Scanning Microscopy Vol. 11, 1997 (Pages 455-465) 0891-7035/97$5.00+.25
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

DIRECT RETRIEVAL OF OBJECT INFORMATION FROM DIFFRACTED ELECTRON WAVES

Kurt Scheerschmidt*

Max Planck Institute of Microstructure Physics, Halle, Germany

Abstract

Electron holography and other wave reconstruction techniques allow one to directly determine 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 either a linearized eigenvalue system or a discretized form of the diffraction equations, the scattering problem can be inverted. This, in principle, enables the direct retrieval of the local thickness and orientation as well as the refinement of potential coefficients or the determination of the atomic displacements, caused by a crystal lattice defect, relative to the atom positions of the perfect lattice. Two special inverse problems as direct solutions of the electron scattering equations can be deduced considering the sample orientation as perturbation, or solely the atomic displacements, which are given by the zeros of a function with an incompletely known Fourier spectrum. The numerical algorithms resulting from the fundamental relations imply ill-posed inverse problems.

Key Words: Electron crystallography, high resolution electron microscopy, electron holography, object retrieval, inverse solutions, improperly posed problems.

*Address for correspondence:
Kurt Scheerschmidt
Max Planck Institute of Microstructure Physics
Weinberg 2, D-06120 Halle, Germany
Telephone number: +49-345-5582910
FAX number: +49-345-5511223
E-mail: schee@mpi-halle.de

Introduction

As often occurring in many physical investigations, in the mathematical sense, the direct solution of the diffraction equations implies an inverse problem. Such inverse problems are difficult, always fascinating, and in most cases ill- or improperly posed [2, 8, 22]. Ill- or improperly posed means that one or all of the requirements are violated usually characterizing physics, i.e. the existence, uniqueness and stability of a solution. Although inverse problems violate especially the existence of unique and continuous solutions for arbitrary data, they are of great practical importance if the trial-and-error solution demands a large variety of possible solutions and models to be tested, as they provide a better insight into the basic relations of the physical phenomena.

For instance, the imaging of crystal defects by either high-resolution transmission electron microscopy or with the help of electron diffraction contrast technique is well known and routinely used. Though theoretical image calculations 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 an almost periodic potential of matter 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 indirect solution to the direct scattering problem, used to analyse the defect nature under investigation.

Electron holography and other reconstruction techniques [3, 9, 10, 11, 27] permit the determination of the scattered wave function at the exit surface of the crystal either directly from the hologram or from defocus series, up to the information limit. For example the side-bands of a Fourier-transformed hologram represent the Fourier spectrum and its conjugate of the complete complex image wave, from which the object wave can be reconstructed by including a reciprocal Scherzer filter in the reconstruction. Thus, both the reconstructed amplitudes and phases can

be compared to trial-and-error calculations [10, 12].

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 at the exit surface of the object instead of using trial-and-error simulation techniques [5, 6, 17, 18, 19, 24]. If all atom positions are assumed to be unknown, e.g., for an amorphous object with random atom positions, the argument of "blocked information channel" [24, 25] does not allow to determine more than one atom per "resolution area". Inverse problems, generally dealing with insufficient measured data, require always a priori physical related information. If the assumption is made that the object is almost a perfect crystal, the inversion problem in general refers to the finding directly the potential from the wavefunction [4, 21], which is generally unsolved, too. Uniqueness is guaranteed if a certain quasi-periodicity of the thickness can be made unambiguous by measuring at different thicknesses. But one can restrict further the ambiguities by increasing the a priori information, e.g., assuming that solely the displacement field of a defect has to be determined [5, 6]. The present paper shows that the knowledge of both the amplitudes and phases of a sufficiently large number of plane waves scattered by the object as well as the partial knowledge of the potential of the perfect crystal structure imply the possibility of directly retrieving object information, instead of using trial-and-error simulation techniques. Two approximations are discussed to solve the resulting inverse scattering problem without reconstructing the whole crystal potential:

First, the special problem of retrieving the local sample orientation is solved on the basis of the perturbation approximation for perfect crystals, and by applying regularized and generalized matrices to invert the resulting linearized problem. The corresponding iteration procedure enables the direct analysis of the moduli and phases if a sufficient number of plane wave amplitudes can be separated yielding local thickness and bending of the object for each image pixel [20].

