Scanning Electron Microscopy

Volume 3
Number 1 *3rd Pfefferkorn Conference*

Article 22

1984

Electron Image Simulation: A Complementary Processing Technique

Michael A. O'Keefe
University of California

Electron Optical Systems (pp. 209-220)
SEM Inc., AMF O'Hare (Chicago), IL 60666-0507, U.S.A.

ELECTRON IMAGE SIMULATION:

A COMPLEMENTARY PROCESSING TECHNIQUE

Michael A. O'Keefe

National Center for Electron Microscopy
Building 72, Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Telephone (415) 486-4610

Abstract

At present it is difficult to use direct image processing techniques to determine the specimen structure from electron micrographs obtained under non-linear imaging conditions, and impossible when the effects of dynamical scattering are strong (as in the case of thicker specimens). However, computing techniques are available to simulate high-resolution transmission electron microscope (HRTEM) images of postulated model structures. With these techniques it is possible to confirm the validity of interpretation of recorded micrographs, to help analyze crystal defects, to characterize microscope parameters, and to determine the ranges of validity of commonly used interpretive approximations. Because processed micrographs can lead to suitable model structures and, in turn, models can indicate optimum directions for processing, the two techniques are excellent complements.

Key Words: Processing, simulation, interpretation, modeling, computing, high-resolution transmission electron microscopy.

Introduction

The ultimate goal of the image processing of electron micrographs is to obtain all possible information about the specimen that may be contained in the micrograph. In fact, since the electron microscope acts somewhat as a spatial frequency filter, more information is available from a "focal series" of micrographs than from any single image. Saxton (1980) shows examples of the application of Schiske's (1973) generalization of the Wiener filter to focal series of micrographs under so-called "linear image" conditions.

Under more general conditions the so-called "second-order" terms can alter the weights of spatial frequencies present in the image and even introduce higher frequencies formed from combinations of the frequencies present in the linear-image contribution to the micrograph. Second-order terms complicate the processing procedure immensely, and no satisfactory method of incorporating them has been published. O'Keefe and Sanders (1976) suggested that it may be possible to remove the second-order component from experimental optimum defocus images by subtracting the minimum contrast (or Gaussian focus) image, since this latter image is formed by selecting the value of focus which produces minimum linear contribution to the image intensity and second-order contributions change only slowly with change in focus. Saxton (1980) has proposed removing the second-order contribution by processing bright-field/dark-field pairs of micrographs, but the method introduces some experimental difficulties. Kirkland (1982) has proposed a computational method, but O'Keefe and Saxton (1983) show that some of the approximations involved are unjustified. At present, therefore, linear transfer can be inverted, but non-linear transfer cannot. Thus in image reconstruction the linear approximation is often chosen even when it is realized that for thicker specimens, the approximation may be poor. After this initial linear step has been used to set up reasonable model structures and yield information on transfer function parameters, simulations can be made both with and without non-linear effects, not only to assist reconstruction, but also in order to assess the importance of the non-linear contribution, and thus the reliability of the initial linear step. In some circumstances the linear approximation can be used to quite high values of crystal thickness (Tanaka and Jouffrey, 1984), whereas in others it fails dramatically by 150Å (O'Keefe and Saxton, 1983).

Present processing methods (and even future ones incorporating the second-order effects) are capable only of determining the complex amplitude (modulus and phase)

of the electron wave leaving the exit surface of the specimen. Under some conditions this electron wave is a simple function of the specimen structure projected in the direction of the electron beam, so that the structure is easily derived from the exit surface wave. This procedure is possible when the specimen scatters sufficiently weakly to be considered a "phase-object" or even a "weak phase object". Conditions of strong scattering occur for thicker specimens composed of heavier atoms and viewed down low index planes where atoms superpose exactly. These conditions are often those desirable in studies of periodic structures, especially of defects in such structures. On the other hand, when atom positions do not superpose, as for the case of defects in the thickness direction, interpretation is often impossible even with the aid of image simulation. Fortunately, biological and organic crystals are weak scatterers and hence can be interpreted as phase objects to thicknesses of several hundred Angstrom units; similarly many oxides and silicate minerals have a thickness limit in the 50A to 100A range, as do the important semiconductors Si and GaAs. However most metals and alloys have limits of only a few tens of Angstroms. For specimens thicker than the above limits, no simple relationship between the projected specimen structure and the exit-surface wave can be found (e.g. Jap and Glaser 1980) and image simulations must be used.

For the simulation of HRTEM images, a model structure is proposed, assembled in the computer, and images computed incorporating the various microscope parameters. In the decade since simulated high-resolution images of known crystal structures first appeared in the "n-Beam Lattice Images" series of papers (Allpress et al., 1972; Lynch and O'Keefe, 1972, Anstis et al., 1973; O'Keefe, 1973; Lynch et al., 1975, O'Keefe and Sanders, 1975), it has become commonplace to interpret HRTEM images of uncharacterized structures by comparing them with simulated images. Such images are currently produced by a variety of programs written in various laboratories around the world; in fact, commercial packages such as Skarnulis' (1979) interactive program and the more general SHRLI suite of programs (O'Keefe et al., 1978) are now available.

Theory of Simulation

The starting point for image simulation is to model the electron microscope as a simple system of electron beam, specimen and lens system (Fig. 1). Generally the initial electron beam is considered to be a parallel beam of plane wave electrons. The microscope lens system is replaced by one spherically-aberrated lens which can be regarded as representing the objective lens. In a real electron microscope the objective lens has the crucial duty of re-assembling the diffracted beams emerging from the specimen into an image which is merely magnified further by subsequent lenses.

In any simulation three functions representing the electron wave amplitude must be computed at three positions within the model microscope; at the exit surface of the specimen, $f(\underset{\sim}{x})$; at the back-focal plane (diffraction plane) of the objective lens, $f_{\underset{\sim}{k}}$; and at the image plane of the lens, $\psi(\underset{\sim}{x})$. The computation of the exit-surface wave $f(\underset{\sim}{x})$, involves mainly the model specimen structure (the only microscope parameter involved is the energy of the electrons in the incident beam). The wave at the back-focal plane, $f_{\underset{\sim}{k}}$, is obtained via a simple Fourier transform of the exit-surface wave $f(\underset{\sim}{x})$. Calculation of the image from the electron wave at the diffraction plane does not involve the specimen but only microscope parameters, such as objective lens defocus and spherical aberration; together with the objective aperture size and position, these modify $f_{\underset{\sim}{k}}$ before it is transformed into $\psi(\underset{\sim}{x})$, the image amplitude.

Calculation of the Exit-Surface Wave

Self et al., (1983) give details on calculating $f_{\underset{\sim}{k}}$ by various methods and conclude that the multislice method (Goodman and Moodie, 1974) is preferable. In this method, the steps involved in calculating the (structure dependent) values of $f_{\underset{\sim}{k}}$ are:

(i) calculation of $V_{\underset{\sim}{k}}$, the Fourier coefficients of potential (structure factors) at reciprocal lattice points $\underset{\sim}{k}$, lying in the zone perpendicular to the electron beam direction. Summing over all atoms in the unit cell:

$$V_{\underset{\sim}{k}} = \frac{h^2}{2\pi m_e e V_c} \sum_j {}^{e}f_j(|\underset{\sim}{k}|) \exp\{-2\pi i(\underset{\sim}{k} \cdot \underset{\sim}{x}_j)\} \qquad (1)$$

where V_c is the volume of the unit cell, and ${}^{e}f_j$ and x_j are the electron scattering factor, and the position of the jth atom respectively. Here the electron scattering factor for each atom is defined as the Fourier transform of the potential distribution for that atom. Inclusion of all V_k within 4Å^{-1} of the origin of reciprocal space provides sufficient accuracy for most calculations.

(ii) Fourier transformation of the V_k values produces $\phi_p(x)$ the crystal potential of one unit cell projected in the direction of the electron beam. The effect of such a thin "slice" of crystal on the electron beam is that of a phase object, and the electron "transmission function" for the slice is

$$q(\underset{\sim}{x}) = \exp\{i\sigma\phi_p(\underset{\sim}{x})\Delta z\} \qquad (2)$$

where Δz is the "slice thickness" and σ is the interaction parameter for electrons of the designated energy.

