

- [9] D. VAN DYCK, *Adv. opt. Electron Phys.* **65**, 295 (1985).
- [10] G. R. ANSTIS, *Computer Simulation of Electron Microscope Diffraction and Images*, Ed. W. KRAKOW and M. O'KEEFE, The Minerals, Metals & Materials Soc., 1989 (p. 229).
- [11] A. HOWIE and Z. S. BASINSKI, *Phil. Mag.* **17**, 1039 (1962).
- [12] P. STADELMANN, *Ultramicroscopy* **21**, 131 (1987).
- [13] M. VAN HEEL and W. KEGSTRA, *Ultramicroscopy* **7**, 113 (1981).
- [14] K. SCHEERSCHMIDT and F. KNOLL, *Internat. Conf. Inverse Problems*, Potsdam 1993, to be published.
- [15] B. N. ZAKHARIEV and A. A. SUZKO, *Direct and Inverse Problems*, Springer-Verlag, Berlin/Heidelberg 1990.

(Received July 15, 1994)

phys. stat. sol. (a) **146**, 491 (1994)

Subject classification: 61.14; 61.16; 61.70

Max Planck Institute of Microstructure Physics, Halle¹⁾

Retrieval of Object Information from Electron Diffraction

I. Theoretical Preliminaries

By

K. SCHEERSCHMIDT and F. KNOLL

Dedicated to Professor Dr. Dr. h.c. HEINZ BETHGE on the occasion of his 75th birthday

Electron holography and other wave reconstruction techniques allow one to determine directly the scattered wave function at the exit surface of an object up to the information limit of the electron microscope. Based on the knowledge of the reconstructed complex electron wave and using an discretized form of the diffraction equations, this in principle enables the direct retrieval of the atomic displacements, caused by a crystal lattice defect, relative to the atom positions of the perfect lattice. A special inverse problem of electron scattering can be deduced considering solely the atomic displacements, which are given by the zeros of a function with an incompletely known Fourier spectrum. The fundamental relations are described, with the problems of solving the ill-posed Fourier transform being discussed.

Mittels Elektronenholographie oder anderen Verfahren zur Wellenrekonstruktion wird die gestreute Elektronenwellenfunktion an der Objektaustrittsseite direkt und vollständig bis zum Informationslimit des Elektronenmikroskops bestimmbar. Unter der Voraussetzung, daß die Elektronenwelle rekonstruiert vorliegt, und daß die Beugungsgleichungen in einer diskretisierten Form verwendet werden, kann damit im Prinzip das atomare Verschiebungsfeld eines Kristalldefektes relativ zum perfekten Gitter direkt zurückgerechnet werden. Ein spezielles inverses Problem für die Elektronenstreuung kann unter diesen Voraussetzungen abgeleitet werden, welches die atomaren Verschiebungen als Nullstellen einer Funktion mit unvollständig bekanntem Fourierspektrum darstellt. Die fundamentalen Beziehungen werden abgeleitet und die Probleme bei der Lösung der unvollständig formulierten inversen Aufgabenstellung diskutiert.

1. Introduction

The imaging of crystal defects by high-resolution transmission electron microscopy or with the electron diffraction contrast technique is well known and routinely used. Though the theoretical image calculations always tend to establish standard rules of interpretation, a direct and phenomenological analysis of electron micrographs is mostly not possible, thus requiring the application of image simulation and matching techniques. Images are modelled by calculating both the interaction process of the electron beam with the almost periodic potential of the matter performed by the multi-slice algorithm or by the direct integration of the basic differential equations and the subsequent Fourier imaging process including the microscope aberrations. The images calculated are fitted to the experiment by varying the defect model and the free parameters. This trial-and-error image matching technique is the solution to the direct scattering problem applied to analyse the defect nature under investigation.

¹⁾ Weinberg 2, D-06120 Halle (Saale), Federal Republic of Germany.

Electron holography or other reconstruction techniques [1, 2] permit the determination of the scattered wave function at the exit surface of the crystal directly out of the hologram or from defocus series up to the microscope information limit owing to the noise in the phase distortion. For instance, the sidebands of a Fourier-transformed hologram represent the Fourier spectrum of the complete complex image wave and its conjugate, respectively, from which the object wave can be reconstructed by separating, centring, and applying the inverse Fourier transform including a reciprocal Scherzer filter, i.e. multiplication with the conjugate complex phase contrast transfer. Thus, both the reconstructed amplitudes and phases can be compared to trial-and-error calculations [3, 4].