Second, based on the knowledge of the reconstructed complex electron wave and using a discretized form of the diffraction equations, an alternative method is developed [18, 19], yielding an algebraic equation system for the complex amplitudes and the elastic displacements. In principle, this system enables the direct retrieval of the atomic displacements, caused by a crystal lattice defect, relative to the atom positions of the perfect lattice. The equations are invertible provided the completeness of the plane waves is valid (continuity of the electron current). A special inverse problem of electron scattering is deduced considering solely those atomic displacements given by the zeros of a function with an incompletely known Fourier spectrum from the scattered electron wave of which the displacement field of a crystal lattice defect can, in principle, be retrieved.

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 matter and, second, by the interference of the plane waves leaving the specimen and being transmitted by the microscope. Assuming that the object wave is reconstructed free of aberrations or under diffraction contrast conditions 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 on the basis of a periodic potential with the electron structure factors as expansion coefficients and the Bloch-wave method for solving the high-energy transmision 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 direct integration or slice techniques, which, in principle, are equivalent descriptions [23]. The object wave in terms of modified plane waves with complex amplitudes ϕ_g yields

$$o(\mathbf{R})=\Sigma_g \phi_g \exp 2\pi i((\mathbf{k}+\mathbf{g})\mathbf{R}+s_g t) \quad (1)$$

with reflections $\mathbf{g}$, excitations s_g, wave vector $\mathbf{k}$, 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 the Ewald pendulum solution as described by the Bloch waves, which are the stationary solution within the periodic potential.

The basic equations of the Bloch wave presentation in forward scattering approximation are given by the eigenvalue system

$$\Sigma A_{gh} C_h - \gamma C_g = 0, \text{ with}$$

$$2k_z A_{gh} = (\mathbf{2K.g}\text{-}g^2)\, \delta_{gh} - V_{g\text{-}h} \quad (2)$$

yielding the amplitudes $C_{\mathbf{g}}^{(l)}$ of the l th partial wave and its "anpassung" $\gamma^{(l)}$ to the dispersion of the lattice as a function of the lattice potential (Fourier coefficient $V_{\mathbf{g}}$) as well as the relative orientation of the object with respect to the electron beam incidence **K.** With these eigenvalues and vectors, for a plane parallel perfect crystal of thickness t the complex amplitudes ϕ_g of Equation (1) are directly given by