(iii) The exit-surface wave at the desired crystal thickness, $H = m\Delta z$, is found from $q(\underset{\sim}{x})$ by iteration. After m slices the electron wave $f(\underset{\sim}{x})$ is given by

$${}^{m}f(\underset{\sim}{x}) = [{}^{m-1}f(\underset{\sim}{x}) \star {}^{m}p(\underset{\sim}{x})]\,{}^{m}q(\underset{\sim}{x}) \qquad (3)$$

where ${}^{m}q(\underset{\sim}{x})$ is the transmission function of the mth slice and ${}^{m}p(\underset{\sim}{x})$ is the small-angle approximation to the free space propagator for the distance between the (m-1)th and mth slices (i.e. the familiar Fresnel propagator); $\star$ represents the convolution operation. The diffraction plane wavefield ${}^{m}f_{\underset{\sim}{k}}$ is obtained from ${}^{m}f(\underset{\sim}{x})$ by Fourier transformation. Note that for heavy atoms, or unit cells large in the direction of the electron beam, the slice thickness Δz may need to be chosen smaller than the unit cell height. In such cases a number of projected potentials must be calculated and appropriate transmission functions stored for each slice. Many structures are formed from atoms sufficiently light to allow slice thicknesses from 3Å to 5Å without intracell slicing.

The iteration of (3) is best performed in reciprocal space, where the wavefunction exists only at the discrete points of the reciprocal lattice. Furthermore, as the Fourier components of the wavefunction fall off with increasing order, it is a good approximation to consider the reciprocal space wavefunction as bandwidth limited (i.e. having intensity in only a finite number of beams). In reciprocal space, (3) becomes

$${}^{m}f_{\underset{\sim}{k}} = [{}^{m-1}f_{\underset{\sim}{k}} \cdot {}^{m}P_{\underset{\sim}{k}}] \star {}^{m}Q_{\underset{\sim}{k}} \qquad (4)$$

where P_k and Q_k are the Fourier transforms of $p(\underset{\sim}{x})$ and $q(\underset{\sim}{x})$. This form gives $^{m}f_k$ directly and is the form used in the SHRLI programs (O'Keefe et al., 1978) and shown in Figure 1.

Now that large amounts of computer memory are available to an individual user, an alternative to both the real-space (3) and diffraction-space (4) methods has become viable. This method uses the fast Fourier transform (FFT) algorithm of Cooley and Tukey (1965) to replace the convolution by an FFT followed by a simple multiplication and an inverse FFT (Ishizuka and Uyeda, 1977). This procedure is faster than direct convolution for large numbers of beams and extremely fast when programmed on an array processor. One problem is that only 2/3 of the extent of the FFT array can be used in order to avoid the aliasing produced by multiplication of two functions in a pseudo-periodic diffraction space, the pseudo-periodicity resulting from sampling the continuous functions $f(\underset{\sim}{x})$ and $q(\underset{\sim}{x})$. Thus only 4/9 of the two-dimensional array can be used for active beams without the danger of overlap of adjacent diffraction cells over-emphasizing the amplitudes of the outer reflections.

Van Dyck (1980) has used an approximate form of (3) to calculate the scattering in real space. Although the diffraction-space method (4) requires a time proportional to N^2 (where N is the number of diffracted beams), and the FFT method requires a time proportional to N log N, the time for Van Dyck's small-block convolution method is proportional to N. Unfortunately practical trials (Self, 1982; Kilaas and Gronsky, 1983) showed that this method required much larger values of N and smaller values of Δz than both the FFT and diffraction-space methods, resulting in calculations that took 4 to 5 times as long as an FFT multislice to produce an equivalent result.

Calculation of the Image-Plane Wave

In order to compute the image plane intensity from the exit surface wave, we need to include the effects of objective lens defocus and spherical aberration. These parameters act merely to change the phases of the diffracted electron beams passing through the aperture of the lens. Thus:

$$\Psi_{\underset{\sim}{k}} = f_{\underset{\sim}{k}} A_{\underset{\sim}{k}} \exp\{-i\gamma(\underset{\sim}{k})\} \qquad (5)$$

where:

$$\gamma(\underset{\sim}{k}) = \pi\lambda\underset{\sim}{k}^2(\Delta f + 1/2\,\lambda^2 C_s \underset{\sim}{k}^2) \qquad (6)$$

and $|\underset{\sim}{k}| = 2s = 2\sin\theta/\lambda$, C_s is the spherical aberration coefficient of the lens, and Δf is the amount of defocus of the objective lens from the Gaussian or in-focus position (here a negative value of Δf has been chosen to correspond to the underfocus condition produced by a weakened lens, and a positive value to overfocus). The objective aperture function A_k is unity for beams passing through the aperture and zero for those outside.

The intensity in the image plane can be calculated by Fourier transform of Ψ_k to obtain the amplitude $\psi(\underset{\sim}{x})$ followed by squaring of the amplitude:

$$I(\underset{\sim}{x}) = \psi(\underset{\sim}{x}) \cdot \psi^*(\underset{\sim}{x}) \; . \qquad (7)$$

Early programs (Lynch and O'Keefe, 1972) used the above procedure to form the image, relying on the objective aperture function to select the correct number of diffracted beams in order to limit the resolution in the simulated image to that of the electron microscope. However, it was soon found that images calculated with an aperture cutoff corresponding to that used experimentally contained more detail (O'Keefe, 1973) than the experimental micrograph.

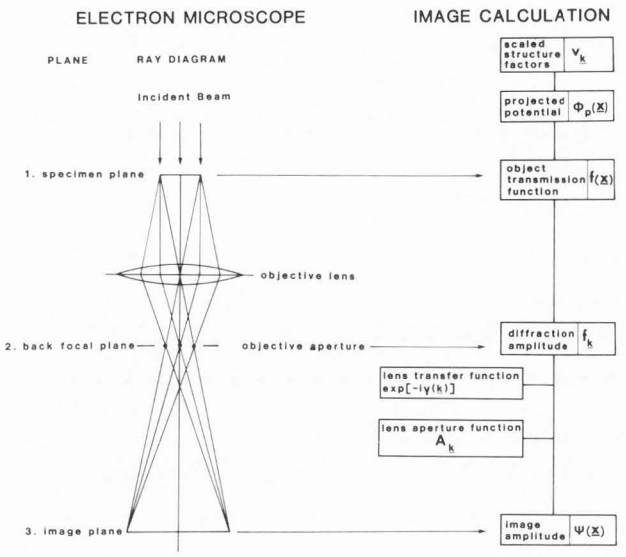


Fig. 1. Model of a simplified electron microscope together with the associated functions for simulating high-resolution images, and showing the three microscope planes at which the electron wavefield is required.

O'Keefe and Sanders (1975) considered two types of "smearing" aberrations which could be responsible for loss of resolution, the convergent character of the incident electron beam and the spread of focus produced by energy spread in the incident electrons. Both these aberrations have the effect of smearing the microscope image by making it a composite of higher-resolution images. Incident beam convergence produces a composite image formed by the summation of many images, each at a different angle within the incident cone. Spread of focus produces a composite formed from images summed over a range of defocus.

O'Keefe and Sanders (1975) measured the experimental value of convergence from a diffraction pattern obtained with the microscope illumination set as for imaging (Fig. 2). Calculation of 49 images at sampling positions within the convergent cone (Fig. 3) produced a set of images (Fig. 4) which could be summed to match the experimental result (Fig. 5). Fejes (1977) included the effect of spread of focus, and also found a degradation in image resolution.

In order to obtain a better understanding of the manner in which resolution is degraded by beam convergence and spread of focus, we can reformulate (7) by taking its Fourier transform to obtain

$$I_{\underset{\sim}{k}} = \Psi_{\underset{\sim}{k}} \star \Psi^*_{-\underset{\sim}{k}} \qquad (8)$$

where I_k is the image intensity spectrum, and $\star$ represents the convolution operation. Writing out the convolution gives

$$I_{\underset{\sim}{k}} = \sum_{\underset{\sim}{k}'} \Psi_{\underset{\sim}{k}'} \Psi^*_{\underset{\sim}{k}'-\underset{\sim}{k}} \; . \qquad (9)$$

Thus I_k, the image intensity spectrum, the Fourier transform of the image intensity, can be calculated directly from the Ψ_k, the aberrated diffraction-plane wavefield; the image intensity is formed by later Fourier transform of the intensity spectrum.

Figure 6 shows the image intensity spectrum formed by an aberrated diffraction-plane wavefield (or image amplitude spectrum) consisting of five colinear diffracted

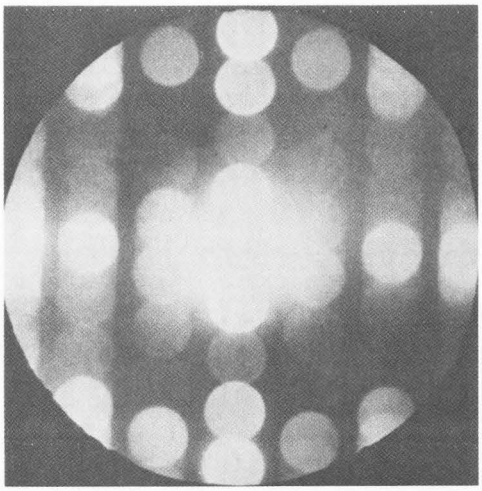


Fig. 2. 001 diffraction pattern of $Nb_{12}O_{29}$ obtained with illumination adjusted for imaging, and showing discs due to electron beam convergence in imaging mode.