The question arises whether it is possible to calculate the object potential or the positions of the atomic scattering centres directly from the wave function reconstructed at the exit surface of the object instead of using trial-and-error simulation techniques [5]. Up to now direct solutions have been given for very thin objects (phase grating approximation [6]) or under the assumption that the crystal potential for the perfect structure is already known and the atomic displacements owing to a crystal lattice defect should solely be determined using the dependence of the three components in the case of plane strains or stresses [7].

The present paper shows that using the knowledge of both the amplitudes and phases of a sufficient number of plane waves scattered by the object as well as the knowledge of the potential of the perfect crystal structure, the displacement field of a crystal lattice defect can in principle be retrieved from the scattered electron wave.

2. The Direct Problem

The HREM image contrast is mainly determined by two processes. First, by the electron diffraction owing to the interaction process of the electron beam with the almost periodic potential of the matter and, second, by the interference of the plane waves leaving the specimen and being transferred by the microscope.

The principles of image formation in the electron microscope have been well established. According to Abbe's theory, the modulus of the interference pattern created by the twofold Fraunhofer diffraction $FT^{-1}\{FT\{o\}e^{D-i\chi}\}$ of the object wave $o(R)$ and modified by the Scherzer phase shifts $\chi(u)$ and the damping envelope $D(u)$ is the so-called high resolution micrograph. The coordinates $R = (x, y)$ and u characterize the real space normal to the beam direction and the reciprocal spatial frequencies, respectively. The point resolution is given by the first zero of the phase contrast transfer, the higher frequencies are transferred with alternating phase shifts and increasing damping and the maximum frequency transferable established the overall information limit of the microscope. More strictly speaking, the instability of the microscope as well as the energy and angular spread of the incoming electrons are described by incoherently summarizing the coherent image intensities. Applying these and including linear effects, the diffractogram $d(u) = \delta(u) * t(u, u') \delta(u')$ is then described by the autocorrelation of the wave spectrum $\delta(u) = FT\{o(R)\}$ weighted with the transmission cross-correlation coefficient $t(u, u')$ [8], the image thus includes nonlinear terms and can no longer be characterized by transfer and damping functions. On the other hand, under the conditions of diffraction contrast the influence of the microscope imaging process itself can be neglected. Thus the image contrast is solely determined by the interaction of the electrons with the object potential.

The interaction of electrons with a crystalline object is described assuming a periodic potential with the electron structure factors as the expansion coefficients and the Bloch-wave

method for solving the high-energy transmission electron diffraction. Different formulations can be given, using Bloch wave or plane wave representations of the scattered waves, applying direct or reciprocal space expansion and using direct integration or slice techniques, which in principle are equivalent descriptions [9]. The object wave in terms of modified plane waves with complex amplitudes φ_g yields

$$o(R) = \sum_g \varphi_g e^{2\pi i((k+g)R + s_g t)} \quad (1)$$

with reflections g , excitations s , wave vector k , and thickness t of a parallel-sided object. The amplitudes φ_g are constant with respect to z in the vacuum outside the object, which means that the plane waves are the stationary solutions of the wave equation. Within the crystal, however, the amplitudes of the modified plane waves φ_g are z -dependent according to Ewald's pendel solution as described by the Bloch waves, which are the stationary solution within the periodic potential. Using furtheron the deformable ion approximation a crystal lattice defect can be included by its elastic displacement field as a phase shift of the Fourier spectrum of the crystal potential. The evaluation of the quantum theoretical scattering problem using the high-energy forward scattering approximation (see, e.g. [10, 11]) yields a parabolic differential equation system for the complex amplitudes of the elastically scattered