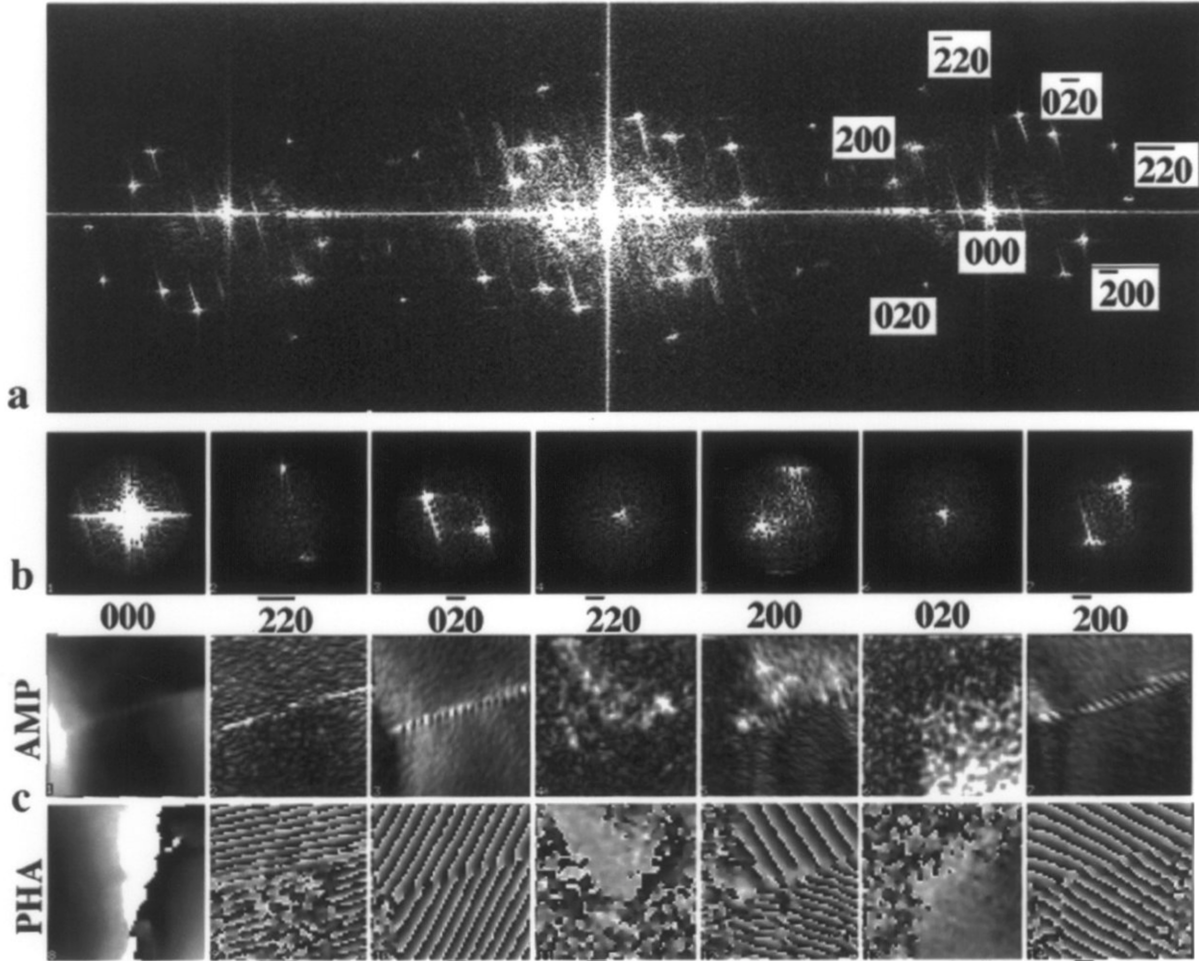

Figure 1. Reconstruction of pairs of reflections of a Σ13-(100) Au grain boundary: (**a**) Fourier spectrum of the hologram (0.05 nm fringes, A. Orchowski, University of Tübingen [18, 19]) with indices of the reflections and asymmetric intensities in the sideband showing the mistilted orientation, (**b**) pairs of reflections of the sideband with Gaussian window filtering, (**c**) reconstructed moduli (AMP) and phases (PHA) of the reflections.

$$\phi_g(t) = \Sigma\, (C^{-1})_{\mathbf{o}}^{(1)} C_g^{(1)} \exp\,(2\pi i\gamma^{(1)}t) \quad (3)$$

or, in matrix form, with diagonal $X = \{\exp\,(2\pi i\gamma^{(l)}t)\}$ representing the diagonalized scattering matrix $\exp(2\pi i\mathbf{A}t)$ as follows:

$$\mathbf{\Phi = CXC^{-1}\theta} \quad (4)$$

where $\Phi=[\phi_{\mathbf{g}}]$ and θ are the vectors of the amplitudes of the exit and the incident waves, respectively.

Using furthermore the deformable ion approximation a crystal lattice defect can be included by its elastic displacement field **v** 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., [1, 7] for the derivation and the explicit form of the equations) yields a parabolic differential equation system for vector Φ of the complex amplitudes of the elastically scattered electron waves:

$$\partial\Phi/\partial z = (\Delta + V[e^{i\mathbf{gv}}])\,\Phi \quad (5)$$

with $\Delta = \{ik_z\nabla^2\text{-}2(\mathbf{k}+\mathbf{g})\nabla\}\Phi/2k'_z+2\pi(s_h\text{-}s_g)z$, $\nabla=(\partial/\partial x,\partial/\partial y,0)$, $k'_z=k_z+g_z+s_g$ and the potential V=V'+iV" including the lattice potential V' and the absorption V" (one electron-optical potential approximation of inelastic scattering). The abbreviation $[e^{i\mathbf{gv}}]$ denotes the diagonal term in the matrix of the Fourier coefficients according to the defect phase term.

In addition, boundary and initial conditions have to be applied: The linearized high-energy approximation directly fits $\phi_g(\mathbf{R},t)$ at the crystal exit surface to $\phi_g(\mathbf{R})$ outside, demanding $|\phi_g(\mathbf{R},0)|=\delta_{g0}$ at the entrance surface, whereas

the continuity of the derivatives has to be omitted in the linearized case. It enables one, however, to estimate the unknown displacements at the exit foil surface by using Equation (5) without potential outside and inverting Equation (5) directly at the exit surface:

$$\{V[e^{igv}]\ \Phi = 2\ \Delta\Phi\}_{z=t} \qquad (6)$$

Instead of boundary conditions one can assume a periodic continuation to describe large extended crystal slabs, i.e., $\phi_g(x,y,z) = \phi_g(x+X,y,z)$ and $\phi_g(x,y,z) = \phi_g(x,y+Y,z)$, with slab extensions X,Y approaching infinity.