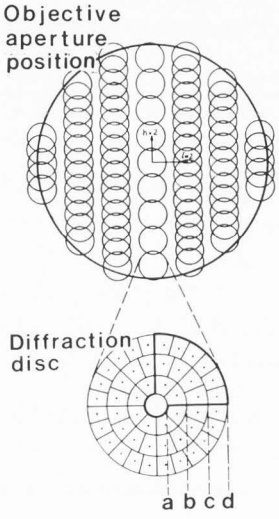


Fig. 3. Method of including beam convergence in the image simulation. Each disc admitted by the objective aperture is sampled at 49 points and images calculated at angles corresponding to each sampling point.

beams. We see that an amplitude spectrum containing terms out to h = ± 2 produces an intensity spectrum with frequencies out to h = ± 4. As shown in the figure, each frequency in the intensity spectrum is made up of a number of terms produced by pairwise multiplication of members of the amplitude spectrum. Terms containing the zeroth order of the amplitude spectrum are called first-order (or linear image) terms (note that they do not occur in frequencies greater than the amplitude spectrum limits of h = ± 2), and those containing no zeroth order are called second-order (or non-linear) terms.

Linear Images

In wavefields (and hence images) from specimens thin enough to be regarded as phase objects, most of the electron intensity resides in the central or zeroth order diffracted beam. In this type of situation the amplitude

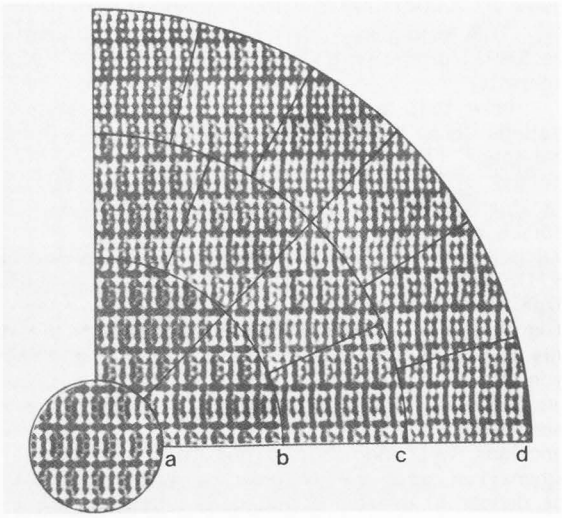


Fig. 4. Tilted-beam images obtained at angles within one quadrant of the convergence cone (the other quadrants are generated from these images by symmetry). The radii (marked a through d) correspond to those so marked in fig. 3. The specimen is 50Å thick $Nb_{12}O_{29}$ imaged at 100keV with an objective aperture size of 0.308 $Å^{-1}$, C_s = 1.8mm, and defocus = -600Å.

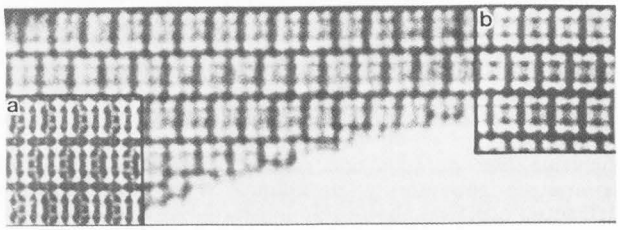


Fig. 5. Experimental image of $Nb_{12}O_{29}$ with inset images simulated without (a) and with (b) inclusion of the convergence effect.

of the zeroth beam is typically 0.9 of the amplitude of the incident beam, while the stronger diffracted beams may attain values of 10^{-3}. The ratio of the weights of the first order terms to those of the second order is thus approximately one thousand times, meaning that the second-order terms may safely be disregarded in interpreting this type of image.

Linear Image Theory

Dropping second-order terms in (9) we find that only two linear terms remain,

$$L_{I_{\underset{\sim}{k}}} = \psi_{\underset{\sim}{k}}\psi^*_0 + \psi_0\psi^*_{-\underset{\sim}{k}} \qquad (10)$$

and this is confirmed by figure 6.

Substituting from (5) and recalling $\psi^*_0 = \psi_0 \doteq 1$ we get

$$L_{I_{\underset{\sim}{k}}} = f_{\underset{\sim}{k}} \exp\{-i\gamma(\underset{\sim}{k})\} + f^*_{-\underset{\sim}{k}} \exp\{+i\gamma(-\underset{\sim}{k})\}. \qquad (11)$$

From (2), the object transmission function for a phase object of thickness H is

$$q(\underset{\sim}{x}) = \exp\{i\sigma\phi_p(\underset{\sim}{x})H\} . \qquad (12)$$

For a weak phase object (WPO) we can make a kinematic scattering approximation and approximate the transmission function as

$$^{WPO}q(\underset{\sim}{x}) = 1 + i\sigma\phi_p(\underset{\sim}{x})H \qquad (13)$$

so that Fourier transformation gives

$$^{WPO}f_{\underset{\sim}{k}} = \delta(\underset{\sim}{k}) + i\sigma H\, V_{\underset{\sim}{k}} \ . \qquad (14)$$

Substituting for $^{WPO}f_k$ in (11) produces the linear image approximation to the image intensity spectrum as,

$$^{L}I_{\underset{\sim}{k}} = \sigma H\, V_{\underset{\sim}{k}} \sin\gamma(\underset{\sim}{k}) + \sigma H\, V^{*}_{-\underset{\sim}{k}} \sin\gamma(-\underset{\sim}{k})$$

and since $V^{*}_{-k} = V_k$, and γ is radially symmetric for a properly-aligned microscope,

$$^{L}I_{\underset{\sim}{k}} = 2\sigma H\, V_{\underset{\sim}{k}} \sin\gamma(\underset{\sim}{k}) \quad \text{for } \underset{\sim}{k} \neq 0$$

For k = 0 the zeroth frequency is approximately $\delta(k)$ so that the full expression is

$$^{L}I_{\underset{\sim}{k}} = \delta(\underset{\sim}{k}) + 2\sigma H\, V_{\underset{\sim}{k}} \sin\gamma(\underset{\sim}{k}) \ . \qquad (15)$$

Thus for a specimen sufficiently thin that (i) the kinematic approximation to electron scattering is adequate, and (ii) the second-order intensity spectrum terms are small, the image intensity may be found by Fourier transformation of (15), or alternatively an experimental micrograph may be interpreted as if the transfer of spatial frequency V_k into the image were controlled by the value of sin γ at that value of k. Since C_s is always positive (for magnetic lenses) and Δf may be chosen to be negative by underfocussing the lens, the shape of the curve of sin $\gamma(k)$ against k can be controlled to some extent by choosing particular values of underfocus.

Scherzer (1949) searched for the optimum linear image and found that it occurs when a value of defocus near to $-\sqrt{1.5C_s\lambda}$ is chosen. In this case, sin $\gamma(k)$ is approximately equal to minus one over a band of frequencies extending to

$$1.5\, C_s^{-1/4} \lambda^{-3/4} \ .$$

Thus for frequencies within this band

$$^{L}I_{\underset{\sim}{k}} = \delta(\underset{\sim}{k}) - 2\sigma H\, V_{\underset{\sim}{k}} \qquad (16)$$

and the image intensity is

$$^{L}I(\underset{\sim}{x}) = I - 2\sigma H\, ^{L}\phi(\underset{\sim}{x}) \qquad (17)$$

where $^{L}\phi(\underset{\sim}{x})$ is a projected potential with resolution limited to the highest frequency V_k falling within the passband. This type of image is extremely important, because it shows directly the projected potential of the specimen, albeit to a limited resolution, with low intensity (black) in regions of high potential (high concentrations of atoms) and high intensity (white) in low potential regions (tunnels). This type of image is often called a structure image, and the particular value of underfocus is referred to as "Scherzer" defocus. Figure 7 shows a sin $\gamma(k)$ curve, or linear image "contrast transfer function" for a value of defocus close to Scherzer focus; the passband and cutoff frequency are marked. The cutoff frequency defines the structure image resolution, or Scherzer resolution, for any particular electron microscope.

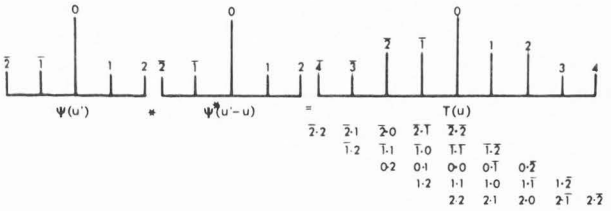


Fig. 6. Representation of the nine spatial frequencies generated in the image intensity spectrum by an amplitude spectrum consisting of five collinear diffracted beams. The terms contributing to each frequency are tabulated in abbreviated form--the term n.m represents multiplication of the pair $\Psi_n \cdot \Psi^*_m$ from the amplitude spectrum.