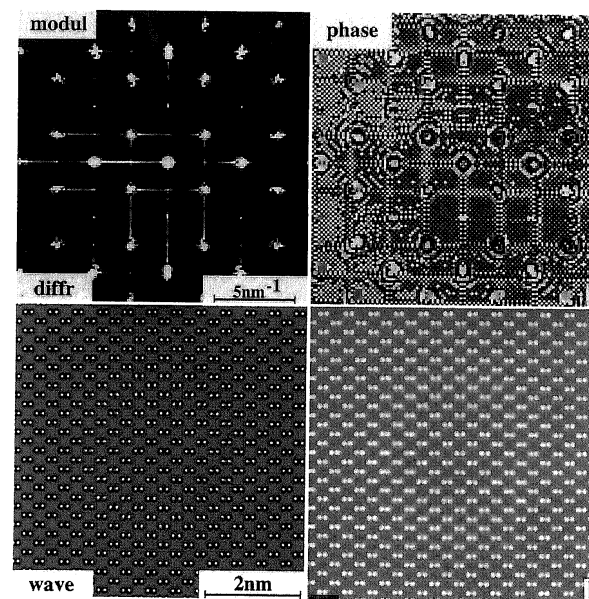
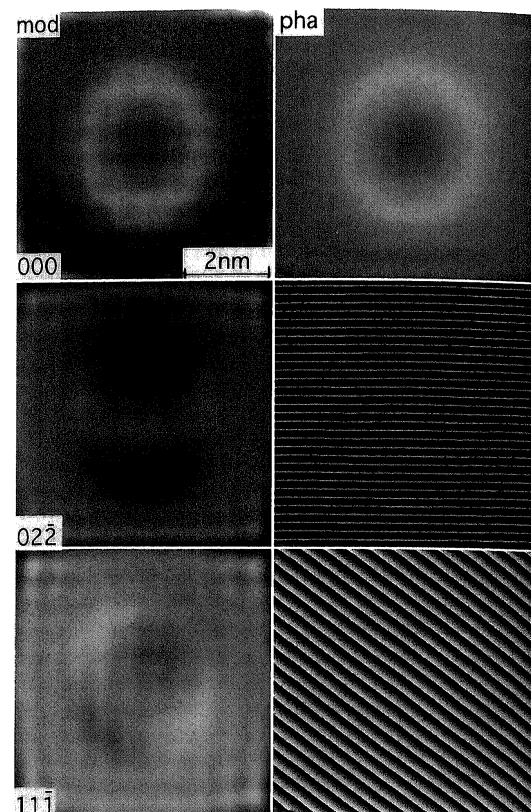


Fig. 1. Simulated modulus (left) and phases (right) of the exit wave function for a spherical inclusion: a) reciprocal-, b) real-space representation ($U = 400$ kV, $\alpha = 20$ nm $^{-1}$, $(12, 8\sqrt{2}, \sqrt{2}/2)$ [011] Si supercell, $t = 8.4$ nm, $R_0 = 1.6$ nm)



electron wave,

$$\frac{\partial \varphi_g}{\partial z} = \{ik_z \nabla^2 - 2(\mathbf{k} + \mathbf{g}) \cdot \nabla\} \frac{\varphi_g}{2k'_z} + i\sigma \sum_h V_{g-h} \varphi_h e^{i\alpha_{gh}}, \quad (2)$$

where $\sigma = 2\pi me/h^2 k_z k'_z$, $\nabla = (\partial/\partial x, \partial/\partial y, 0)$, $k'_z = k_z + g_z + s_g$, and $\alpha_{gh} = 2\pi[(s_h - s_g)z + (\mathbf{g} - \mathbf{h}) \cdot \mathbf{v}(x, y, z)]$ with the elastic displacement field \mathbf{v} and the potential $V = V' + iV''$ including the lattice potential V' and the absorption V'' (one-electron optical potential approximation of inelastic scattering).

Boundary and initial conditions have to be applied additionally: The linearized high-energy approximation fits directly $\varphi_g(\mathbf{R}, t)$ at the crystal exit face to $\varphi_g(\mathbf{R})$ outside, demanding $|\varphi_g(\mathbf{R}, 0)| = \delta_{g0}$ at the entrance face, whereas the continuity of the derivatives has to be omitted in the linearized case. Instead of boundary conditions one can assume a periodic continuation for large extended crystal slabs, i.e. $\varphi_g(x, y, z) = \varphi_g(x + X, y, z)$ and $\varphi_g(x, y, z) = \varphi_g(x, y + Y, z)$, with slab extensions X, Y tending to infinity.