The Wave Reconstruction

Holography with electrons offers one of the possibilities to increase the resolution limit by correcting for 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, e.g., the back focal plane and the intermediate image plane of the objective lens [9, 10, 11, 12]. The object is arranged so that a reference wave outside of 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 the fringe 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 $d_H(\mathbf{u})$ of the intensity distribution of the hologram generates three distinct spectral patterns if the carrier frequency $\mathbf{u}_c$ is sufficiently high:

$$d_H(\mathbf{u}) = \{\delta(\mathbf{u})+d(\mathbf{u})\} + \{\hat{o}\cdot e^{D-i\chi}\}*\delta(\mathbf{u}-\mathbf{u}_c) + \{\hat{o}^*\cdot e^{D+i\chi}\}*\delta(\mathbf{u}+\mathbf{u}_c) \qquad (7)$$

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 means of an inverse Fourier transform and using a reciprocal Scherzer filter $e^{D-i\chi}$ with damping D and aberrations χ [10, 16]. Nevertheless, the information limit ($D(\mathbf{u})\approx 0$) determines the maximum transferred spatial frequency owing to the noise in the phase distortion. Examples are discussed in [14], where the reconstruction of holograms of a Σ=13 (100) tilt grain boundary in gold demonstrates the aberration free wave reconstruction and yields amplitudes and phases of the corresponding HREM image. A comparison with respective simulations enables one to analyse the orientation of the grains and to determine the particular atomic displacements at the grain boundary. Furthermore, microdiffraction patterns from the holographic reconstructed waves and respective defocus series can be studied to investigate the local deviations from the perfect lattice structure.

In the following it is important to note that, besides the whole sideband, each single reflection of sufficient intensity can be reconstructed separately. This provides the possibility of noise reduction if suitable windows and filtering are applied and if the pixels are precisely cen-tred to avoid additional phase shifts. The environment of the reflections included in the filtering process has to be chosen such that the information of local distortions folded with the reflections will be transferred to the reconstructed partial waves. The reconstruction of the single reflections causes modulus and phase to be dis-tributed in the partial waves, which is the presupposition of the inversion algorithm discussed in the following. Figure 1 shows the Fourier spectrum of the hologram (upper row) of a Σ = 13 (100) tilt grain boundary in gold (θ=22.6°, see [14] and preliminary joint work [15]). The second row of Figure 1 presents the intensities of the pairs of reflections as indicated in the spectrum of the hologram filtered through a Gaussian mask. The third row shows modulus (mod) and phases (pha) of the particular reflections chosen of types 000, {002}, and {220}, thus presenting the reconstruction of the cor-responding amplitudes ϕ_g out of the hologram. The re-construction of the higher-order reflections is impossible here because of the lower intensity of the latter and the mutual overlap of the autocorrelation and the side-band. The reconstructed amplitudes of the reflections can directly be interpreted as bright and dark field images of the grain boundary. The fringes remaining in the phases are due to the term $e^{2\pi i\Delta\mathbf{g}\mathbf{R}}$ according to the tilt $\Delta\mathbf{g}$ of the reflection pairs of the different grains. Figure 2 presents the reconstruction in more detail and using single reflections filtered, i. e., as denoted by 1 and 2 the same reflections are considered as before but separated now according to grain 1 and 2, respectively. The additional fringes in the phases are thus omitted, as can be seen more clearly in the presentation of the real and imaginary part of the wave functions in the corresponding lower rows. The shift of the fringes at the grain boundary directly indicates the phase shift owing to the crystal defect. The modulation in the lower frequency range is either due to the local bending of the sample or to thickness oscillations.