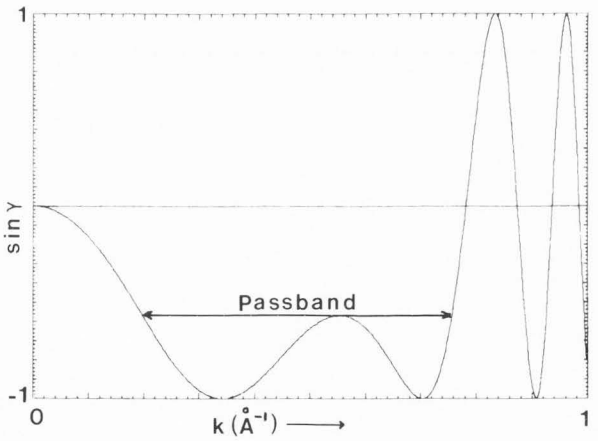


Fig. 7. Linear image contrast transfer function (plot of sin $\gamma(\underset{\sim}{k})$ against $|\underset{\sim}{k}|$) for an objective lens close to Scherzer defocus. Spatial frequencies are transferred with weights proportional to sin γ. Frequencies within the passband (marked) are considered to contribute without significant attenuation. The cutoff frequency that determines resolution in structure images lies at the upper limit of the passband.

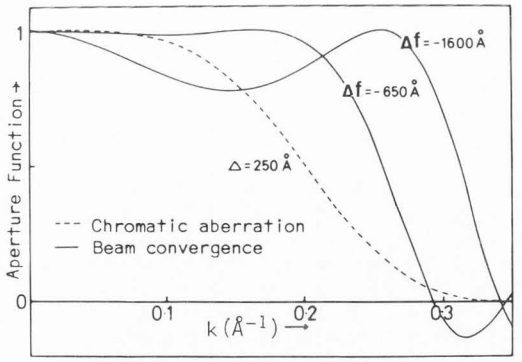


Fig. 8. Plots of the linear damping functions (aperture functions) due to a spread of focus with a gaussian half-width of 250Å (dashed) and a beam convergence with a semi-angle of 1.4 milliradian (solid line) for a 100keV electron microscope with C_s = 1.8mm.

Linear Image Resolution

Frank (1973) re-formulated the resolution-limiting effects of both beam convergence and spread of focus into expressions in reciprocal space. He showed that the effect on linear images is to impose envelope functions or soft apertures on the linear image amplitude, limiting transference of the higher spatial frequencies. Aperture functions for a Gaussian spread of focus of half-width 250Å , and for a beam convergence semi-angle of 1.4 milliradian are shown in figure 8. The aperture function due to spread of focus (or chromatic aberration) slopes gradually to zero, transferring higher frequencies only partially. The convergence aperture function is steeper, and its shape and cutoff frequency change with changes in defocus; at higher degrees of underfocus the cutoff frequency is higher, allowing more higher frequencies to contribute to the image, but mid-range frequencies are damped.

Whereas an accurate value of the beam convergence α can be measured directly from a diffraction pattern obtained with focussed illumination (fig. 2), the spread of focus Δ for a particular electron microscope must either be estimated by matching an experimental image with ones computed using different values of Δ, or approximated from known values of the chromatic aberration coefficient C_c and the high voltage and lens current ripple. A good approximation to Δ is obtained from

$$\Delta = C_c\sqrt{(\delta V/V)^2 + 4(\delta I/I)^2 + (\delta E/E)^2} \qquad (18)$$

where $\delta V/V$ and $\delta I/I$ are the high voltage and lens current ripple (usually quoted in ppm) respectively, and $\delta E/E$ is the fractional energy spread of the electron beam.

Figure 9 shows experimental images of a thin crystal of the block oxide $Nb_{12}O_{29}$ obtained on two different electron microscopes. Below each micrograph is the linear image envelope function corresponding to that microscope, and simulated images are inset. Notice that the factor limiting the resolution of the 100keV microscope is the large convergence factor (semi-angle 1.4 milliradian), while the 1MeV microscope is limited by the envelope due to a spread of focus of 500A. The value of 500A was obtained by application of equation 18 to a chromatic aberration coefficient of 4.4mm, and high voltage and lens current ripples of 5ppm; the energy spread was taken as 1eV, producing a Δ value of $\Delta = 44\sqrt{25 + 100 + 1} = 494$Å. As it happened, $Nb_{12}O_{29}$ images computed for this 1MeV microscope are not very sensitive to moderate changes in the value of Δ. Images computed for a range of Δ values and compared with experiment show that the correct value of Δ lies between 400A and 600A (Fig. 10).

It appears quite common that high voltage electron microscopes have linear-image resolutions limited by the spread of focus effect, presumably because of the difficulty of obtaining low values of ripple. Conversely, lower voltage (100keV) microscopes are usually limited by the effects of the relatively large convergence angles used to obtain sufficient brightness on the fluorescent screen. Of course, spread of focus becomes a significant resolution-limiting factor even in lower-voltage microscopes if Δ is made immoderately large by increasing the energy spread of the electron beam leaving the thermionic filament. Whereas at 1MeV a value of δE of 1eV produces an insignificant contribution to Δ, such a value at 100keV represents 10ppm and contributes a larger proportion than the voltage and current ripple. Krivanek (1975) demonstrated that a high beam current of 30μA (corresponding to an energy spread of 5eV half-width) reduced the linear image resolution of a 125keV electron microscope to worse than 3.4Å.

Non-Linear Images

Images obtained from thicker regions of specimens, as well as dark-field images, cannot be described in terms of linear contrast transfer functions, and hence the concept of the envelope function does not apply.

Non-Linear Resolution

O'Keefe (1979) extended the concept of the linear-image envelope function to include the non-linear terms, and showed that neglect of the "cross-terms" present in the general non-linear (but not the linear) damping functions could lead to the overdamping and loss of higher frequencies in the image intensity spectrum. Figure 11(a) shows how the full damping function due to beam convergence acts on the terms contributing to each frequency in the intensity spectrum, while (b) shows the effect on these terms if linear damping envelope functions are applied. The degree of damping of any intensity spectrum term is read off the plot by considering the two members of the amplitude spectrum making up the term (fig. 6), and reading up or across from the origin of the damping function (located at the center of the plot). The 0.0 term (contributing to the 0 frequency of the intensity spectrum) is located at the origin of the plot and is undamped (white). Similarly, all the n.n terms (falling along the bottom-left to top-right diagonal) are also undamped in the general case (a), but are lost if linear damping is applied (b). Note that those terms which contribute to the linear image (and lie along horizontal and vertical axes through the origin of each plot) are damped equally by both the general and linear damping functions (as indicated by the inset cross-section in (b) showing the familiar linear envelope profile of fig. 8).

A plot of the difference between the general and linear damping functions for convergence (fig. 11c) shows that the main effect of using the linear form to include the effects of beam convergence and spread of focus in a simulated image is to overdamp the n.n terms contributing to the zero frequency. This has the effect of lowering the mean image intensity level, occasionally producing negative values of intensity, but not otherwise seriously affecting the appearance of the simulated image (negative intensities arise because the unphysical nature of the linear damping function violates the principle of particle conservation, as was pointed out by Rose, 1977).

The image changes occasioned by using linear damping for the spread of focus effect are much more serious. Figure 12(a) shows that, as in the convergence case, the general damping function for spread of focus also does not damp the n.n terms. In addition, however, terms of the form $\bar{n}.n$ are not damped by the general spread of focus function. Thus the intensity spectrum will always contain frequencies out to twice the order of the highest members of the amplitude spectrum (unless these happen to be damped by a sufficiently large convergence).

We can apply the spread of focus function of figure 12(a) to the intensity spectrum terms listed in figure 6. Consider the $\bar{2}$ frequency in the intensity spectrum; its contributing terms are the linear $\bar{2}.0$ and 0.2 interactions, together with the non-linear $\bar{1}.1$ term. For a sufficiently large value of Δ, the linear terms would be heavily damped, but the non-linear one would not (as marked in figure 12a) so that this frequency would be present in the image. Using the linear damping function (figure 12b) the non-linear term is damped and the frequency will be missing from the image.

Previously we discussed the structure image (or Scherzer) resolution limit determined by the cutoff frequency at which the dominant aperture function blocks linear transfer into the intensity spectrum (as in figure 8).

