^{(j)}(K_i)$ is the $\boldsymbol{g}$-th Fourier component of the j-th K_i-excited Bloch wave, whose eigenvalue gives the longitudinal zone-axis wave vector component, $k_z^{(j)}(K_i)$. The rapid variation of the high-energy electron wave in the z direction denoted by an overall phase factor, $\exp(-i2\pi k_0 z)$, where k_0 is the incident electron wave vector, is omitted in this description ([23], which, however, adopted a different convention for free-space plane waves, i.e. $\exp(ik_0 z)$).

Equation (1) is a completely reciprocal-space formulation. By going back to the primitive form of eq. (2), one can clearly see the physics of the image formation process. First, we write down two important property equations for Bloch waves [23]

$$\phi_g^{(j)}(K_i + \boldsymbol{g}') = \phi_{g+g'}^{(j)}(K_i) \quad (3)$$

$$k_z^{(j)}(K_i + \boldsymbol{g}) = k_z^{(j)}(K_i) \quad (4)$$

which arise because Bloch functions satisfy $\psi^{(j)}(K_i, r) = \psi^{(j)}(K_i + \boldsymbol{g}, r)$ (to within a phase factor, which is also the case for eq. (3)). The interference term described by eq. (2) was derived utilizing eqs (3) and (4) from

$$F(\boldsymbol{g}, K_i, Q, z) = \sum_{j,k} \phi_0^{(j)*}(K_i)\phi_g^{(j)}(K_i)\phi_0^{(k)}(K_i + Q)\phi_{g-Q}^{(k)*}(K_i + Q) \times \exp\{-i2\pi z[k_z^{(j)}(K_i) - k_z^{(k)}(K_i + Q)]\} \quad (5)$$

where $\phi_0^{(j)*}(K_i)$, due to the orthogonality between the Bloch states, is the amplitude of excitation of the j-th Bloch wave for plane wave K_i illumination. Equation (5) represents the interference between two Q-related incident partial plane waves, K_i and $K_i + Q$. Taking into account the integration over K_i in eq. (1), one can see that the image contrast at a spatial frequency of Q arises from the interference between all pairs of incident partial plane waves whose transverse wave vectors are separated from each other by Q. This is also why the objective aperture diameter should not be less than the distance between neighboring reciprocal $\boldsymbol{g}$ points in order to obtain image contrast. Each of two such partial plane waves, due to the scattering by the specimen, excites a set of depth-dependent Bloch states (represented in eq. (5) by the summation over j and k, respectively), which, by definition, can be expanded into a series of Bragg beams in the reciprocal $\boldsymbol{g}$ space. As eq. (5) shows, interference occurs between any two scattered Bragg beams that have the same final wave vector (i.e. $K_i + \boldsymbol{g}$) and many such interfering Bragg beams are summed over by the detector to give the final image contrast. For a HAADF detector, the summation of a large number of coherently interfering Bragg beams gives rise to incoherent imaging.

Lattice imaging in STEM has been understood as the interference between overlapping coherent convergent beam discs through optical diffraction analysis [24]. Since the Bragg discs centered on $\boldsymbol{g}$ vectors in the detector plane are images of the probe-forming illumination cone, the concept of overlapping discs is the same as interfering Bragg beams described here. For $Q = 0$, the interference term described by eq. (5) would result from the identical incident plane wave, K_i, thus corresponding to the non-overlapping or self-interfering contribution. On the other hand, since the zero frequency Fourier component of $I(R_0, z)$ is independent of the probe position, R_0, the value of eq. (1) at zero spatial frequency contributes a uniform background to the image. That is, the intensity contributed by the non-overlapping interference is insensitive to the probe position and image contrast comes only from the overlapping interference (i.e. the non-zero frequency Fourier components in eq. (1)) falling on the detector, which is in good agreement with Spence and Cowley's description [24].

An image function (IF) is the entire evaluation of eq. (1). We call this a full image function (FIF) as being the sum over all the $j = k$ self terms and $j \neq k$ cross terms, and a so-called residual image function (RIF) is defined as including the summation over $j = k$ only. From eq. (2) it can be seen that the RIF is independent of depth, whereas the cross terms are depth-dependent. The FIF, therefore, represents the image intensity contribution from an elementary slice of crystal at depth z. It is obvious that at zero thickness the intensity detected by a HAADF detector would be zero, which can also be verified from eq. (2) by considering the orthonormality of the Bloch states. For $z = 0$, the right side of eq. (2) reduces to $\delta_{0g}\delta_{Qg}$ ensuring $\boldsymbol{g} = 0$, and $D(K_i + \boldsymbol{g})$ in eq. (1) would remain zero for all values of K_i since a HAADF detector inner angle is larger than the probe-forming aperture angle, thus, leading to zero detected intensity. At the limit of large depth, all the cross terms would oscillate infinitely quickly in the integral over K_i, due to the differences between the j and k dispersions in the argument of the exponential term. Since the dispersion surfaces will eventually cancel out the cross terms and leave the self terms only, the RIF turns out to be the large depth limit of the FIF.