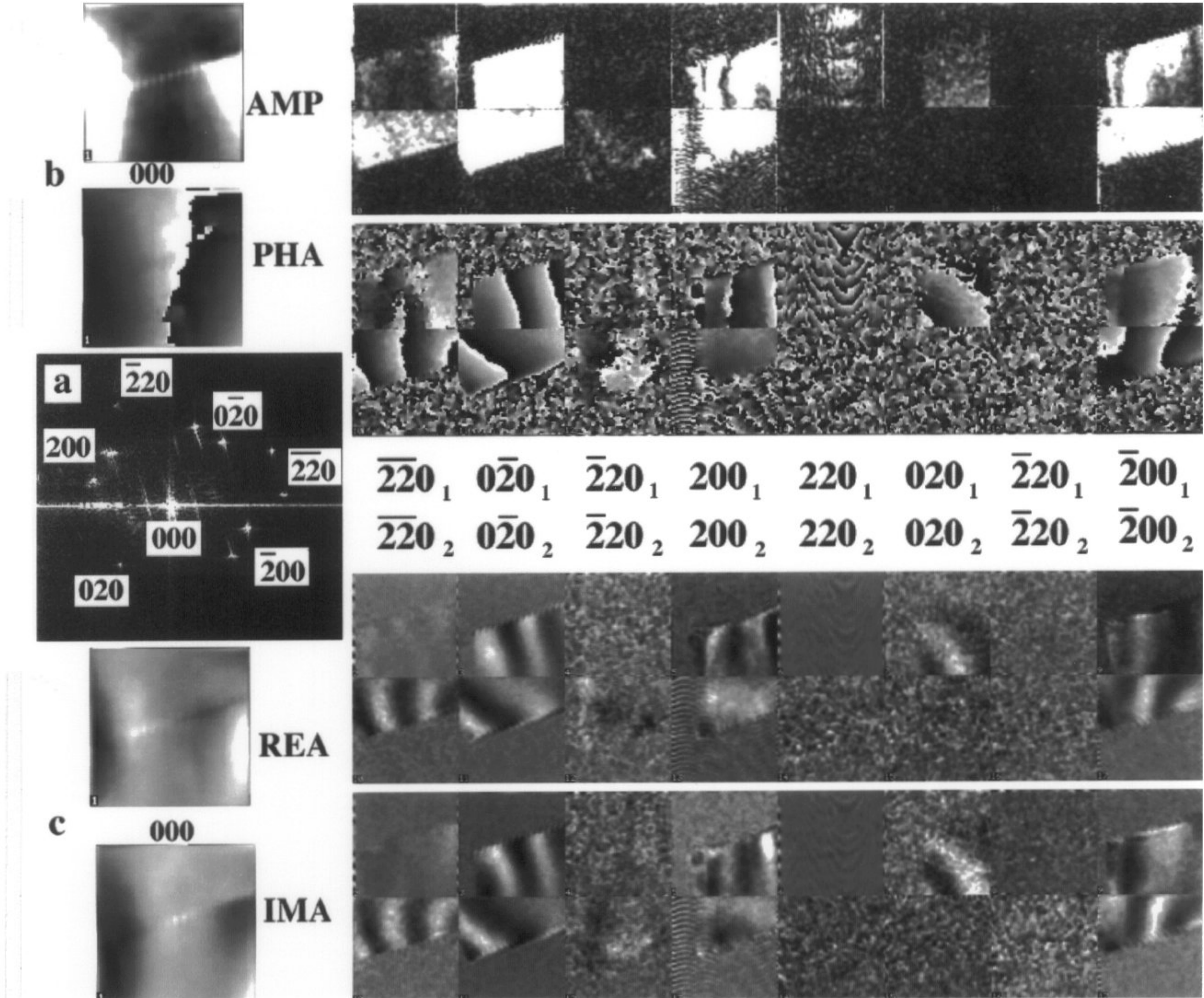

Figure 2. Reconstruction of single reflections of a Σ13-(100) Au grain boundary separately for both grains denoted by 1 and 2, respectively: (**a**) Fourier spectrum of the hologram (0.05 nm fringes, A. Orchowski, University of Tübingen [18, 19]), (**b**) reconstructed moduli (AMP) and phases (PHA), (**c**) reconstructed real (REA) and imaginary (IMA) part of the partial waves.

Inversion by Linearization of the Ewald Pendulum Solution

Equation (2) can be linearized applying perturbation methods. Assuming that the eigenvalues γ are non-degenerated, and by analogy with Equation (4), the perturbation solution may read

$$\phi = \Gamma\, \Xi\, \Gamma^{-1}\, \theta \qquad (8)$$

where the matrices are given by

$$\Gamma = \mathbf{C}\,(1+\Delta),\ \Xi = \{\exp(2\pi i \lambda t)\},\ \text{and} \qquad (9)$$
$$\lambda=\gamma+\Delta\{\delta_{ij}\}+\Delta^{-1}\{1/(\gamma_i-\gamma_j)\}\ \Delta$$

As diagonal elements the perturbation matrix $\Delta_{gh} = (\Delta\mathbf{K}.\mathbf{g})\{\delta_{gh}\}+i\Delta V_{gh}$ contains the deviation of the orien-ta-tion $\Delta\mathbf{K}$ from that of the original eigenvalue system $\mathbf{K}$. The non-diagonal elements describe a perturbation of the potential as, e.g., according to optical absorption. Figure 3 demonstrates the validity of the perturbation solution comparing Equation (8) with the exact solution (4) of the two-beam case, which is solely chosen for the sake of simplicity and because an explicite solution exists for this