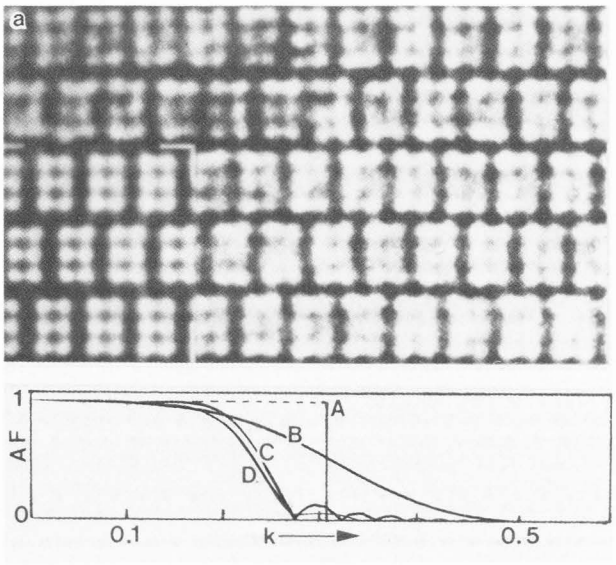


Fig. 9. Structure images of $Nb_{12}O_{29}$ taken at (a) 100kV and (b) 1MV. The inserts at lower left of each micrograph are calculated images for a 38Å thick crystal of $Nb_{12}O_{29}$. The aperture functions (AF, plotted below each image as a function of $\underset{\sim}{k}$ in reciprocal Å) show the resolution conditions under which each calculation was carried out. At 100kV the physical aperture (A) at $\underset{\sim}{k}$ = 0.308Å limits resolution to 3.2Å: the aperture function due to a defocus halfwidth of 100Å (B) limits resolution to 2.4Å; while the aperture function due to an incident beam convergence of 1.4 milliradian (C) restricts it to 3.8Å. The combined effect of these functions (D) results in an image of 3.8Å resolution. At 1MV the physical aperture (A) and convergence aperture function (C) limit resolutions to 1.9Å and 1.5Å respectively. For the calculated image to match the experimental result required a defocus-depth halfwidth of 500Å, resulting in the B curve shown. The combined effect (D) is virtually identical to B and yields an image of 2.5Å resolution.

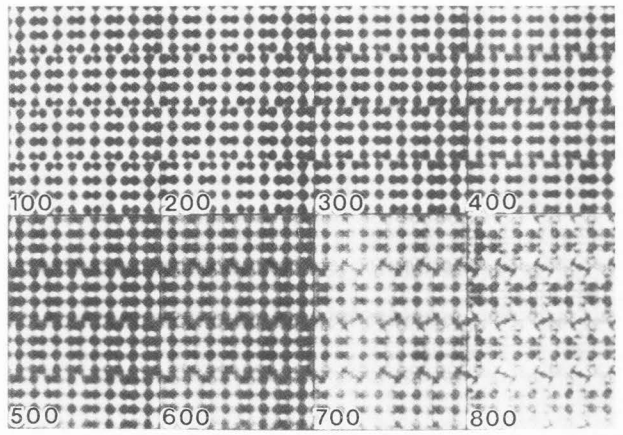


Fig. 10. Series of $Nb_{12}O_{29}$ images simulated for increasing values of Δ, the spread of focus halfwidth (marked in Å). Image resolution is degraded as Δ increases.

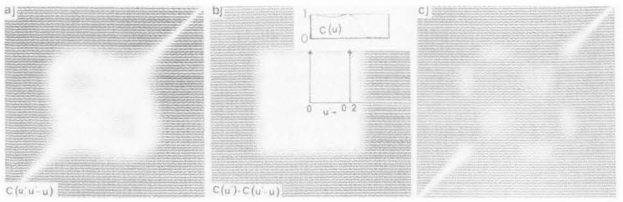


Fig. 11. Intensity spectrum damping functions due to beam convergence: (a) general (linear plus non linear) convergence damping function; (b) linear damping function with inset trace along a linear-image-term direction; (c) the inaccuracy in using linear damping (plot (a) minus plot (b)).

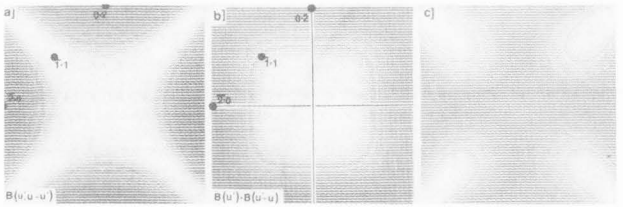


Fig. 12. Intensity spectrum damping functions for spread of focus: (a) general damping function; (b) linear damping function; (c) the difference. The inset spots show interferences contributing to the 2nd order frequency of the intensity spectrum of fig. 6.

Non-linear transfer into the intensity spectrum gives rise to yet another resolution limit. This non-linear resolution limit is sometimes called the "fringe resolution" of the microscope, and is given by the highest frequency transferred into the intensity spectrum. The general damping functions (figs. 11, 12) show that the highest order term transferred will be of the form n.n since this type of interference term is undamped by the spread of focus function (fig. 12) and passed by the convergence function (fig. 11) at a maximum value determined only by the value of focus; if the value of focus is chosen to allow the nth order member of the amplitude spectrum to contribute to the image via a linear interference term of

0.n, then the non-linear interference $\bar{n}.n$ will also be passed to appear in the image, even if the linear term 0.n is blocked by the spread of focus function. The fringe resolution of an electron microscope is thus twice as good as the limit set by its linear convergence aperture function (provided, of course, that the microscope is vibration-free and is aligned sufficiently accurately that the $\bar{n}.n$ interference falls within the narrow white arm of the spread of focus function, that is that the $\bar{n}$ and n diffracted beams form equal angles with the optic axis of the microscope).

Real-Space Second-Order Effects

Whereas images from thin crystals show detail to the linear-image resolution of the electron microscope, images from thicker crystals contain linearly-transferred detail to the same resolution plus non-linear detail to the non-linear resolution limit. In structure images the phases of the linear contribution are chosen to be negative (by selecting Scherzer focus) producing black spots at positions corresponding to areas of high potential (atoms or unresolved groups of atoms). In general, strong second-order contributions tend to have positive phase and produce white peaks at positions of high potential. These white peaks are sharper (coming from the high-frequency region of the intensity spectrum) than the linear black peaks, and produce the characteristic 'black donut with white hole' image in thicker structure images. Pirouz (1981) has shown that inclusion of the second-order terms is equivalent to adding a term proportional to the square of the projected crystal potential, and yet another proportional to the projected charge density, to the thin-crystal projected potential image; a negative-going (black) peak in the linear image (due to an atom or group of unresolved atoms) is thus squared to add a sharper positive-going (white) peak into the total (linear plus non-linear) image.

This effect can be seen in both experimental and computed images. Typically an image of a wedge crystal will show periodic black spots at the thin edge and the same black spots with white centers in thicker regions. Similarly, white spots appear in computed images of organic crystals as crystal thickness increases (O'Keefe et al., 1983). A pair of images of an organic crystal computed using the general and linear damping functions is shown in figure 13 together with their difference (formed by subtracting the images using the SEMPER programs of Saxton, 1979). Because the crystal is thin (19Å) the two images appear identical, but subtraction reveals a 3% difference in contrast. This difference appears at both the positions of minima (blacks) and maxima (whites) in the image, since the non-linear contribution arises from a squaring of the linear image contrast (Cowley, 1975). Although the difference (the non-linear contribution) has only 3% of image contrast for a crystal 19Å thick, this figure rises to 20% at 56Å thick. At a thickness of 94Å the contrast difference is 30% and differences in the images are easily visible (fig. 14).

In some structures non-linear contributions can produce images which appear like high-resolution structure images (Smith and O'Keefe, 1983); the "dumbbell" images of CdTe and Si viewed in [110] orientation are good examples of images dominated by non-linear contributions. Fig. 15 shows a series of "weak-phase-object images" of β-SiC in [110] orientation for increasing resolution. These images are the ones expected from very thin crystals imaged at Scherzer defocus; they show that each silicon/carbon atom pair appears as an unresolved black spot at 2.17Å resolution (or alternatively, the tunnels in the structure appear white). As resolution is improved the black spots elongate at 1.54Å resolution then split at 1.09Å resolution. Images computed for the Cambridge University HREM show dumbbell images with this splitting at crystal thicknesses of 150Å, although the linear-image resolution of the microscope is certainly worse than 1.09Å. Smith and O'Keefe (1983) show how the dumbbell images arise from a combination of non-linear and linear terms producing white spots at the atom positions over a limited range of crystal thickness and (non-Scherzer) microscope defocus. Interestingly, images from thick crystals at Scherzer defocus produce a type of dumbbell image where one white spot falls on an atom position, while the other is centered on a tunnel site.

It is considered generally well-known that the use of linear damping functions in image simulations is equivalent to an assumption that the electron beam is perfectly coherent. Rose (private communication - 1984) points out that the entire theory of STEM image formation is based on it--nevertheless it bears repeating to avoid misconceptions (e.g. O'Keefe and Saxton, 1983). Perfect spatial coherence (i.e. a condenser aperture filled with coherent electron waves) is possible in the case of an electron microscope equipped with a field emission gun rather than the more common thermionic kind. In this case the general form for the intensity spectrum (fig. 16, equation 1) may be replaced by one using a linear convergence function (fig. 16, equation 2). However, perfect temporal coherence appears impossible to achieve, so that a linear spread of focus form (fig. 16, equation 3) should not be used for image simulation (except within the unphysical weak phase object domain). Images of β-SiC calculated with linear damping terms (fig. 16, equation 3) do not produce the well-known split white dumbbells at a crystal thickness of 150Å, whereas images calculated with the general form do (fig. 17).