Calculation and discussion

Throughout this article, all the Bloch wave calculations were performed at 300 kV for GaAs (lattice constant $a = 5.656$ Å at room temperature), which has classic 'dumbbell' pairs of Ga and As columns in its <110> projection. To take into account

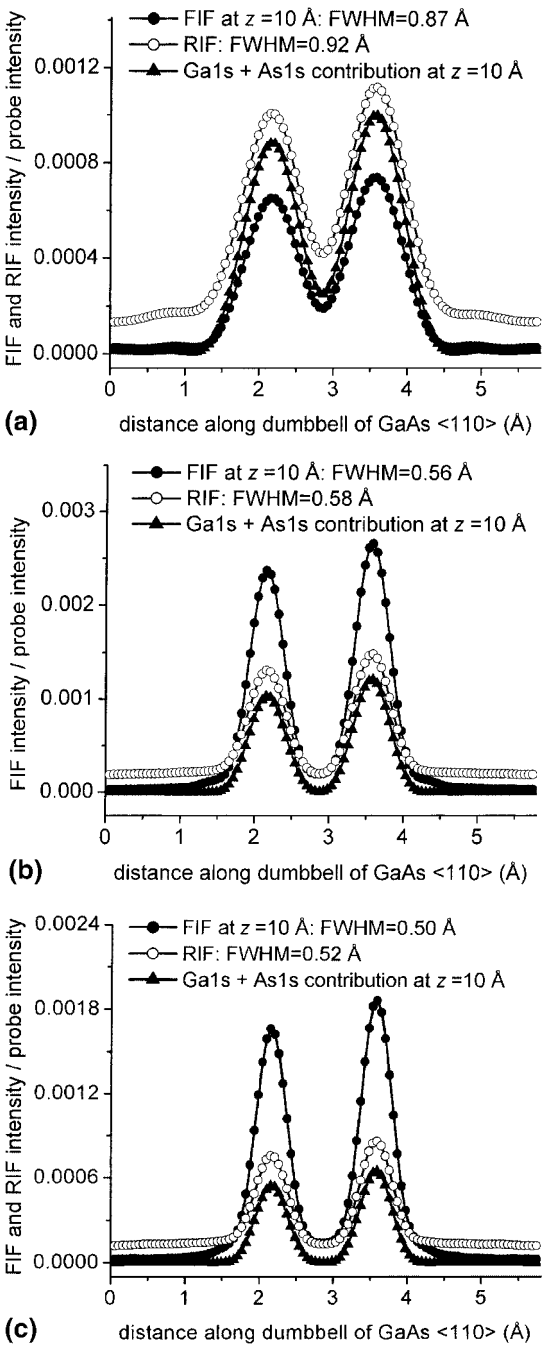

Fig. 1 Profile plots of the RIF and FIF at 10 Å depth across the dumbbell pair of <110> GaAs compared with the contribution of As 1s and Ga 1s only, calculated with 849 reflections, 300 kV acceleration voltage and no absorption. The objective aperture semi-angle, detector angle range and number of incident beams are, respectively: (a) 12.2 mrad, 49–96 mrad, 219; (b) 20.9 mrad, 45–96 mrad, 637; and (c) 24.4 mrad, 49–96 mrad, 871.

the thermal vibration effect on the atomic potential at finite temperature but avoid possible artifacts induced by the choice of the somewhat elusive Debye–Waller factors, we set the Debye–Waller factors as 0.5 Å^2 for both Ga (Z = 31) and As (Z = 33) atoms. The Bloch wave calculations were carried out for all sampled incident partial plane waves within the illumination cone as a preliminary step for the integral over K_i in eq. (1). The probe size is sometimes defined as the full width at half maximum (FWHM) of its intensity distribution and, in the absence of lens aberrations, can be estimated by

$$\mathrm{FWHM} = 0.515\frac{\lambda}{\theta} \qquad (6)$$

where λ is the incident electron wavelength and θ the objective aperture semi-angle, and the coefficient is derived from the Airy function for the probe.

1s dominance and incoherent imaging

We start by presenting simulated results for three probes with illumination semi-angles of 12.2, 20.9 and 24.4 mrad, respectively corresponding to probe intensity FWHM of 0.83, 0.49 and 0.42 Å at 300 kV acceleration voltage. These probes are aberration-free with defocus set to zero, i.e. focused on the crystal entrance surface. Figure 1 compares FIF intensity profiles for a slice near the entrance surface of the crystal (10 Å depth) with the Ga 1s and As 1s contributions at 10 Å and with the RIF representing the limit of large depth. In all cases, clear peaks appear at the positions of the two dumbbell columns with 1.4 Å spacing and the relative intensities show the expected contrast, with the As column having 11–19% more intensity (close to the Rutherford scattering value of $(33/31)^2$ = 1.13). By comparing the 15 contributions with the RIF, the 1s dominance is obvious and the non-1s states only contribute an almost uniform background, raising the entire profile. This background can be understood as a weak contribution from the non-1s states, which have dispersed at greater depths. This dispersion also leads to some probe spreading through the thickness of the crystal, giving a slightly broadened peak width to the RIF. Nevertheless, the dumbbell remains clearly resolved with a distinctive signal/background ratio of 6.0–7.0.

With the smaller probe angle of 12.2 mrad (Fig. 1a), the image contribution at 10 Å is reduced compared with the 1s contributions alone. With the aperture angle increased to 20.9 mrad, this behavior is reversed (Fig. 1b), with the 10 Å slice now providing a higher image contribution than the 1s states. The increased peak intensity from the thin region of crystal represents the increased contribution of all non-1s states, which have not dispersed yet. In the thin region the *entire incident probe* is contributing to the image, whereas in the thick region only the 1s state contributions remain in the form of the dumbbell. As the probe is further sharpened, this behavior is further enhanced (Fig. 1c). Of course, the peak widths are also reduced with the wider objective apertures, reflecting the improved resolution of the sharper probes.

This behavior suggests that the 1s state remains the only single dominant Bloch state contributing to the HAADF image, even for a sub-angstrom probe, but there is an increasing importance from the sum of all other Bloch states. We next show that these are, in fact, plane wave states located around the periphery of the aperture and determine the optimum aperture for maximum 1s state excitation.