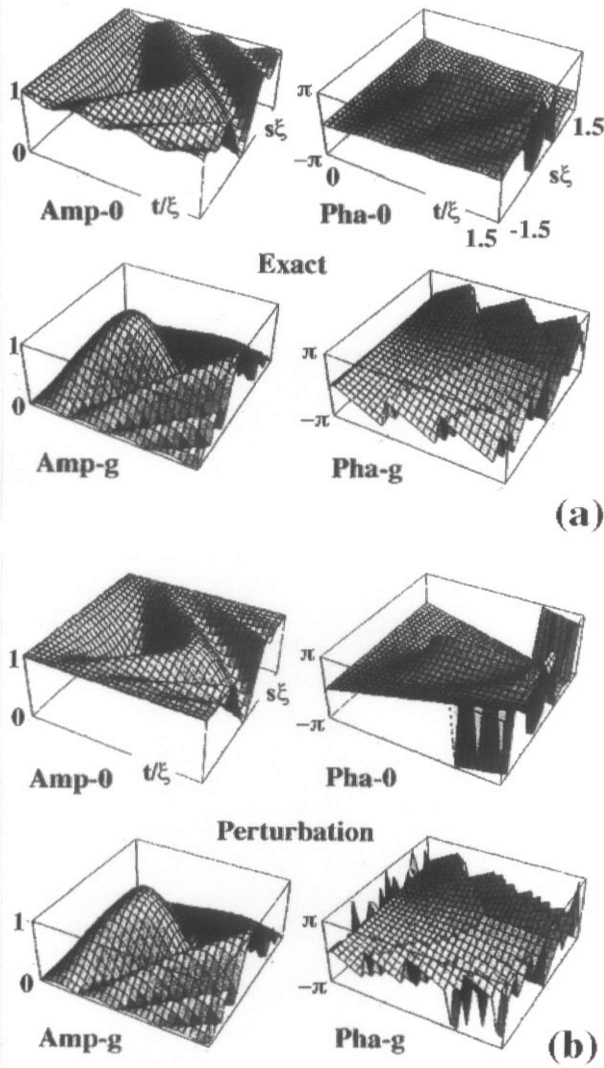

Figure 3. Comparison of the exact two-beam solution (**a**) of moduli (Amp) and phases (Pha) of transmitted (0) and diffracted beam (**g**) with the corresponding perturbation solution (**b**). Differences occur for orientations with $|s\xi| \gg 1$, where the perturbation is no longer valid (s = Bragg deviation, t = crystal thickness, ξ = extinction distance).

situation. As for moduli and phases, for both reflections there are remarkable deviations for $|s\xi| > 1$ almost independent of thickness t around the orientation of the pole $s\xi=0$ of the exact two-beam excitation.

Starting from approximate values of thickness t_o and beam orientation (k_{xo},k_{yo}) gained either from a priori knowledge or by analysing, e.g., the asymmetry of the single reflections reconstructed from the holographically retrieved wave function, the perturbation solution is valid within certain intervals around t_o and (k_{xo},k_{yo}). Equation (8) can be expanded in a Taylor series yielding

$$\phi(t,k_x,k_y) = \phi(t_o,k_{xo},k_{yo}) + (t-t_o)\delta\phi/\delta t + (k_x-k_{xo},k_y-k_{yo})\,\mathrm{grad}_k\phi \quad (10)$$

The derivatives can directly be gained from Equations (9) using equivalent abbreviations:

$$\delta\phi/\delta t = \Gamma\cdot\delta\Xi/\delta t\cdot\Gamma^{-1}\theta \text{ and}$$
$$\nabla_k\phi = (\nabla_k\Gamma\Xi - \Gamma^{-1}\nabla_k\Gamma\Xi + \Gamma\nabla_k\Xi)\Gamma^{-1}\theta \quad (11)$$