Dark-Field Images

The general form of the intensity spectrum can be used to simulate images in both bright-field and dark-field modes. In dark-field calculations all contributions to the intensity spectrum come from non-linear terms, and incorrect over-damping of these terms must be avoided by using the correct general damping functions. In the case of small-aperture dark field imaging, a large proportion of the contributing beams will have convergence cones that intersect the objective aperture and the simulation program must be capable of modelling this situation. Figure 18 shows diffraction patterns from a crystal of $4Nb_2O_5.9WO_3$ in [001] projection. When a small aperture is centered on the 130 reflection, the defocussed pattern (a) shows 9 diffracted beams passing through the aperture, whereas under imaging conditions portions of 19 discs lie inside the aperture (b). Taking this effect into account Iijima and O'Keefe (1977) successfully matched dark-field images (fig. 19a,b). Centering the aperture on $1\bar{3}0$ produced the expected switch in the positions of bright spots in the image (fig. 19c).

Misalignment and Crystal Tilt

One problem that arises in high resolution electron microscopy is accurate alignment of the electron microscope and of the specimen. While determining the structure of takeuchiite by HRTEM, Bovin et al., (1981), found that their best electron micrograph showed lower symmetry than images simulated from proposed models. Examination of the diffraction pattern revealed that the specimen was tilted 23 milliradian off the zone axis so that the center of the Laue circle was at 770. Figure 20 shows the experimental image with inset simulations for crystal thicknesses of 60 Å and 150Å. Note that the lower symmetry is more obvious in the simulation for the thicker crystal.

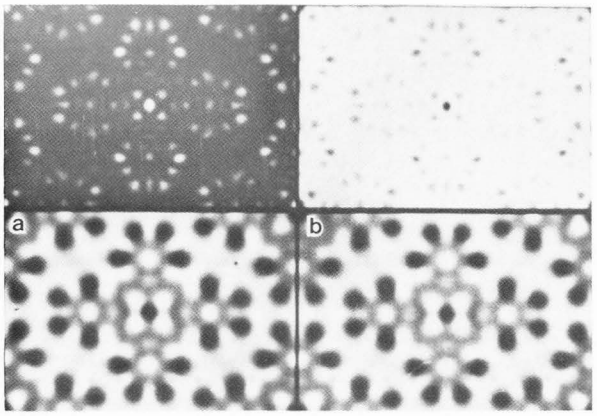


Fig. 13. Images of a thin (19Å) crystal of copper hexadecachlorophthalocyanine using (a) general and (b) linear damping functions. Simulation conditions are as published for the Kyoto HAREM (e.g. Kirkland, 1982). Difference images are shown above; difference contrast is 3% of image contrast.

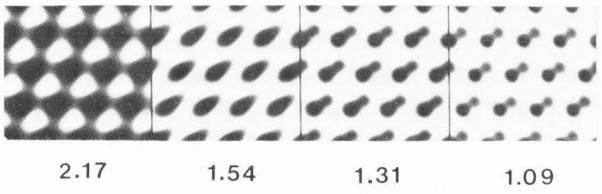


Fig. 15. Resolution series of weak-phase object images of β-SiC in [110] projection. Resolution (Å) is marked.

Fig. 16. Image intensity spectrum equations for (1) general damping functions, (2) linear convergence and general spread of focus, (3) linear convergence and spread of focus. Here the damping functions due to lack of lateral and longitudinal coherence (i.e. convergence and spread of focus) are designated by B and E.

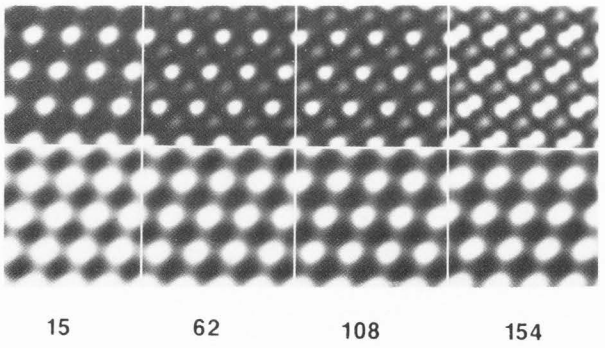


Fig. 17. Images of β-SiC [110] calculated for typical operating conditions (500kV, C_s = 3.5mm, Δf = -1100Å underfocus, convergence half-angle 0.3mrad, rms focus spread 250Å) for crystal thickness (from left) 15, 62, 108 and 154Å ; upper and lower rows calculated using equations (1) and (3) of fig. 16 respectively, and scaled individually to maximum contrast. Note the contrast reversal from the WPO images in fig. 15, due to the non-Scherzer value of -1100Å defocus.

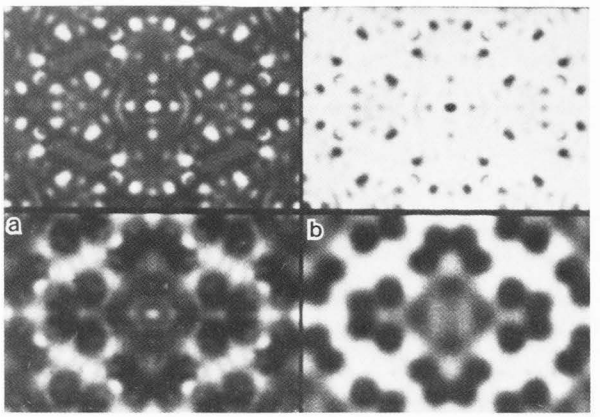


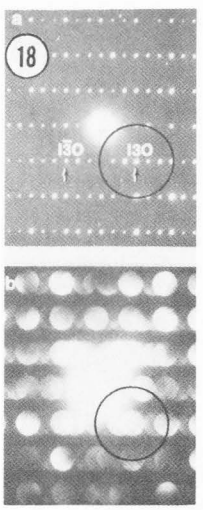
Fig. 14. Images of thick (94Å) copper hexadecachlorophthalocyanine computed with (a) general and (b) linear damping functions. Simulation conditions are for the Cambridge HREM (e.g. Smith and O'Keefe, 1983). Difference images (above) have 30% of image contrast.

16

$$I_{\underline{k}} = \sum_{\underline{k}'} f_{\underline{k}'} f^{*}_{\underline{k}'-\underline{k}} \exp\{-i[\gamma(\underline{k}') - \gamma(\underline{k}'-\underline{k})]\} \cdot B\{[\nabla\gamma(\underline{k}') - \nabla\gamma(\underline{k}'-\underline{k})]/2\pi\}\, E\{\tfrac{1}{2}(k'^2 - |\underline{k}'-\underline{k}|^2)\} \quad (1)$$

$$I_{\underline{k}} = \sum_{\underline{k}'} f_{\underline{k}'} \exp\{-i\gamma(\underline{k}')\}\, B\{\nabla\gamma(\underline{k}')/2\pi\} \cdot f^{*}_{\underline{k}'-\underline{k}} \exp\{+\gamma(\underline{k}'-\underline{k})\}\, B\{\nabla\gamma(\underline{k}'-\underline{k})/2\pi\}\, E\{\tfrac{1}{2}(k'^2-|\underline{k}'-\underline{k}|^2)\} \quad (2)$$

$$I_{\underline{k}} = \sum_{\underline{k}'} f_{\underline{k}'} \exp\{-i\gamma(\underline{k}')\}\, B\{\nabla\gamma(\underline{k}')/2\pi\}\, E\{\tfrac{1}{2}k'^2\} \cdot f^{*}_{\underline{k}'-\underline{k}} \exp\{+i\gamma(\underline{k}'-\underline{k})\}\, B\{\nabla\gamma(\underline{k}'-\underline{k})/2\pi\}\, E\{\tfrac{1}{2}|\underline{k}'-\underline{k}|^2\} \quad (3)$$



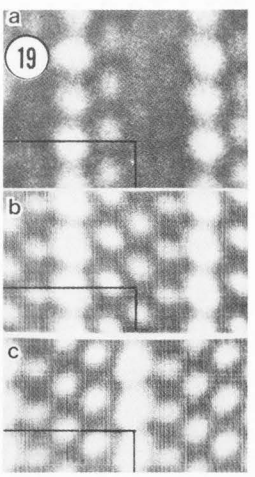


Fig. 18. Diffraction patterns from [001] $4Nb_2O_5.9WO_3$ for parallel (a) and focussed (b) illumination. The objective aperture size and position are shown. Note the many diffraction cones intersecting the aperture in (b).

