International Tables for Crystallography (2006). Vol. C, Section 4.3.8.5, pp. 425–427.

Misalignment of the electron beam, optic axis, and crystal axis in bright-field HREM work becomes increasingly important with increasing resolution and specimen thickness. The first-order effects of optical misalignment are an artifactual translation of spatial frequencies in the direction of misalignment by an amount proportional to the misalignment and to the square of spatial frequency. The corresponding phase shift is not observable in diffractograms. The effects of astigmatism on transfer functions for inclined illumination are discussed in Saxton (1978).

The effects of misalignment of the beam with respect to the optic axis are discussed in detail by Smith, Saxton, O'Keefe, Wood & Stobbs (1983), where it is found that all symmetry elements (except a mirror plane along the tilt direction) may be destroyed by misalignment. The maximum allowable misalignment for a given resolution δ in a specimen of thickness t is proportional to

$$\alpha = \delta/8t. \qquad (4.3.8.24)$$

Misalignment of a crystalline specimen with respect to the beam may be distinguished from misalignment of the optic axis with respect to the beam by the fact that, in very thin crystals, the former does not destroy centres of symmetry in the image.

The use of known defect point-group symmetry (for example in stacking faults) to identify a point in a HREM image with a point in the structure and so to resolve the black or white atomic contrast ambiguity has been described (Olsen & Spence, 1981). Structures containing screw or glide elements normal to the beam are particularly sensitive to misalignment, and errors as small as 0.2 mrad may substantially alter the image appearance.

A rapid comparison of images of amorphous material with the beam electronically tilted into several directions appears to be the best current method of aligning the beam with the optic axis, while switching to convergent-beam mode appears to be the most effective method of aligning the beam with the crystal axis. However, there is evidence that the angle of incidence of the incident beam is altered by this switching procedure.

The effects of misalignment and choice of beam divergence Θ_c on HREM images of crystals containing dynamically forbidden reflections are reviewed by Nagakura, Nakamura & Suzuki (1982) and Smith, Bursill & Wood (1985). Here the dramatic example of rutile in the [001] orientation is used to demonstrate how a misalignment of less than 0.2 mrad of the electron beam with respect to the crystal axis can bring up a coarse set of fringes (4.6 Å), which produce an image of incorrect symmetry, since these correspond to structure factors that are forbidden both dynamically and kinematically.

Crystal thickness is most accurately determined from images of planar faults in known orientations, or from crystal morphology for small particles. It must otherwise be treated as a refinement parameter. Since small crystals (such as MgO smoke particles, which form as perfect cubes) provide such an independent method of thickness determination, they provide the most convincing test of dynamical imaging theory. *The ability to match the contrast reversals and other detailed changes in HREM images as a function of either thickness or focus (or both) where these parameters have been measured by an independent method gives the greatest confidence in image interpretation.* This approach, which has been applied in rather few cases [see, for example, O'Keefe, Spence, Hutchinson & Waddington (1985)] is strongly recommended. The tendency for n-beam dynamical HREM images to repeat with increasing thickness in cases where the wavefunction is dominated by just two Bloch waves has been analysed by several workers (Kambe, 1982).

Since electron scattering factors are proportional to the difference between atomic number and X-ray scattering factors, and inversely proportional to the square of the scattering angle (see Section 4.3.1), it has been known for many years that the low-order reflections that contribute to HREM images are extremely sensitive to the distribution of bonding electrons and so to the degree of ionicity of the species imaged. This observation has formed the basis of several charge-density-map determinations by convergent-beam electron diffraction [see, for example, Zuo, Spence & O'Keefe (1988)]. Studies of ionicity effects on HREM imaging can be found in Anstis, Lynch, Moodie & O'Keefe (1973) and Fujiyoshi, Ishizuka, Tsuji, Kobayashi & Uyeda (1983).