The linearized Equation (10) together with the analytical expressions (11) enable the inverse solution:

$$(t,k_x,k_y) = \mathbf{M}_{inv}\cdot[\phi^{exp}-\phi^{pert}] \quad (12)$$

where the inverse matrix generalized and regularized by, e.g., the Penrose-Moore generalization $\mathbf{M}_{inv} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T$, is also given analytically using the matrix of the coefficients $\mathbf{M} = (\delta\Phi/\delta t,\mathrm{grad}_k\Phi)$ of Equation (10). The series expansion (10) as well as the resulting formalism (12) can be extended to include also the derivatives of deviations from potential coefficients, which are omitted here for the sake of simplicity. That means, additional unknown object parameters can be included in the retrieval procedure as far as the problem remains overdetermined with respect to the unknowns. The possibility of the inversion proposed here is based on linearization and the fact that the problem is overdetermined with respect to the unknowns but underdetermined if the noise is included, resulting in a least square minimization of a suitable vector norm of the defect $||\Phi^{exp}-\Phi^{pert}|| = \mathrm{Min}$ [2, 13]. The stability of the procedure may be enhanced by using more general regularizations as, e.g., the Tichonov- Philips regularization: While the Moore-Penrose inverse minimizes the defect, an additional constraint allows one to weight the measured data and to smooth the solution yielding, e.g., $\mathbf{M}_{inv} = (\mathbf{M}^T\mathbf{C}_1\mathbf{M}+\gamma\mathbf{C}_2)^{-1}\mathbf{M}^T$ with the suitable regularization factor γ and matrices $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively (see, e.g., [2, 13] and a forthcoming paper studying the numerical stability of the procedure).

The algorithm (12) is the main result of the present work, because the regularized inverse iteration can directly be applied to each pixel in the real space representation of the single reflections reconstructed from the hologram. On the assumption of the basic eigenvalue system (8) and starting with suitable local thickness t_o as well as incident beam orientation (k_{xo},k_{yo}) the values of thickness t and orientation (k_x,k_y) are probably enhanced if Equation (12) is applied to the measured amplitudes and phases of each image pixel and each reflection **g**. Figures 4 and 5 demonstrate the applicability using pairs of reflections from the wave reconstruction of Figure 1. In both cases the same