Fig. 19. Dark-field images of $4Nb_2O_5.9WO_3$ for the conditions of fig. 18. (a) experimental; (b) computed with the aperture centered on 130; (c) computed with the aperture at $1\bar{3}0$.

Because the effects of crystal tilts and electron beam misalignment are easily incorporated in simulation programs, the programs are useful in exploring the results of tilts and misalignments on high resolution images. Computations by Smith et al., (1983), revealed that misalignments (the effect of having the electron beam at a slight angle to the optic axis of the objective lens) as small as one half milliradian could change the symmetry in thin (26Å) crystal images (fig. 21), whereas crystal tilts of up to 32 milliradian had little effect on such images. As the crystal thickness was increased the effect of crystal tilt increased; for a crystal of thickness 76A , significant changes occurred for tilts of 8 milliradian (fig. 22). Thus experimental images in which symmetry is lower merely in the thicker regions (as in fig. 20) indicate specimen tilt (or buckling), while an extension of lower symmetry to the thin edge probably calls for re-alignment of the microscope.

Calculation of Images of Defects

While high-resolution image simulation theory is usually presented for perfect crystals, it is possible to compute images of periodic and non-periodic faults. In programs using a multislice algorithm the method used is to create a large "supercell" with perfect crystal surrounding the fault. Such a calculation was made by O'Keefe and Iijima (1978) for a tetragonal tungsten bronze element embedded in a matrix of ten by ten unit cells of WO_3 (fig. 23). The simulated images matched an experimental focal series of micrographs of such a defect (fig. 24). This "periodic extension" method can be used to extend the multislice calculation to include not only faults, but also crystal edges (e.g. Marks and Smith, 1983).

In simulating images of defects by the periodic extension method outlined above, we are calculating the diffuse scattering from the defect only at certain points in reciprocal space--in effect we are "sampling" what should be a continuous distribution of diffuse scattering at a finite number of points on a grid whose interval is the reciprocal of the real-space supercell. In the calculation of O'Keefe and Iijima (1978), the choice of a ten by ten supercell produced a sampling of the diffuse scattering on a grid ten times finer than the reciprocal lattice containing the Bragg beams from the perfect WO_3 crystal. A finer sampling could have been produced by using a larger supercell, but since calculation size is always restricted (in this case to an array of 128 x 128 points) the calculation would not have extended out far enough in reciprocal space. The trade-off in sampling fineness and extent in reciprocal space restricts 128 x 128 calculations to supercell sizes of less than 40A and ideally less than 25A. Larger programs using arrays of 256 x 256 points (Krakow, 1980) and 512 x 512 points could safely handle defect cells to areas of 50Å by 50Å and 100Å by 100Å respectively. These larger areas will enable the extent of the lattice relaxation around the defect to be broadened.

Conclusion

Besides the obvious use of simulated images in testing models of specimen structure, simulation programs are also useful in analysing the effects of microscope parameters like spread-of-focus, in determining the importance of maladjustments such as misalignment and specimen tilt, and in setting limits on how far approximations like the "structure image" concept may be trusted. Modified programs can be used to explore unconventional imaging modes like dark-field hollow-cone (DFHC) and BFHC modes (e.g. O'Keefe and Pitt, 1980).

With the advent of cheaper, larger memory, 32-bit, "super-mini" computers, it should be possible to run large (perhaps 512 x 512) simulation programs on-line in interactive mode. With the addition of an associated array processor, such calculations should produce an image only a few minutes after a new structural model is read in. The image would, of course, be displayed on a video monitor in "split screen" mode for comparison with a digitized experimental micrograph.

Acknowledgement

The author is grateful for helpful comments from F. Lenz and J. Hren and equally helpful discussions with H. Rose. This work is supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U. S. Department of Energy under Contract No. DE-AC03-76SF00098.

References

Allpress JG, Hewat EA, Moodie AF, Sanders JV. (1972). n-beam lattice images. I. Experimental and computed images from $W_4Nb_{26}O_{77}$. Acta Cryst. A28, 528-536.

Anstis GR, Lynch DF, Moodie AF, O'Keefe MA. (1973). n-beam lattice images. III. Upper limits of ionicity in $W_4Nb_{26}O_{77}$. Acta Cryst. A29, 138-147.

Bovin J-O, O'Keeffe M, O'Keefe MA. (1981). Electron microscopy of oxyborates. III. On the structure of takeuchiite. Acta Cryst. A37, 42-46.

Cooley JW, Tukey JW. (1965) An algorithm for the machine calculation of complex Fourier series. Math. Comput. 19, 297-301.

Cowley JM. (1975). Diffraction Physics, North Holland/ American Elsevier, New York, 286-287.

Fejes P. (1977). Approximations for the calculation of high-resolution electron-microscope images of thin films. Acta Cryst. A33, 109-113.

Frank J. (1973) The envelope of electron microscope transfer functions for partially coherent illumination. Optik 38, 519-536.

Goodman P, Moodie AF. (1974) Numerical evaluation of N-beam wave functions in electron scattering by the multi-slice method. Acta Cryst. A30, 280-290.

Iijima S, O'Keefe MA. (1977). Interpretation of dark field images of crystals, in: 35th Ann. Proc. Electron Microscopy Soc. Amer., G.W. Bailey (ed), Claitor's, Baton Rouge, 190-191.

Ishizuka K, Uyeda N. (1977). A new theoretical and practical approach to the multislice method. Acta Cryst. A33, 740-749.

Jap BK, Glaser RM. (1980). The scattering of high-energy electrons. II. Quantitative validity of the single-scattering approximations for organic crystals. Acta Cryst. A36, 57-67.

Kilaas R, Gronsky R. (1983). Real space image simulation in high resolution electron microscopy. Ultramicroscopy 11, 289-298.

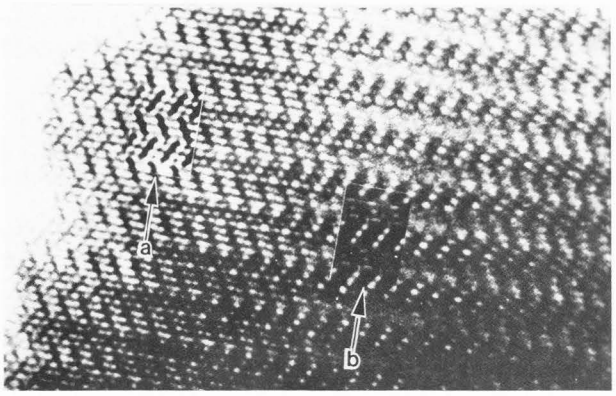


Fig. 20. Experimental image of takeuchiite with inset images computed for the experimentally-determined crystal tilt of 23 milliradians at crystal thicknesses of 60Å (a) and 150Å (b). Note the lower symmetry in the thicker simulated image.

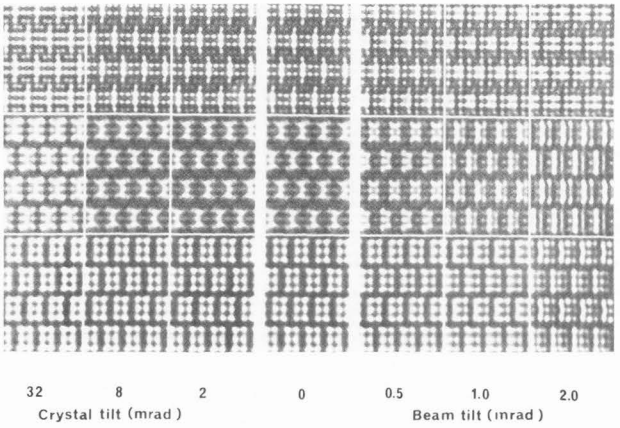


Fig. 21. Simulated images of thin (26Å) $Nb_{12}O_{29}$ for increasing degrees of beam and crystal tilt (marked) around the axis lying vertical on the page. Simulated at 100keV for C_s = 1.8mm, convergence semi-angle α = 1.4mrad, Δ = 140Å, and defocus values of -1527Å (upper), -1000Å (center) and -625Å (lower). Small misalignments produce changes in image symmetry. Crystal tilt effects are not as severe.

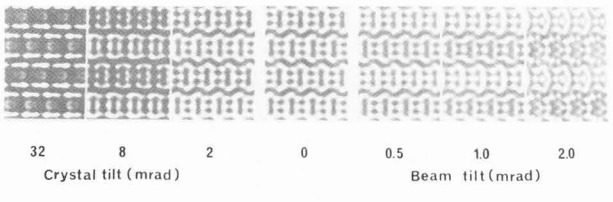


Fig. 22. Simulated images of thicker (76Å) $Nb_{12}O_{29}$ for increasing beam and crystal tilt (marked). Conditions as for fig. 21, with a defocus of -600Å. For this thicker crystal, tilt effects are stronger.