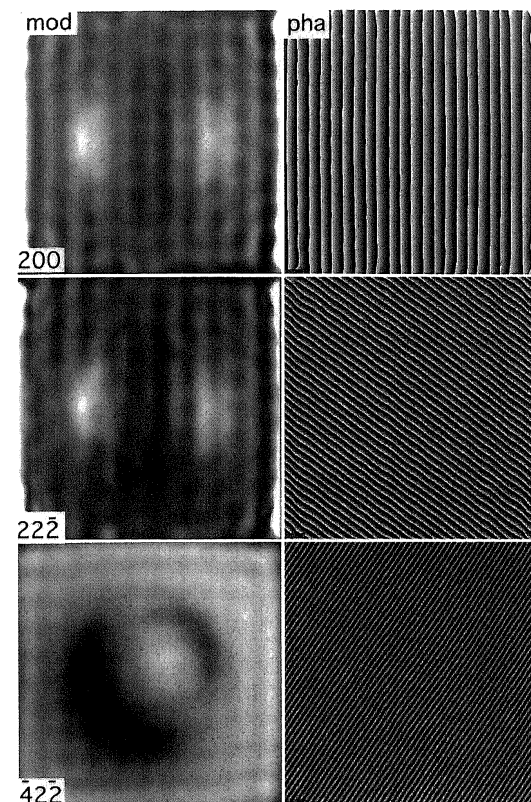


Fig. 2. Modulus (left) and phases (right) of the plane wave amplitudes corresponding to the wave function of Fig. 1 for the transmitted beam (000) and reflections of the type {200}, {022}, {222}, {111}, and {422} from the zero-order Laue zone

Fig. 1 shows the modulus (left) and phases (right) of the exit wave function in reciprocal (Fig. 1a) and real (Fig. 1b) space representation, simulated for a spherical inclusion with the Ashby-Brown displacement field and linear displacements within the defect. Fig. 2 shows the corresponding partial plane wave amplitudes φ_g according to (1) and the initial and boundary conditions. The amplitudes φ_g are used as input for the reconstruction algorithm with the diffusion-like description of the diffraction process (see below). The wave function is calculated using the EMS package for multi-slice simulations [12]; the partial plane waves are generated similar to the reconstruction described later on by applying the IMAGIC image processing software [13] extended by own scripts. The data correspond to a 400 kV microscope and [011] silicon in high-symmetric incidence with a sample thickness of $t = 8.4$ nm. Besides the transmitted beam (000) reflections of the type {200}, {022}, {222},

$\{111\}$, and $\{422\}$ from the zero-order Laue zone are shown in modulus (mod) and phases (pha). The spherically symmetric displacements give rise to asymmetric variations in modulus by analogy with the black-white lobes of precipitates in diffraction contrast images and weak phase shifts in addition to that of $\exp(2\pi i g R)$ characterized by the small fringe bendings.

3. The Wave Reconstruction

Holography with electrons offers one of the possibilities of increasing the resolution by avoiding the microscope aberrations. It also enables the complete complex object wave to be restored. Image plane off-axis holograms are recorded in a microscope which is equipped with a Möllenstedt-type electron biprism inserted between the back focal plane and the intermediate image plane of the objective lens [1, 3]. The object is arranged so that a reference wave outside it is transferred through the microscope, and owing to a positive voltage of the biprism both waves mutually overlap in the image plane creating additional interference fringes. The intensity of the latter is modulated by the modulus of the object wave, whereas their position is varied by the phase of the object wave. Thus the recorded interference pattern is an electron hologram from which both the modulus and the phase of the object wave can be reconstructed by optical diffraction or numerical reconstruction.

A Fourier transform $d_H(u)$ of the intensity distribution of the hologram $i_H(R)$ generates three distinct spectral patterns if the carrier frequency u_c is sufficiently high,

$$d_H(u) = \{\delta(u) + d(u)\} + \{\delta e^{D-ix}\} * \delta(u - u_c) + \{\delta e^{D+ix}\} * \delta(u + u_c). \quad (3)$$

In the central region of the spectrum the zero peak and autocorrelation occur, representing the conventional diffractogram $d(u)$ of the object intensity, completely identical with that obtained from a corresponding HREM micrograph. The sidebands represent the Fourier spectrum of the complete complex image wave and its conjugate, respectively, from which the object wave $o(x, y)$ can thus be reconstructed by separating, centring, and applying the inverse Fourier transform including a reciprocal Scherzer filter, i.e. multiplication by $\exp(-D + ix)$ because of the always linear transfer to the sideband. Nevertheless, the information limit ($D(u) \approx 0$) determines the maximum transferred spatial frequency owing to the noise in the phase distortion.