The depletion of the elastic portion of the dynamical electron wavefunction by inelastic crystal excitations (chiefly phonons, single-electron excitations, and plasmons) may have dramatic effects on the HREM images of thicker crystals (Pirouz, 1974). For image formation by the elastic component, these effects may be described through the use of a complex 'optical' potential and the appropriate Debye–Waller factor (see Section 2.5.1). However, existing calculations for the absorption coefficients derived from the imaginary part of this potential are frequently not applicable to lattice images because of the large objective apertures used in HREM work. It has been suggested that HREM images formed from electrons that suffer small energy losses (and so remain 'in focus') but large-angle scattering events (within the objective aperture) due to phonon excitation may contribute high-resolution detail to images (Cowley, 1988). For measurements of the imaginary part of the optical potential by electron diffraction, the reader is referred to the work of Voss, Lehmpfuhl & Smith (1980), and references therein. All evidence suggests, however, that for the crystal thicknesses generally used for HREM work ($t < 200$Å) the effects of 'absorption' are small.

In summary, the general approach to the matching of computed and experimental HREM images proceeds as follows (Wilson, Spargo & Smith, 1982). (i) Values of Δ, Θ_c, and C_s are determined by careful measurements under well defined conditions (electron-gun bias setting, illumination aperture size, specimen height as measured by focusing-lens currents, electron-source size, *etc*). These parameters are then taken as constants for all subsequent work under these instrumental conditions (assuming also continuous monitoring of electronic instabilities). (ii) For a particular structure refinement, the parameters of thickness and focus are then varied, together with the choice of atomic model, in dynamical computer simulations until agreement is obtained. Every effort should be made to match images as a function of thickness and focus. (iii) If agreement cannot be obtained, the effects of small misalignments must be investigated (Smith *et al.*, 1985). Crystals most sensitive to these include those containing reflections that are absent due to the presence of screw or glide elements normal to the beam.

4.3.8.5. *Computing methods*

The general formulations for the dynamical theory of electron diffraction in crystals have been described in Section 5.2 of *IT* B (1992). In Section 4.3.6, the computing methods used for calculating diffraction-beam amplitudes have been outlined.

Given the diffracted-beam amplitudes, $\Psi_{\rm g}$, the image is calculated by use of equations (4.3.8.2), including, when appropriate, the modifications of (4.3.8.13*b*).

The numerical methods that can be employed in relation to crystal-structure imaging make use of algorithms based on (i) matrix diagonalization, (ii) fast Fourier transforms, (iii) real-space convolution (Van Dyck, 1980), (iv) Runge-Kutta (or similar) methods, or (v) power-series evaluation. Two other solutions, the Cowley–Moodie polynomial solution and the

Copyright © 2006 International Union of Crystallography

Feynman path-integral solution, have not been used extensively for numerical work. Methods (i) and (ii) have proven the most popular, with (ii) (the multislice method) being used most extensively for HREM image simulations. The availability of inexpensive array processors has made this technique highly efficient. A comparison of these two N-beam methods is given by Self, O'Keefe, Buseck & Spargo (1983), who find the multislice method to be faster (time proportional to $N \log_2 N$) than the diagonalization method (time proportional to N^2) for $N > 16$. Computing space increases roughly as N^2 for the diagonalization method, and as N for the multislice. The problem of steeply inclined boundary conditions for multislice computations has been discussed by Ishizuka (1982).

In the Bloch-wave formulation, the lattice image is given by

$$I(\mathbf{r}) = \sum_{i,j} \sum_{\mathbf{h},\mathbf{g}} C_0^{(i)} C_0^{(j)} C_\mathbf{g}^{(i)} C_\mathbf{h}^{(j)} \exp\{i[2\pi(\gamma^{(i)} - \gamma^{(j)})t + 2\pi(\mathbf{g} - \mathbf{h}) \cdot \mathbf{r} - \chi(\Delta f, C_s, \mathbf{g}) + \chi(\Delta f, C_s, \mathbf{h})]\}, \quad (4.3.8.25)$$

where $C_\mathbf{g}^{(i)}$ and $\gamma^{(i)}$ are the eigenvector elements and eigenvalues of the structure matrix [see Hirsch, Howie, Nicholson, Pashley & Whelan (1977) and Section 4.3.4].