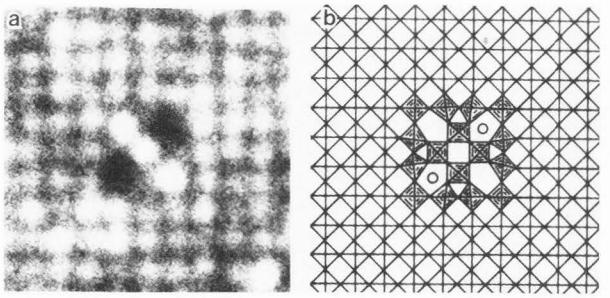


Fig. 23. (a) Structure image of a TTB element in a WO_3 matrix, and (b) the model derived from it.

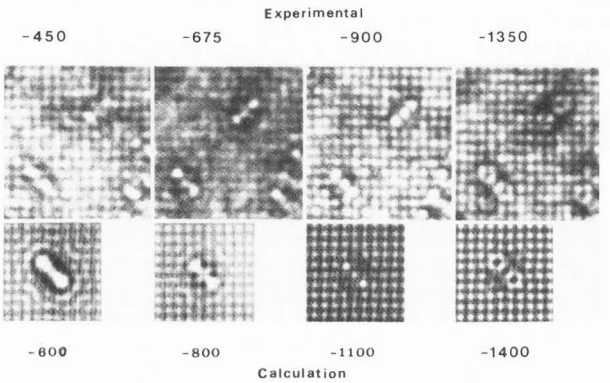


Fig. 24. Focal series of experimental TTB images with defocus values marked above in Å, together with best-match simulations from the model of fig. 23 computed for values of defocus marked below.

Kirkland EJ. (1982). Nonlinear high resolution image processing of conventional transmission electron micrographs. Ultramicroscopy 9, 45-74.

Krakow W. (1980). Computer modeling of high resolution electron microscope images using 65,536 beams, in: 38th Ann. Proc. Electron Microscopy Soc. Amer., G.W.Bailey (ed), Claitor's, Baton Rouge, 178-179.

Krivanek OL. (1975). The influence of beam intensity on the electron microscope contrast transfer function. Optik 43, 361-372.

Lynch DF, Moodie AF, O'Keefe, MA. (1975). n-beam lattice images. V. The use of the charge-density approximation in the interpretation of lattice images. Acta Cryst. A31, 300-307.

Lynch DF, O'Keefe MA. (1972). n-beam lattice images. II. Methods of calculation. Acta Cryst. A28, 536-548.

Marks LD, Smith DJ. (1983) Direct surface imaging in small metal particles. Nature 303, 316-317.

O'Keefe MA. (1973). n-beam lattice images. IV. Computed two-dimensional images. Acta Cryst. A29, 389-401.

O'Keefe MA. (1979). Resolution-damping functions in non-linear images, in: 37th Ann. Proc. Electron Microscopy Soc. Amer., G.W. Bailey (ed), Claitor's, Baton Rouge, 556-557.

O'Keefe MA, Buseck PR, Iijima S. (1978). Computed crystal structure images for high resolution electron microscopy. Nature 274, 322-324.

O'Keefe MA, Fryer JR, Smith DJ. (1983). High-resolution electron microscopy of molecular crystals. II. Image simulation. Acta Cryst. A39, 838-847.

O'Keefe MA, Iijima S. (1978). Calculation of structure images of crystalline defects, in: Electron Microscopy 1978, J.M. Sturgers (ed), Microscopical Society of Canada, Toronto, 282-283.

O'Keefe MA, Pitt AJ. (1980). The WPO image of $Ti_2Nb_{10}O_{29}$ as a measure of resolution, in: Electron Microscopy 1980, vol. 1, P. Brederos and G. Boom (eds), Seventh European Congress on Electron Microscopy Foundation, Leiden, 122-123.

O'Keefe MA, Sanders JV. (1975). n-beam lattice images. VI. Degradation of image resolution by a combination of incident-beam divergence and spherical aberration. Acta Cryst. A31, 307-310.

O'Keefe MA, Sanders JV. (1976). The phase component of lattice images of a zeolite crystal. Optik 46, 421-430.

O'Keefe MA, Saxton WO. (1983) The 'well known' theory of electron image formation, in: 41st Ann. Proc. Electron Microscopy Soc. Amer., G. W. Bailey (ed.), San Francisco Press, San Francisco, 288-289.

Pirouz P. (1981). Thin-crystal approximations in structure imaging. Acta Cryst. A37, 465-471.

Rose H. (1977) Nonstandard imaging methods in electron microscopy. Ultramicroscopy 2, 251-267.

Saxton WO. (1979). SEMPER--A portable image-processing system applied to electron microscopy, in: Inst. Phys. Conf. Ser. No. 44, Institute of Physics, London, 78-87.

Saxton WO. (1980). Correction of artifacts in linear and nonlinear high resolution electron micrographs. J. Microsc. Spectrosc. Electron. 5, 661-670.

Scherzer O. (1949) The theoretical resolution limit of the electron microscope. J. Appl. Phys. 20, 20-29.

Schiske P. (1973). Image processing using additional statistical information about the object, in: Image processing and computer-aided design in electron optics, P.W. Hawkes (ed), Academic, London, 82-90.

Self PG. (1982). Observations on the real space computation of dynamical electron diffraction intensities. J. Microscopy 127, 293-299.

Self PG, O'Keefe MA, Buseck PR, Spargo AEC. (1983). Practical computation of amplitudes and phases in electron diffraction. Ultramicroscopy 11, 35-52.

Skarnulis AJ. (1979). A system for interactive electron image calculation. J. Appl. Cryst. 12, 636-638.

Smith DJ, O'Keefe MA. (1983) Conditions for direct structure imaging in silicon carbide polytypes. Acta Cryst. A39, 139-148.

Smith DJ, Saxton WO, O'Keefe MA, Wood GJ, Stobbs WM. (1983) The importance of beam alignment and crystal tilt in high resolution electron microscopy. Ultramicroscopy 11, 263-282.

Tanaka M, Jouffrey B. (1984) Lattice-image interpretation of a relatively-small-unit-cell crystal. Acta Cryst. A40, 143-151.

Van Dyck D. (1980) Fast computational procedures for the simulation of structure images in complex or disordered crystals: a new approach. J. Microscopy 119, 141-152.

Additional Discussion

1. Question by J. Hren - Have you (or others) attempted to take into account relaxations associated with the crystal thickness. And, if so, please describe the results.
Answer - Although it is quite possible to include surface relaxations in the simulations (given an appropriate model), I have not done so. The paper by Marks and Smith (1983) describes some results of relaxation at a surface parallel to the beam, but I have never seen any simulations including relaxations at surfaces perpendicular to the incident beam direction.

2. Question by J. Hren - Under what conditions could you expect to image the effects of defects in the thickness direction?
Answer - Again, it is easily done using existing programs, but since an electron micrograph is largely a projection of the specimen structure, defects which produce overlapping structures in the beam direction would lead to confused contrast in both experimental micrograph and simulated image. Given a reasonable thickness of specimen, it might be very difficult to distinguish between simulations produced by different defect models.

3. Comment by F. Lenz - In the earlier literature the "smearing" of "damping" effects due to "convergence" of "spread of focus" were often described as lack of coherence, where "convergence" corresponds to lateral and "spread of focus" to longitudinal coherence.
Answer - Yes, these are certainly effects due to partial coherence. I used the terms "convergence" and "spread of focus" to describe how they were included in the simulation programs.

4. Comment by F. Lenz - The reason why so many authors have neglected second-order terms in the image transfer relations is not only the wish to simplify them. It was well known from the beginning that contrast transfer is a non-linear process, whereas amplitude transfer is linear. For image simulation it is relatively easy to conclude from object properties on image contrast without having to linearize, and it has been done before even for non-periodic objects. The inverse procedure, i.e. the deduction from image contrast of object properties is, however, much more complicated. Linear transfer can be inverted, non-linear transfer cannot. In image reconstruction one has to linearize even if one knows that for thicker specimens this may be a poor approximation. Since our original source of information is usually an electron micrograph, the first step in a reconstruction procedure must be linear. Its results will not only help to set up reasonable model structures but also yield information on the parameters describing the transfer function, such as the coefficients of spherical aberration and axial astigmatism and the defocus at which a micrograph was taken.
Answer - I agree. Image reconstruction and image simulation make excellent complementary techniques for obtaining maximum information from electron micrographs. Of course linear image reconstruction is the ideal method to use with images from thin crystals; however image simulation can also be useful in checking the degree of linearity of the experimental conditions under which the micrograph was obtained, as well as its more usual function of using the structure determined by the reconstruction to simulate images for comparison with experimental images obtained under conditions known to be non-linear, e.g. images from thick crystals, and dark-field images.