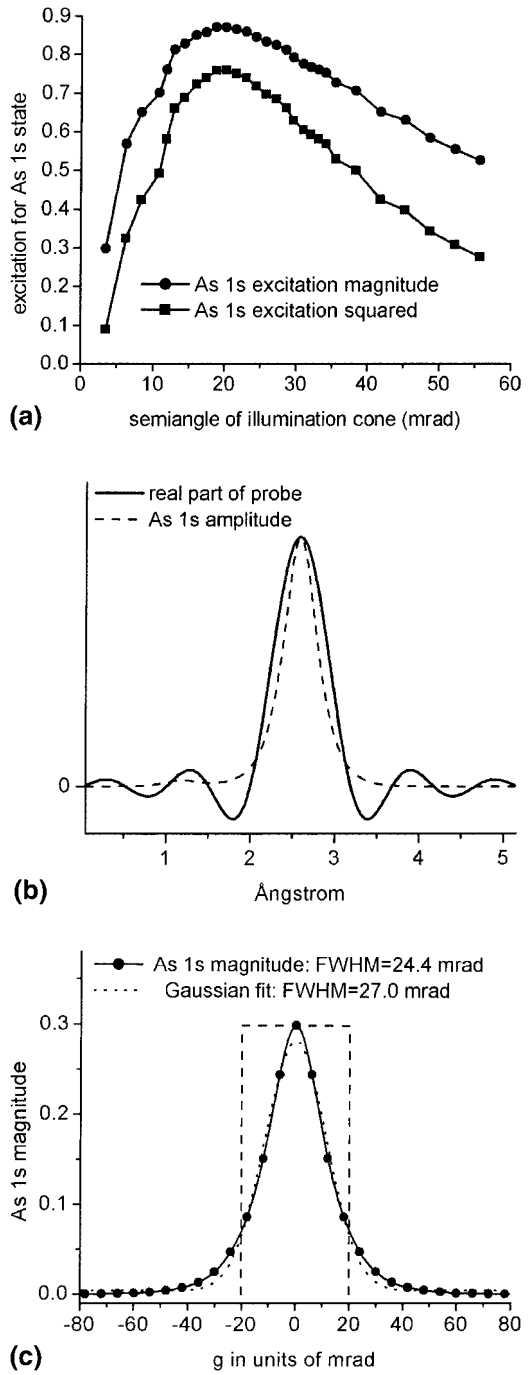

Fig. 2 (a) Angular dependence of the coupling between the incident probe located over the As column and the As 1s state. The optimum probe-forming aperture for maximum 1s state excitation is ~20.2 mrad. Calculations are for K_i= (0,0,0) including all Q-related points in the incident cone. For other K_i's in the first Brillouin zone, very similar plots would be obtained. (b) Real part of the 300 kV aberration-free probe profile (solid line) formed with an aperture semi-angle of 20.2 mrad compared with the As 1s state amplitude (dashed line). For a probe without aberrations its imaginary part is zero. (c) Same as (b), but in reciprocal space. The dotted line is the Gaussian fit to the As 1s state magnitude and the dashed square denotes the optimal coupling of objective aperture with the As 1s state.

Excitation of 1s state by the probe due to the similarity between them

An aberration-corrected probe may have a very similar form to the 1s states. In general, if a particle takes an eigenstate of a filter device, then it will be directly transmitted. Hence, if the probe enters the crystal over an atomic column matching the 1s states, it will travel down the column with a fraction reaching the annular dark-field detector; this is the so-called channeling effect. If the probe is located between the atomic columns, its coupling with the 1s state is now weak and the probe will decompose into a large number of less-localized Bloch states that do not scatter efficiently into the detector. Furthermore, if the illumination angle is increased, the probe will become narrower than a 1s state and the coupling on the column will become less efficient. Thus, there should exist an optimum illumination angle for which the probe fits best with the 1s states of the specimen. Beyond this optimum angle, the contrast in the image, and possibly the image resolution itself, might even decrease.

A real-space excitation is first introduced for each Bloch state. We replace the integral in eq. (5) of [7] by a sum over all the Q-related K_i's within the objective aperture for a single initial K_i^{1BZ} in the first Brillouin zone. By utilizing the symmetry properties described by eqs (3) and (4), the resultant excitation coefficient for the j-th Bloch wave branch with the probe located over the respective column is then a function of K_i^{1BZ}

$$\alpha^{(j)}(K_i^{1BZ}) = C_N \sum_Q A(K_i^{1BZ} + Q)\phi_0^{(j)*}(K_i^{1BZ} + Q)$$
$$= C_N \sum_Q \phi_Q^{(j)*}(K_i^{1BZ}) \qquad (7)$$

The Q summation is performed over the entire objective aperture disc and the normalization coefficient C_N equals the reciprocal of the number of eligible Q-values.

The illumination angle evolution of the overlap integral between the probe and the 1s state is then shown in Fig. 2a for the As 1s state with the probe located over the As column. A maximum normalized excitation of 0.87 for the As 1s state appears with a ~20.2 mrad aperture. At this optimum angle, the probe size is 0.50 Å according to eq. (6), which is significantly larger than the FWHM for the As 1s (0.31 Å). Interestingly, van Aert *et al.* [25] reached a similar conclusion based on an optimal signal/noise criterion. Intuitively, one would anticipate the optimum probe to have a similar FWHM to the 1s state. The apparent discrepancy of ~0.2 Å originates from the different forms of the probe and the 1s state: the 1s state bears a Gaussian-like shape, whereas the probe-forming aperture has a top-hat form in reciprocal space, leading to a real-space probe described by an Airy function. This is illustrated for the optimal coupling in real (Fig. 2b) and reciprocal space (Fig. 2c).