Using modern personal computers or workstations, it is now possible to build efficient single-user systems that allow interactive dynamical structure-image calculations. Either an image intensifier or a cooled scientific grade charge-coupled device and single-crystal scintillator screen may be used to record the images, which are then transferred into a computer (Daberkow, Herrman, Liu & Rau, 1991). This then allows for the possibility of automated alignment, stigmation and focusing to the level of accuracy needed at 0.1 nm point resolution (Krivanek & Mooney, 1993). An image-matching search through trial structures, thickness and focus parameters can then be completed rapidly. Where large numbers of pixels, large dynamic range and high sensitivity are required, the Image Plate has definite advantages and so should find application in electron holography and biology (Shindo, Hiraga, Oikawa & Mori, 1990).

For the calculation of images of defects, the method of periodic continuation has been used extensively (Grinton & Cowley, 1971). Since, for kilovolt electrons traversing thin crystals, the transverse spreading of the dynamical wavefunction is limited (Cowley, 1981), the complex image amplitude at a particular point on the specimen exit face depends only on the crystal potential within a cylinder a few ångströms in diameter, erected about that point (Spence, O'Keefe & Iijima, 1978). The width of this cylinder depends on accelerating voltage, specimen thickness, and focus setting (see above references). Thus, small overlapping 'patches' of exit-face wavefunction may be calculated in successive computations, and the results combined to form a larger area of image. The size of the 'artificial superlattice' used should be increased until no change is found in the wavefunction over the central region of interest. For most defects, the positions of only a few atoms are important and, since the electron wavefunction is locally determined (for thin specimens at Scherzer focus), it appears that very large calculations are rarely needed for HREM work. The simulation of profile images of crystal surfaces at large defocus settings will, however, frequently be found to require large amounts of storage.

A new program should be tested to ensure that (*a*) under approximate two-beam conditions the calculated extinction distances for small-unit-cell crystals agree roughly with tabulated values (Hirsch *et al.*, 1977), (*b*) the simulated dynamical images have the correct symmetry, (*c*) for small thickness, the Scherzer-focus images agree with the projected potential, and (*d*) images and beam intensities agree with those of a program known to be correct. The damping envelope (product representation) [equation (4.3.8.17)] should only be used in a thin crystal with $\Phi_0 > \Phi_\mathbf{g}$; in general, the effects of partial spatial and temporal coherence must be incorporated using equation (4.3.8.13*a*) or (4.3.8.13*b*), depending on whether variations in diffraction conditions over θ_c are important. Thus, a separate multislice dynamical-image calculation for each component plane wave in the incident cone of illumination may be required, followed by an incoherent sum of all resulting images.

The outlook for obtaining higher resolution at the time of writing (1997) is broadly as follows. (1) The highest point resolution currently obtainable is close to 0.1 nm, and this has been obtained by taking advantage of the reduction in electron wavelength that occurs at high voltage [equation (4.3.8.16)]. A summary of results from these machines can be found in *Ultramicroscopy* (1994), Vol. 56, Nos. 1–3, where applications to fullerenes, glasses, quasicrystals, interfaces, ceramics, semiconductors, metals and oxides and other systems may be found. Fig. 4.2.8.6 shows a typical result. High cost, and the effects of radiation damage (particularly at larger thickness where defects with higher free energies are likely to be found), may limit these machines to a few specialized laboratories in the future. The attainment of higher resolution through this approach depends on advances in high-voltage engineering. (2) Aberration coefficients may be reduced if higher magnetic fields can be produced in the pole piece, beyond the saturation flux of the specialized iron alloys currently used. Research into superconducting lenses has therefore continued for many years in a few laboratories. Fluctuations in lens current are also eliminated by this method. (3) Electron holography was originally developed for the purpose of improving electron-microscope resolution, and this approach is reviewed in the following section. (4) Electron–optical correction of aberrations has been under study for many years in work by Scherzer, Crewe, Beck, Krivanek, Lanio, Rose and others – results of recent experimental tests are described in Haider & Zach (1995) and Krivanek, Dellby, Spence, Camps & Brown (1997). The attainment of 0.1 nm point resolution is considered feasible. Aberration correctors will also provide benefits other than increased resolution, including greater space in the pole piece for increased sample tilt and access to X-ray detectors, *etc*.