Fig. 3 shows calculated holograms (hol) assuming a perfect microscope without aberrations (Fig. 3a) and for two different defoci (Fig. 3b: $\Delta = -50$ nm, Fig. 3c: $\Delta = 50$ nm) to demonstrate the reconstruction of the wave function in reciprocal space, i.e. the Fourier transform of the hologram and the selected sideband in intensity (diffraction pattern) and phase are given. Further on the corresponding modulus (mod) and phases (pha) of the complete reconstructed sideband are shown, which should be equivalent to the exit wave function (Fig. 3a) and HREM images (Fig. 3b, c). Fig. 4 shows the modulus (mod) and phases (pha) of the selected particular reflections of type $\{000\}$, $\{200\}$, $\{111\}$, and $\{022\}$, thus giving the reconstruction of the corresponding amplitudes φ_θ out of the holograms. The holograms are generated assuming a reference beam with damping 0.2 and a carrier frequency of 13.2 nm^{-1} (i.e. located approximately at reflection $1/10(-44, 43, -43)$ as in the experimental situation of Fig. 2 in [4]). The reconstruction of the $\{400\}$ reflections is impossible here because of the overlap of the autocorrelation and the sideband. Thus the aperture and the damping are chosen to exclude the $\{113\}$ reflections, which also omits the dumbbells in the HREM reconstruction resulting in differences between the original HREM images and the corresponding reconstructions. The reconstructed amplitudes of the single

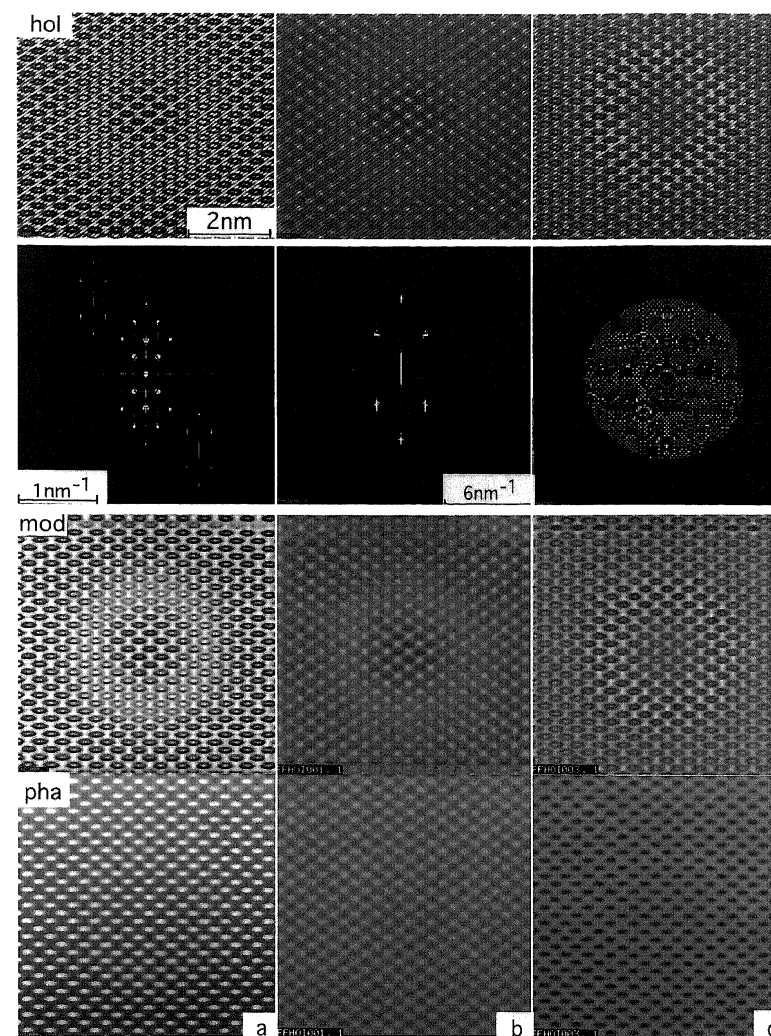


Fig. 3. Calculated holograms (hol) assuming a perfect microscope without aberrations (a) and for two different defoci (b) $\Delta = -50$ nm, (c) $\Delta = 50$ nm, Fourier transform of the hologram and the selected sideband in intensity and phase (diffraction pattern, second row), as well as modulus (mod) and phases (pha) of the complete sideband giving the perfect exit wave function (a) and HREM images (b, c). The hologram is generated assuming a reference beam with damping 0.2 and a carrier frequency of 13.2 nm^{-1} .