It is interesting to note that although the As 1s state appears close to Gaussian in form and takes a slightly smaller FWHM than the latter (Fig. 2c), the 1s state contains slightly higher values at larger angles. In real space, therefore, the 1s state is narrower than the Fourier transform of the Gaussian, with

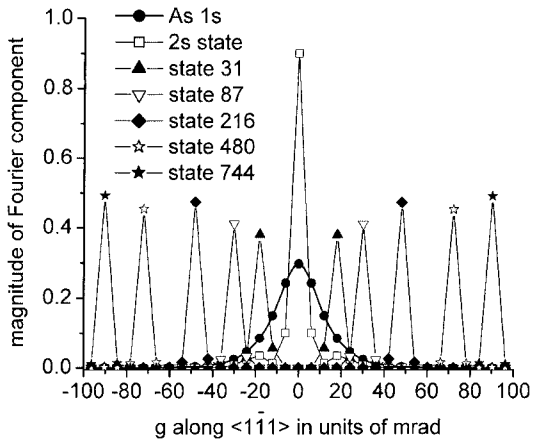

Fig. 3 Distribution of the Fourier components $\phi_g^{(j)}(K_i)$ along $<1\bar{1}1>$ for various Bloch waves in <110> GaAs, plotted for K_i= (0,0,0) with 849 reflections.

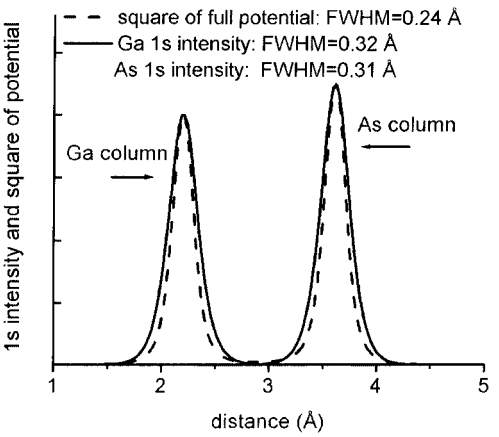

Fig. 4 Profile plots of the intensity of the As and Ga 1s Bloch states (solid) compared with the square of projected full potential (dashed) for GaAs <110>, scaled to the same peak values.

FWHM of 0.47 and 0.63 Å, respectively. The difference would be further enhanced if one simply adopts a Gaussian with identical reciprocal-space FWHM to replace the 1s state and, thus, care should be taken with such an approximation.

The strong excitation of the 1s state by a STEM probe can also be demonstrated by comparison with the non-1s Bloch states, as shown in Fig. 3 for K_i = (0,0,0). In view of eq. (3), one can easily see $\phi_0^{(j)*}(K_i = Q) = \phi_Q^{(j)*}(K_i = 0)$, which says that the excitation amplitude of a Bloch state for non-normal incidence is just the corresponding Bloch wave component for normal incidence. That is, Fig. 3 can also be regarded as an excitation plot as a function of incident K_i for different Bloch states. Moreover, eq. (7) has stated that the real-space excitation before normalization can be expressed by the summation of reciprocal excitations for all Q-related K_i values over the objective aperture. It is clear that the 1s state is the broadest in reciprocal space, while the non-1s states, especially those of higher order, have a very localized distribution, closely resembling plane waves. Although the reciprocal space excitation $\phi_0^{(1s)*}(K_i)$ is maximum for normal incident K_i, the optimum illumination angle for maximum 1s real-space excitation occurs with a ~20.2 mrad aperture as a result of competition between the 1s and all the non-1s states excited by the probe. Beyond this angle, the excitations for more and more higher-order states would be gradually switched on at the cost of reduced normalized excitations for 1s and lower-order states.

Simulation evidence for the convolution description

A perfectly incoherent image is given by a convolution between an object function representing the scattering power of the specimen and a resolution function. It is, thus, appealing to test the validity of the simple incoherent description for the available data from Bloch wave calculations. We would expect the small-thickness image slice to approximate a convolution of the probe with the detector-filtered high-angle scattering potential, whereas the high-thickness image slice would approximate a convolution of the probe with the 1s state. In our case, the FWHM of the intensity profile for the As 1s state is 0.31 Å, which is slightly broader than the square of the projected potential (0.24 Å), as shown in Fig. 4. The FWHM of the convolutions and calculated FIF contributions (as showed in Fig. 1) are compared in Table 1 for the As column. The error in the FWHM is only 0.04 Å for the 10 Å slice and 0.03 Å for the 1000 Å slice. Also parenthesized in the last row of Table 1 is the FWHM value for the RIF, which is equal to the FIF at the limit of infinite depth, providing good agreement again with the 1s convolution results.

For very thin samples of tens of angstroms, the resolution would basically rest upon the probe size since the projected

Table 1. Test of the convolution description in terms of FWHM of the As column

Illumination semi-angle	12.2 mrad	20.9 mrad	24.4 mrad
Probe intensity size	0.83 Å	0.49 Å	0.42 Å
Detector inner radius	49 mrad	45 mrad	49mrad
Square of detector-filtered projected object potential	0.12 Å	0.14 Å	0.12 Å
(High-angle potential)^2 ⊗ probe[a]	0.86 Å	0.53 Å	0.46 Å
FWHM of FIF at 10 Å	0.87 Å	0.56 Å	0.50 Å
FWHM of As 1s intensity	0.31 Å	0.31 Å	0.31 Å
As 1s intensity ⊗ probe[a]	0.89 Å	0.59 Å	0.53 Å
FWHM of FIF at 1000 Å	0.90 Å (0.92 Å)	0.60 Å (0.58 Å)	0.56 Å (0.52 Å)

[a]Symbol ⊗ denotes the convolution operation.