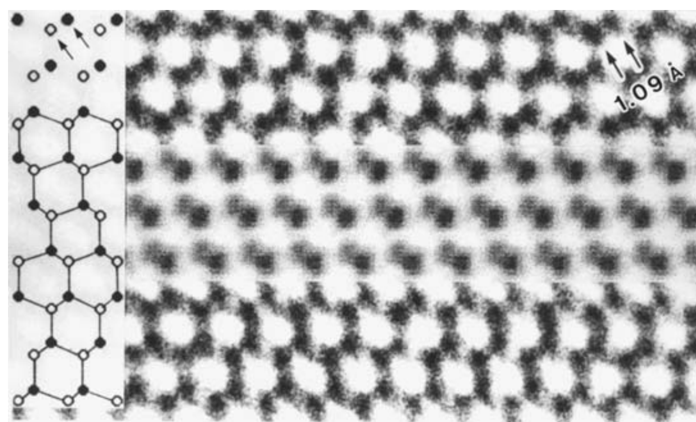

Fig. 4.3.8.6. Structure image of a thin lamella of the 6*H* polytype of SiC projected along [110] and recorded at 1.2 MeV. Every atomic column (darker dots) is separately resolved at 0.109 nm spacing. The central horizontal strip contains a computer-simulated image; the structure is sketched at the left. [Courtesy of H. Ichinose (1994).]

The need for resolution improvement beyond 0.1 nm has been questioned – the structural information retrievable by a single HREM image is always limited by the fact that a projection is obtained. (This problem is particularly acute for glasses.) Methods for combining different projected images (particularly of defects) from the same region (Downing, Meisheng, Wenk & O'Keefe, 1990) may now be as important as the search for higher resolution.

4.3.8.6. *Resolution and hyper-resolution*

Since the resolution of an instrument is a property of the instrument alone, whereas the ability to distinguish HREM image features due to adjacent atoms depends on the scattering properties of the atoms, the resolution of an electron microscope cannot easily be defined [see Subsection 2.5.1.9 in *IT* B (1992)]. The Rayleigh criterion was developed for the incoherent imaging of point sources and cannot be applied to coherent phase contrast. Only for very thin specimens of light elements for which it can be assumed that the scattering phase is $-\pi/2$ can the straightforward definition of point resolution d_p [equation (4.3.8.16)] be applied. In general, the dynamical wavefunction across the exit face of a crystalline sample bears no simple relationship to the crystal structure, other than to preserve its symmetry and to be determined by the 'local' crystal potential. The use of a dynamical 'R factor' between computed and experimental images of a known structure has been suggested by several workers as the basis for a more general resolution definition.

For weakly scattering specimens, the most satisfactory method of measuring either the point resolution d_p or the information limit d_i [see equation (4.3.8.21)] appears to be that of Frank (1975). Here two successive micrographs of a thin amorphous film are recorded (under identical conditions) and the superimposed pair used to obtain a coherent optical diffractogram crossed by fringes. The fringes, which result from small displacements of the micrographs, extend only to the band limit d_i^{-1} of information common to both micrographs, and cannot be extended by photographic processing, noise, or increased exposure. By plotting this band limit against defocus, it is possible to determine both Δ and θ_c. As an alternative, for thin crystalline samples of large-unit-cell materials, the parameters Δ, θ_c, and C_s can be determined by matching computed and experimental images of crystals of known structure. It is the specification of these parameters (for a given electron intensity and wavelength) that is important in describing the performance of high-resolution electron microscopes. We note that certain conditions of focus or thickness may give a spurious impression of ultra-high resolution [see equations (4.3.8.7) and (4.3.8.8)].