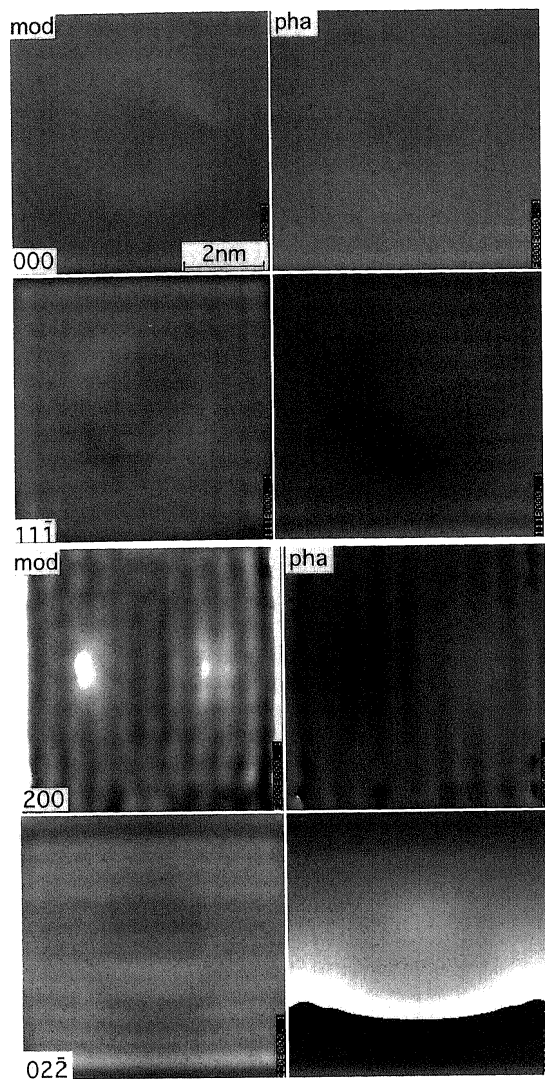


Fig. 4. Modulus (mod) and phases (pha) of the single reflections (000), {200}, {111}, and {022}, i.e. the corresponding plane wave amplitudes, reconstructed from hologram (a) of Fig. 3

reflections can be directly compared with the original simulations in Fig. 1; the moduli show good agreement despite of the resolution reduced; in the phases the term $\exp(2\pi i g R)$ is omitted owing to the centring in the reconstruction.

4. Forward-Backward Iteration

The differential equations (1) allow the diffusion-like interpretation and can be discretized using standard difference algorithms [14], with the help of

$$\begin{aligned} \nabla^2 \varphi &= \frac{1}{\Delta x^2} [\varphi(x + \Delta x, y, z) - 2\varphi(x, y, z) + \varphi(x - \Delta x, y, z)] \\ &\quad + \frac{1}{\Delta y^2} [\varphi(x, y + \Delta y, z) - 2\varphi(x, y, z) + \varphi(x, y - \Delta y, z)], \\ \frac{\partial \varphi}{\partial z} &= \frac{1}{\Delta z} [\varepsilon(\varphi(x, y, z + \Delta z) - \varphi(x, y, z)) + (1 - \varepsilon)(\varphi(x, y, z) - \varphi(x, y, z - \Delta z))] \end{aligned} \quad (4)$$

an algebraic equation system results for the complex amplitudes and the elastic displacements at the (xyz)-grid points (i, j, k) , $(i \pm 1, j, k)$, $(i, j \pm 1, k)$, and $(i, j, k \pm 1)$. Using the abbreviations

$$\begin{aligned} A^\pm &= \pm \frac{\sigma k_z z_0 I^2}{x_0^2 K} + 2\pi(k_x + g_x)(\pm \varepsilon_x - 1/2), \\ B^\pm &= \pm \frac{\sigma k_z z_0 J^2}{y_0^2 K} + 2\pi(k_y + g_y)(\pm \varepsilon_y - 1/2), \\ C &= A^- - A^+ + B^- - B^+ + 1/2 - \varepsilon_z, \\ \alpha(i, j, k) &= 2\pi \left[(s_h - s_g) \frac{kz_0}{K} + (g - h) v(i, j, k) \right], \end{aligned} \quad (5)$$