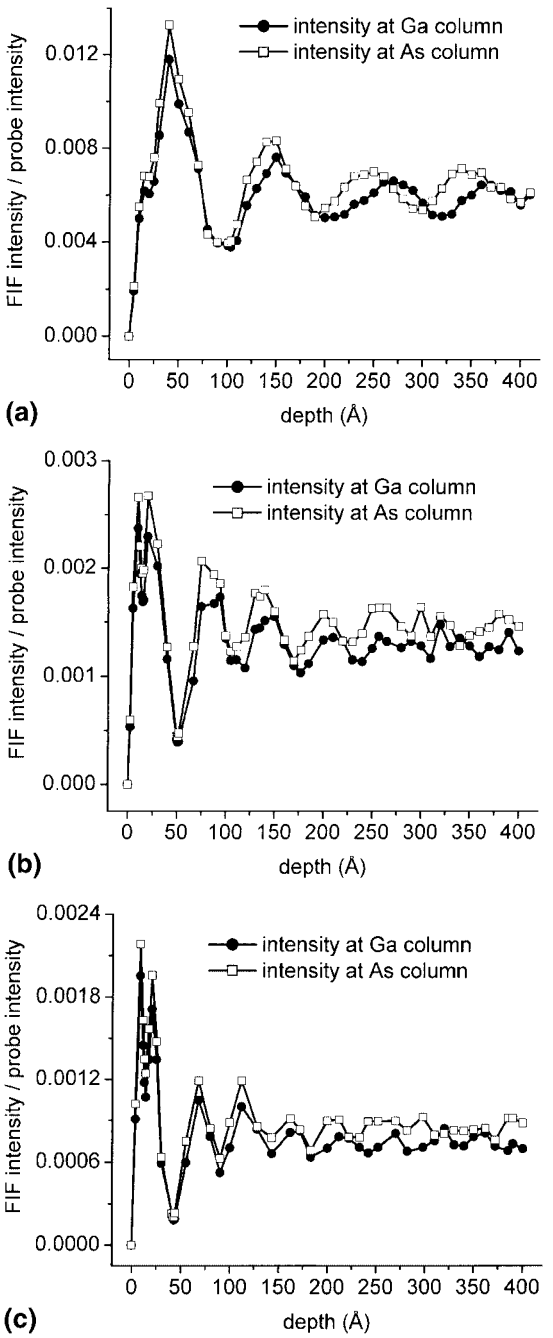

Fig. 5 The FIF intensity contributions as a function of depth within a GaAs <110> crystal for a 300 kV probe located over a Ga column and an As column, calculated without absorption. Objective aperture, detector angle range, number of reflections and incident beams are, respectively: (a) 12.2 mrad, 35–66 mrad, 411, 219; (b) 20.9 mrad, 45–96 mrad, 849, 637; and (c) 24.4 mrad, 49–96 mrad, 849, 871.

potential filtered by HAADF detectors is usually much sharper. In the case of thicker specimens, the width of the image again appears to be in accord with the convolution description, now involving the 1s state. However, questions arise with even sharper probes, because of the reduced contribution of the 1s Bloch state compared with all other states. Above a certain 'critical' probe angle, the non-1s states and the high-angle incident K_i's give more and more significant contribution to the non-local background of the image, and may finally counteract the 1s dominance. We, therefore, examine the depth dependence in more detail for selected aperture angles.

Depth evolution of FIF intensity from Bloch wave calculation

Since the specimen depth z is encoded in the exponential term of eq. (2), the FIF intensity contribution can be calculated for a slice at any desired depth by including the off-diagonal terms within the Bloch wave model. Figure 5 gives an evolution of the depth dependence of the FIF intensity for GaAs <110> with the three illumination angles used earlier, i.e. (a) conventional 12.2 mrad, (b) 20.9 mrad close to the optimal aperture and (c) smaller probe of 24.4 mrad. An obvious feature is the oscillatory intensity resulting from the phase term of eq. (2). The oscillation period or effective extinction distance remains approximately constant for the first few periods, in agreement with the channeling model. However, the extinction distance covers a range of ~40–100 Å and turns out to be probe size-related. That is, the oscillation period becomes compressed with increased illumination angle; the sharper the probe, the smaller the extinction distance. This can be attributed to the fact that higher K_i's for increased illumination angle would introduce more non-1s components and, as a result, the extinction distance shrinks due to the increased separation between wave vectors for the 1s and the average of the high-order Bloch states. In the next subsection, a simple model is proposed for estimating the effective extinction distance.

In the thinnest part of the crystal, the 1s dominance has yet to establish itself, so the highest peak arising in the first period represents the contribution from all Bloch states. At increased depth, one can see from Fig. 5 that the oscillations tend to become non-periodic and attenuate in amplitude. This indicates the disorder and decay in the cross terms as a function of depth. It is worth mentioning that the first major valley appearing around 101, 50 and 42 Å in Figs 5a–c, respectively, corresponds to diffusive images for As and Ga columns. We show in the next subsection that this represents probe propagation in the thin specimen. When the objective aperture size is sufficiently large, this diffusive image can take a doughnut shape, which can be double the size of the usual image spot.