Within the domain of linear imaging, implying, for the most part, the validity of the WPO approximation, many forms of image processing have been employed. These have been of particular importance for crystalline and non-crystalline biological materials and include image reconstruction [see Section 2.5.4 in *IT* B (1992)] and the derivation of three-dimensional structures from two-dimensional projections [see Section 2.5.5 in *IT* B (1992)]. For reviews, see also Saxton (1980*a*), Frank (1980), and Schiske (1975). Several software packages now exist that are designed for image manipulation, Fourier analysis, and cross correlation; for details of these, see Saxton (1980*a*) and Frank (1980). The theoretical basis for the WPO approximation closely parallels that of axial holography in coherent optics, thus much of that literature can be applied to HREM image processing. Gabor's original proposal for holography was intended for electron microscopy [see Cowley (1981) for a review].

The aim of image-processing schemes is the restoration of the exit-face wavefunction, given in equation (4.3.8.13*a*). The reconstruction of the crystal potential $\varphi_p(\mathbf{r})$ from this is a separate problem, since these are only simply related under the approximation of Subsection 4.3.8.3. For a non-linear method that allows the reconstruction of the dynamical image wavefunction, based on equation (4.3.8.13*b*), which thus includes the effects of multiple scattering, see Saxton (1980*b*).

The concept of holographic reconstruction was introduced by Gabor (1948, 1949) as a means of enhancing the resolution of electron microscopes. Gabor proposed that, if the information on relative phases of the image wave could be recorded by observing interference with a known reference wave, the phase modification due to the objective-lens aberrations could be removed. Of the many possible forms of electron holography (Cowley, 1994), two show particular promise of useful improvements of resolution. In what may be called in-line TEM holography, a through-focus series of bright-field images is obtained with near-coherent illumination. With reference to the relatively strong transmitted beam, the relative phase and amplitude changes due to the specimen are derived from the variations of image intensity (see Van Dyck, Op de Beeck & Coene, 1994). The tilt-series reconstruction method also shows considerable promise (Kirkland, Saxton, Chau, Tsuno & Kawasaki, 1995).

In the alternative off-axis approach, the reference wave is that which passes by the specimen area in vacuum, and which is made to interfere with the wave transmitted through the specimen by use of an electrostatic biprism (Möllenstedt & Düker, 1956). The hologram consists of a modulated pattern of interference fringes. The image wavefunction amplitude and phase are deduced from the contrast and lateral displacements of the fringes (Lichte, 1991; Tonomura, 1992). The process of reconstruction from the hologram to give the image wavefunction may be performed by optical-analogue or digital methods and can include the correction of the phase function to remove the effects of lens aberrations and the attendant limitation of resolution. The point resolution of electron microscopes has recently been exceeded by this method (Orchowski, Rau & Lichte, 1995).

The aim of the holographic reconstructions is the restoration of the wavefunction at the exit face of the specimen as given by equation (4.3.8.13*a*). The reconstruction of the crystal potential $\varphi(\mathbf{r})$ from this is a separate problem, since the exit-face wavefunction and $\varphi(\mathbf{r})$ are simply related only under the WPO approximations of Subsection 4.3.8.3. The possibility of deriving reconstructions from wavefunctions strongly affected by dynamical diffraction has been considered by a number of authors (for example, Van Dyck *et al.*, 1994). The problem does not appear to be solvable in general, but for special cases, such as perfect thin single crystals in exact axial orientations, considerable progress may be possible.

Since a single atom, or a column of atoms, acts as a lens with negative spherical aberration, methods for obtaining superresolution using atoms as lenses have recently been proposed (Cowley, Spence & Smirnov, 1997).

4.3.8.7. *Alternative methods*

A number of non-conventional imaging modes have been found useful in electron microscopy for particular applications. In scanning transmission electron microscopy (STEM), powerful electron lenses are used to focus the beam from a very small bright source, formed by a field-emission gun, to form a small probe that is scanned across the specimen. Some selected part of the transmitted electron beam (part of the coherent convergent-

references