and denoting by I, J, K the maximum number of grid nodes in x, y , and z direction, respectively, yields

$$\begin{aligned} &(1/2 - \varepsilon_z) \varphi_g(i, j, k - 1) - (1/2 + \varepsilon_z) \varphi_g(i, j, k + 1) \\ &= A^+ \varphi_g(i + 1, j, k) - A^- \varphi_g(i - 1, j, k) + B^+ \varphi_g(i, j + 1, k) \\ &\quad - B^- \varphi_g(i, j - 1, k) + C \varphi_g(i, j, k) - i\sigma(z_0/K) \sum_h V_{g-h} \varphi_h(i, j, k) e^{i\alpha(i, j, k)} \end{aligned} \quad (6)$$

which is equivalent to forward $(k + 1)$ and backward $(k - 1)$ integration with respect to the beam propagation, i.e. using $\varepsilon_z = 1/2$ or $-1/2$, respectively.

The periodic boundary conditions and the initial conditions can simply be written $\varphi_g(i, j, k) = \varphi_g(i + I, j, k)$, $\varphi_g(i, j, k) = \varphi_g(i, j + J, k)$, and $|\varphi_g(i, j, 0)| = \delta_{0g}$, $\varphi_g(i, j, t) = F_g(i, j)$, respectively, where F_g is known from the wave reconstruction for a certain number of reflections.

The difference equations (6) are equivalent for backward $(k - 1)$ and forward $(k + 1)$ integration, thus being insufficient for determining both the wave amplitudes $\varphi(i, j, k)$ and

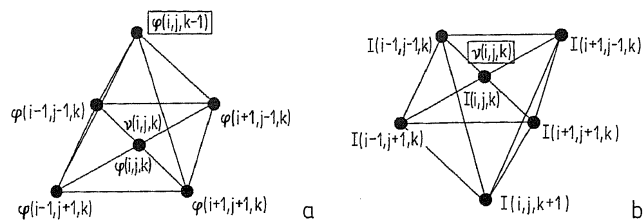


Fig. 5. Relations of the variables in a) forward ($\epsilon = -1/2$) and b) backward ($\epsilon = 1/2$) integration

the elastic displacement field $v(i, j, k)$ at the grid points (i, j, k) considered. This can be seen also by simply numbering the unknowns and the equations at each node (see Fig. 5): for N beams, there are N unknown amplitudes and three unknown displacements and N relations according (6) using either $(k-1)$ or $(k+1)$. One of the difference equations, however, can be replaced as follows: While the optical potential in reciprocal space representation is generally non-hermitian, the hermiticity of the potential V' and of the "absorption" V'' yields the equation of continuity for the whole current $I = \varphi_g \varphi_g^*$. The continuity equation can be denoted as

$$\frac{\partial I}{\partial z} = 9I - 2 \sum_g V''_{gh} \varphi_g \varphi_g^* e^{izg_h} \quad (7)$$

with the abbreviation $9I = \sum_g [k_z(\varphi_g \nabla^2 \varphi_g^* - \varphi_g^* \nabla^2 \varphi_g) + 2(\mathbf{k} + \mathbf{g}) \nabla(\varphi_g \varphi_g^*)]/k'_z$.

The equation of continuity can be discretized by analogy with the discretization of the differential equations above, the differential operator 9 , however, yields mixed terms with respect to different nodes (i, j, k) and $(i \pm 1, j \pm 1, k)$. By analogy with the Gelfand-Levitan algorithm (see, e.g., [15]) an additional equation results, which is a kind of completeness relation, yielding

$$\sum_g Q_g e^{2\pi i g v} = 0. \quad (8)$$

The coefficients $Q_g = \sum_h k_z/k'_z V''_{gh} \varphi_g^* \varphi_h \exp[2\pi i(s_g - s_{g-h})z]$ for $\mathbf{g} \neq (000)$ and the corresponding $Q_{000} = 9I$ are given at the nodes (i, j, k) from (8) in forward scattering. Equation (8) can replace one of equations (6) in backward integration enabling the determination of the displacements v at (i, j, k) by inverting the equation of continuity as an independent additional equation. Thus, in principle, the retrieval of the displacement is given by the remaining inverse problem (8), which is equal to find the root of a function given by an incomplete Fourier transform.