A small dip arises at 15 Å in Fig. 5b and 14 Å in Fig. 5c, and it is incipient even in Fig. 5a for the 0.83 Å probe. The physical origin can be understood as the kinematical z-coherence effect discussed in [6]. Based on the scattering geometry shown in Fig. 12b of [6], destructive interference would occur at a depth of $z = 2\lambda / \theta_i^2$, where λ is the electron wavelength and θ_i the detector inner angle. For θ_i = 45 and 49 mrad, as in Figs 5b and 5c, respectively, the above formula would give a dip at a respective column length of around 19 and 16 Å, providing a good account for the corresponding results from the coherent Bloch wave calculations. In addition, such sharp dips do not occur in the subsequent oscillation periods, which can be attributed to the fact that the dynamical scattering dephases

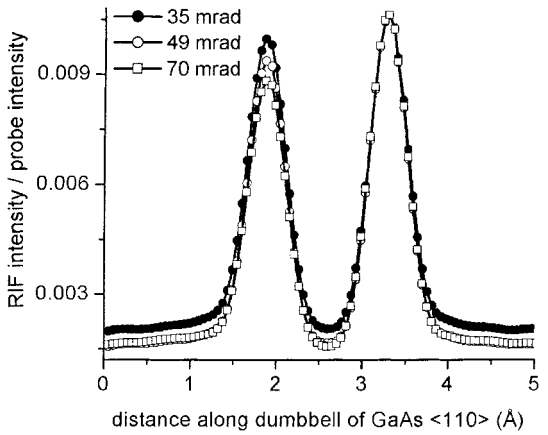

Fig. 6 Profiles of the RIF intensity for detector inner radius of 35 mrad, 49 mrad and 70 mrad scaled to the same peak value at the As column. Other calculation parameters are the same as in Fig. 1c.

the Bloch states and dominates over the kinematical effect at larger depth.

Finally, it should be pointed out that, although the curves in Fig. 5 are not an accurate representation of the physical image intensity that would be recorded in the presence of absorption and TDS, they do illustrate how interference between all the excited Bloch states leads to an extinction distance in the order of ~40–100 Å. As for the TDS effect, if we assume that all the high-angle beams reaching the detector suffer diffuse scattering, the total detected signal would be proportional to the integral of eq. (1) over the entire specimen thickness, providing an alternative way to account for the TDS within the coherent Bloch wave model.

Estimating the depth oscillations

Since the form of Fig. 5 also suggests that the non-1s part of the wave function propagates through the crystal largely in the form of plane waves, the depth oscillation of FIF intensity can be understood as the beating between the 1s Bloch state and the non-1s part of the probe as they propagate in the specimen. We write the wave function in terms of 1s and non-1s components

$$\psi(R, z) = \psi_{\text{non-1s}}(R, z) + \psi_{1s}(R) \times \exp(-i2\pi k_z^{[1s]} z) \quad (8)$$

Applying eq. (8) at zero depth, we can write the non-1s component in the equation as

$$\psi_{\text{non-1s}}(R, z = 0) = P(R) - \psi_{1s}(R) \quad (9)$$

where $P(R)$ is the probe wave-function. The non-1s component of the wave function will consist mostly of plane wave components at angles close to the aperture edge, as implied by eq. (9) and Fig. 2c. We can assume, therefore, that these plane waves have a wave vector of

$$k_z^{\text{plane}} = \cos\theta \cdot {}^{1}\!/\!_{\lambda} \quad (10)$$

Table 2. Estimated depth for the first valley compared with calculation

Illumination semi-angle	12.2 mrad	20.9 mrad	24.4 mrad
As column: calculation	101 Å	50 Å	42 Å
As column: estimation	103 Å	59 Å	48 Å
Ga column: calculation	104 Å	51 Å	43 Å
Ga column: estimation	112 Å	62 Å	50 Å

where θ is the semi-angle of the probe-forming aperture.

The depth plot will be again a minimum when the plane waves and the 1s state interfere at the extinction distance, i.e.

$$k_z^{[1s]} z - \cos\theta \cdot {}^{z}\!/\!_{\lambda} = 1 \quad (11)$$

where $k_z^{[1s]}$ can be calculated using the eigenvalue of the 1s state that comes out of the Bloch wave calculations, as shown in the following. Since the eigenvalues from the Bloch calculations are –23.8689 Å^{-2} for the As 1s state and –20.7850 Å^{-2} for the Ga 1s state, the corresponding wave vectors $k_z^{(1s)}$ are, respectively, 0.0060 and 0.0052 Å^{-1} according to (compare eq. (11) of [23])

$$k_z^{(1s)} = -\frac{\text{Eigenvalue}}{2 \cdot ({}^{2\pi}\!/\!_{\lambda})} \frac{1}{2\pi} \quad (12)$$

The total wave vector along the z direction for the 1s state can be obtained by putting back the overall phase factor $\exp(-i2\pi k_0 z)$ omitted in eq. (2)

$$k_z^{[1s]} = k_0 + k_z^{(1s)} \quad (13)$$

The wavelength for incident electrons at 300 kV is $\lambda = 0.0197$ Å and the corresponding wave vector is $k_0 = 1 / \lambda = 50.7614$ Å^{-1}. Hence, we have $k_z^{[1s]} = 50.7674$ Å^{-1} for the As 1s and 50.7666 Å^{-1} for the Ga 1s. Substituting them into eq. (11) gives rise to the depth corresponding to the first minima for different probe semi-angles, providing a good fit to the values from the Bloch wave calculations, as shown in Table 2.

In addition to the good agreement between the two different approaches, it can be noticed also that the effective extinction distance for Ga is slightly larger than that for As. This is because the As 1s state is lower in energy than the Ga 1s state, due to its deeper potential well, thus, leading to a larger $k_z^{[1s]}$ and smaller extinction distance for heavier atoms according to eqs (13) and (11). Similar results [26] were reported within a simple channeling description testified by multi-slice calculations.

Detector inner angle range

Before concluding this section, some test results are presented concerning the minimum detector inner radius. Generally, a factor of at least 3 for the ratio of detector inner radius to the illumination convergence semi-angle has been regarded as necessary for a HAADF-STEM image to be well described as incoherent [27]. Here, we temporarily neglect the cross terms

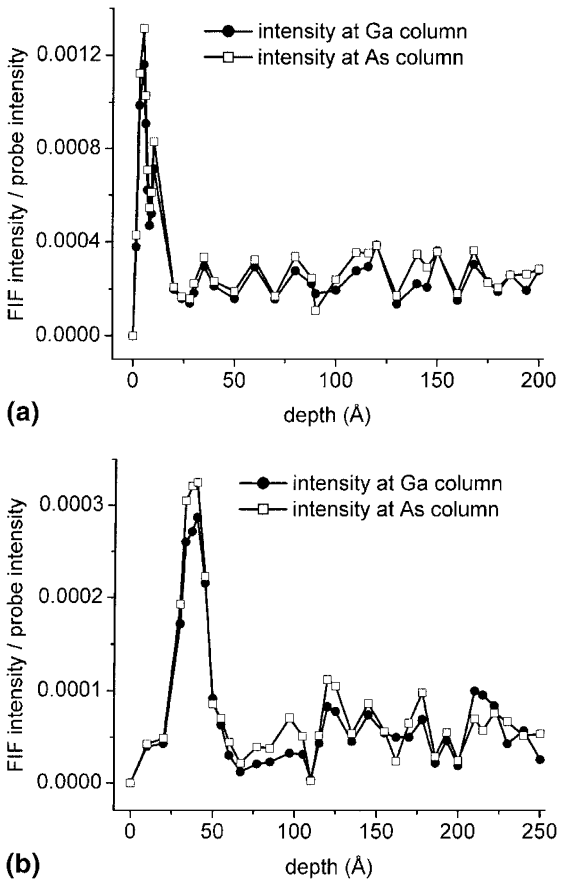

Fig. 7 Similar to Fig. 5, but for a probe semi-angle of 34.8 mrad (a) without defocus and (b) with defocus = 50 Å, calculated using 1405 reflections, detector angles of 70–128 mrad, 117 incident beams and no absorption.

to focus on the basic image features in the case of sub-angstrom probes. Figure 6 shows calculated RIFs for a probe-forming semi-angle of 24.4 mrad. It is seen that the forms of the RIFs remain almost unchanged with the inner radius above a ratio of 2. This is reasonable, since the semi-angle of a sub-angstrom probe is already greater than a typical Bragg scattering angle. A detector inner angle of twice the size of the illumination cone is sufficient to guarantee the requirement that the detector should eliminate the unscattered or low-angle scattered electron beams. On the other hand, it can be seen from Fig. 6 that for a given probe, a larger inner angle gives lower background and a contrast that increases towards Z^2. This is achieved at the expense of reduced signal.

Future prospects

We have shown that a HAADF-STEM image can still be a direct image of the 1s states even with sub-angstrom probes, but the non-1s component, which propagates essentially as plane waves, gains in contribution as the objective aperture angle is increased above that for optimal 1s coupling. Therefore, it is instructive to reflect on the nature of the image formed from this component.

Figure 5 shows that with increasing probe aperture there is an increasingly enhanced contribution from slices close to the entrance surface, and the oscillatory behavior at higher thickness slices is reduced both in intensity and in period. From Fig. 1 we know that this behavior comes from the enhanced non-1s component with increased probe-forming aperture. For an objective aperture as large as, for example, ~60 mrad, one can extrapolate from Fig. 5 and eq. (11) that the peak for the first oscillation period would fall within 10 Å, with a peak value much higher than those for the subsequent oscillation periods. This behavior indicates a reduced depth of field and a changeover from a channeling to a depth-slicing mode of imaging.

Therefore, if we defocus a 'superprobe' so that it is brought to focus within the specimen, what the HAADF detector finally sees, due to the much reduced depth of field in the order of angstroms, would be mostly the image from that specific depth. Depth slicing and three-dimensional imaging, thus, may be realized as images are generated at different focal planes with next generation aberration-corrected STEM.

As a justification of the above extrapolation, Fig. 7 compares the depth-dependence of the FIF intensity for an illumination semi-angle of 34.8 mrad with and without defocus. For this small probe of 0.29 Å FWHM, the depth-dependence without defocus (Fig. 7a) has a compressed oscillation period and the first major valley lies at 28 Å, agreeing well with the estimation of eq. (11). Meanwhile, the small dip occurs at 8 Å, in good comparison once again with the z-coherence prediction (8.0 Å). Moreover, the difference between peak values at 5 Å and any other regions is significantly enhanced to reach a ratio of more than 3. With the introduction of a defocus of 50 Å (Fig. 7b), since the probe is expected to have a less effective focusing inside the specimen, it is reasonable that the FIF results have a reduced contrast, resolution and intensity compared with the case without defocus. In spite of that, Fig. 7b still shows a steep peak in the region of 40 Å depth, where the dumbbell structure can be distinctively resolved and the real image, thus, would be dominated by this characteristic depth region. In future generation aberration-corrected microscopes, larger illumination angles will allow depth slicing over a larger range of defocus with better depth resolution.

Concluding remarks

We have used a full Bloch wave model of coherent dynamical diffraction of an incident STEM probe by a crystal to understand the physics of HAADF image formation with sub-angstrom probes. The Bloch wave approach is extended to include the depth-dependent cross terms and it is confirmed that the dominance of the 1s state is maintained with a sub-angstrom probe down to an optimum size, which corresponds to the maximum real-space 1s state excitation.

Beyond the optimum aperture for 1s state excitation, new behavior appears as the 1s dominance is replaced by dominance of the non-1s part of the probe. This part of the probe is

formed predominantly from high-angle plane wave components of the probe. With large probe-forming apertures these components are incident many Bragg angles from the zone axis and, therefore, are relatively unaffected by dynamical diffraction that occurs for the lower-angle components of the probe. They come, therefore, to a focus at a specific depth even in a zone-axis crystal. Interference occurs between this component and the 1s Bloch state giving an extinction distance that depends not only on the specimen, but also on the probe size. With increasing probe-forming apertures, the behavior of the beam becomes increasingly classical as the 1s part of the probe becomes progressively reduced compared with the high-angle components. The image comes increasingly from a specific depth within the specimen.