At the exit surface a further equation is given applying the forward integration outside the crystal to determine $\varphi(i, j, k+1)$ from $\varphi(i, j, k)$ where the potential is assumed to be vanishing because of the vacuum propagation. The backward integration, however, using (8) then enables the determination of $v(i, j, k)$ at the exit surface.

5. Conclusions: A Particular Inverse Problem

The inverse problem (8) is ill-posed for two reasons: Only one equation has to be solved for the vectorial root $v(i, j, k)$ at node (i, j, k) , and the spectrum $Q_g(i, j, k)$ is incomplete

and noisy. This results in unstable numerical solutions using standard algorithms to find the roots, owing to the existence of a large number of subsidiary roots. Different algorithms are tested, viz. the Newton-Raphson algorithm itself to solve (8), and to transform (8) in an iterative form as a kind of quasi-regularization, e.g. using relations for the arguments yielding

$$v_x^{n+1} = \frac{1}{2} \pi \{ \arg [Q \exp(2\pi i v_x^n) + \sum_g Q_g \exp(2\pi i g v^n)] - \arg [Q] \} \quad (9)$$

and similar for v_y, v_z . Both algorithms demand the iteration for linearly independent coefficients g , thus coplanar vectors g leave one component unconsidered.

Analytical solutions of (8) can be performed if four terms at maximum are considered. For, e.g. non-vanishing $Q_{000}, Q_{\lambda 00}, Q_{0\mu 0}$, and if the other coefficients can be neglected, one obtains $v_x = 1/2\pi\lambda\{\varphi_{000} - \varphi_{\lambda 00} + \arccos[(b_{\lambda 00}^2 - b_{0\mu 0}^2 - b_{000}^2)/2b_{000}b_{\lambda 00}]\}$ and similar for v_y , whereas v_z is arbitrary. Here, φ and b are the arguments and amounts, respectively, of Q . The solution for four non-vanishing terms has in principal the same structure, allows to determine the component v_z , too, and has to fulfil further restrictions.

The retrieval of the atomic displacements from a reconstructed electron wave function at the exit of an object results in the particular inverse problem (8) with the difficulties of finding the roots as discussed. Thus, from the mathematical point of view the retrieval procedure is an ill-posed inverse problem requiring additional information about the unknown reconstructed displacements in order to make the process stable and continuous, to avoid singularities, and to restrict the manifold set of solutions possible. Open questions arise, e.g., owing to the assumptions of cyclic boundary conditions, the applicability of the completeness relation to the backward iteration and for depths, where the equations for the displacement retrieval cannot be inverted because of singular coefficients. The procedure described has transformed these difficulties to the mathematical problem of determining the roots of a function with incomplete Fourier transform. Nevertheless, the uniqueness and the stability of the solutions are determined by the spectrum Q_g given by the reconstructed amplitudes, this should be discussed in more detail in a forthcoming paper.

Acknowledgement

We are grateful to the Volkswagenstiftung for financial support.

References

- [1] H. LICHTE, *Ultramicroscopy* **20**, 293 (1986).
- [2] W. COENE, G. JANSSEN, M. OP DE BEECK, and D. VAN DYCK, *Phys. Rev. Letters* **69**, 3743 (1992).
- [3] H. LICHTE, *Adv. Electron Microscopy* **12**, 25 (1991).
- [4] H. LICHTE, E. VÖLKL, and K. SCHEERSCHMIDT, *Ultramicroscopy* **47**, 231 (1992).
- [5] K. SCHEERSCHMIDT and R. HILLEBRAND, *Proc. 32nd Course Internat. Centre of Electron Microscopy, High-Resolution Electron Microscopy — Fundamentals and Applications, Halle 1991* (p. 56).
- [6] D. VAN DYCK and W. COENE, *Scanning Microscopy, Suppl.* **2**, 131 (1988).
- [7] A. K. HEAD, *Austral. J. Phys.* **22**, 43 (1969).
- [8] K. ISHIZUKA, *Ultramicroscopy* **5**, 55 (1980).