The next generation superprobe, therefore, presents a very interesting opportunity. By exploiting the plane wave-like, non-1s Bloch wave components, it appears possible to achieve three-dimensional information via depth slicing, even in a zone-axis crystal, and potentially to achieve atomic resolution in three dimensions with future generations of aberration-corrected STEM.

Acknowledgments

The authors should like to thank Dr A. R. Lupini for his generous expertise in programming and S. D. Findlay for useful discussions. Both are gratefully acknowledged for critically reading the manuscript. This research was sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory managed by UT-Battelle, LLC for the US Department of Energy under contract no. DE-AC05-00OR22725, and by appointment to the ORNL Postdoctoral Research Program administered jointly by ORNL and ORISE.

References

1 Pennycook S J and Boatner L A (1988) Chemically sensitive structure imaging with a scanning transmission electron microscope. *Nature* **336**: 565–567.

2 Pennycook S J and Jesson D E (1990) High-resolution incoherent imaging of crystals. *Phys. Rev. Lett*. **64**: 938–941.

3 Nellist P D and Pennycook S J (1996) Direct imaging of the atomic configuration of ultra-dispersed catalysts. *Science* **274**: 413–415.

4 Voyles P M, Muller D A, Grazul J L, Citrin P H, and Gossmann H-J L (2002) Atomic-scale imaging of individual dopant atoms and clusters in highly *n*-type bulk Si. *Nature* **416**: 826–829.

5 Lupini A R and Pennycook S J (2003) Localization in elastic and inelastic scattering. *Ultramicroscopy* **96**: 313–322.

6 Jesson D E and Pennycook S J (1993) Incoherent imaging of thin specimens using coherently scattered electrons. *Proc. Roy. Soc. London A* **441**: 261–281.

7 Nellist P D and Pennycook S J (1999) Incoherent imaging using dynamically scattered coherent electrons. *Ultramicroscopy* **78**: 111–124.

8 Krivanek O L, Nellist P D, Dellby N, Murfitt M F, and Szilagyi Z (2003) Towards sub-0.5 Å electron beams. *Ultramicroscopy* **96**: 229–237.

9 Batson P E, Dellby N, and Krivanek O L (2002) Sub-Ångstrom resolution using aberration corrected electron optics. *Nature* **418**: 617–620.

10 Dinges C, Berger A, and Rose H (1995) Simulation of TEM images considering phonon and electronic excitations. *Ultramicroscopy* **60**: 49–70.

11 Nakamura K, Kakibayashi H, Kanehori K, and Tanaka N (1997) Position dependence of the visibility of a single gold atom in silicon crystals in HAADF-STEM image simulation. *J. Electron Microsc*. **46**: 33–43.

12 Amali A and Rez P (1997) Theory of lattice resolution in high-angle annular dark-field images. *Microsc. Microanal*. **3**: 28–46.

13 Kirkland E J (1998) *Advanced Computing in Electron Microscopy.* (Plenum Press, New York.)

14 Watanabe K, Yamazaki T, Hashimoto I, and Shiojiri M (2001) Atomic-resolution annular dark-field STEM image calculations. *Phys. Rev. B* **64**: 115432.

15 Ishizuka K (2002) A practical approach for STEM image simulation based on the FFT multislice method. *Ultramicroscopy* **90**: 71–83.

16 Anstis G R, Cai D Q, and Cockayne D J H (2003) Limitations on the s-state approach to the interpretation of sub-angstrom resolution electron microscope images and microanalysis. *Ultramicroscopy* **94**: 309–327.

17 Allen L J, Findlay S D, Oxley M P, and Rossouw C J (2003) Lattice-resolution contrast from a focused coherent electron probe. Part I. *Ultramicroscopy* **96**: 47–63.

18 Findlay S D, Allen L J, Oxley M P, and Rossouw C J (2003) Lattice-resolution contrast from a focused coherent electron probe. Part II. *Ultramicroscopy* **96**: 65–81.

19 Buxton B F, Loveluck J E, and Steeds J W (1978) Bloch waves and their corresponding atomic and molecular orbitals in high energy electron diffraction. *Phil. Mag A* **38**: 259–278.

20 Dwyer C and Etheridge J (2003) Scattering of Å-scale electron probes in silicon. *Ultramicroscopy* **96**: 343–360.

21 Möbus G and Nufer S (2003) Nanobeam propagation and imaging in a FEGTEM/STEM. *Ultramicroscopy* **96**: 285–298.

22 Pennycook S J and Jesson D E (1991) High-resolution Z-contrast imaging of crystals. *Ultramicroscopy* **37**: 14–38.

23 Bird D M (1989) Theory of zone axis electron diffraction. *J. Electron Microsc. Techniq.* **13**: 77–97.

24 Spence J C H and Cowley J M (1978) Lattice imaging in STEM. *Optik* **50**: 129–142.

25 van Aert S, den Dekker A J, van Dyck D, and van den Bos A (2002) Optimal experimental design of STEM measurement of atom column positions. *Ultramicroscopy* **90**: 273–289.

26 van Dyck D and Op de Beeck M (1996) A simple intuitive theory for electron diffraction. *Ultramicroscopy* **64**: 99–107.

27 Hartel P, Rose P, and Dinges C (1996) Conditions and reasons for incoherent imaging in STEM. *Ultramicroscopy* **63**: 93–114.