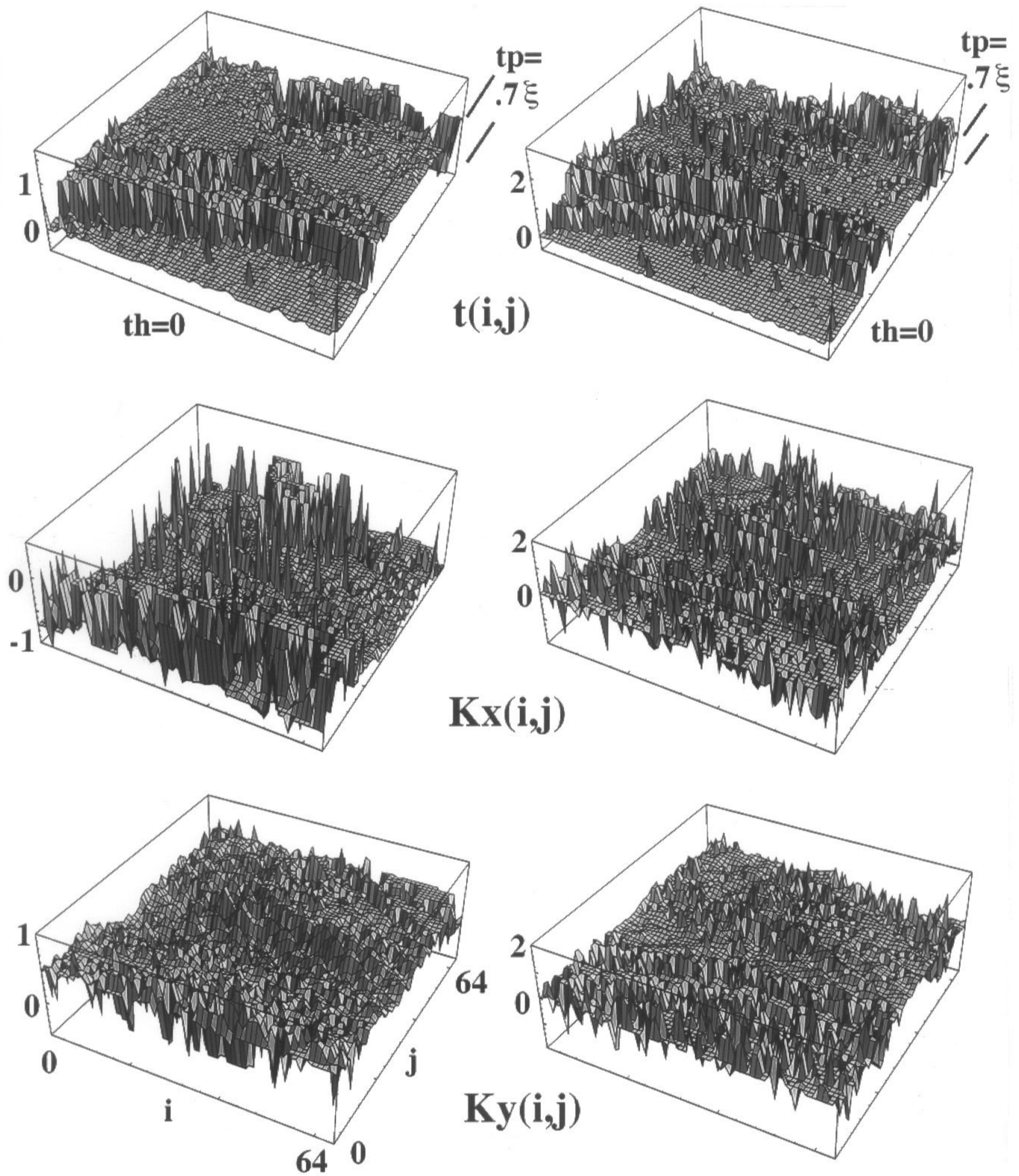

Figure 4. Non-stabilized iteratively determined local sample thickness t and beam orientation (K_x,K_y) retrieved from the reconstructed pairs of reflections of Figure 1 for arbitrary start values of thickness t (resulting in stable solutions th=0 in the hole, and tp = 0.77ξ on the plateau) and given start values of orientation **K**=(0.51, 0.71, 0) and **K**=(-0.28, 1.21, 0) for the left (**a**) and right (**b**) columns, respectively.

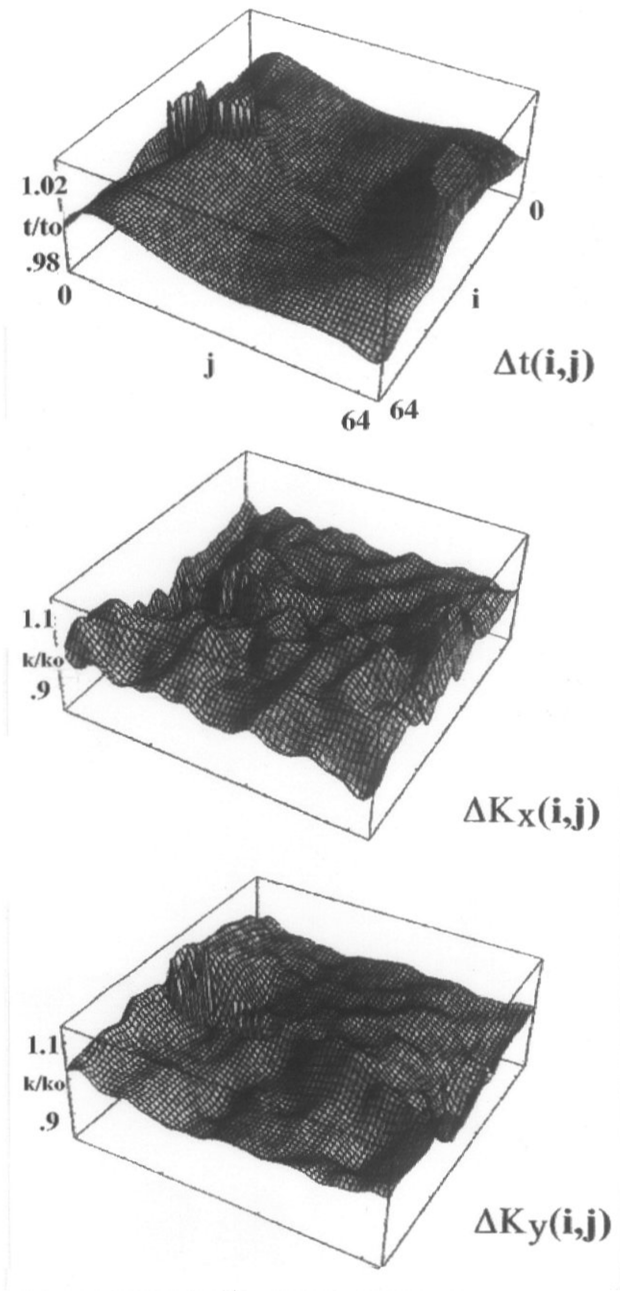

Figure 5. Local thickness variations Δt and distribution of the local sample bendings (ΔK_x,ΔK_y) reconstructed for Figure 3 after reduction of the noise oscillations owing to thresholds $0.98t_o<\Delta t<1.02t_o$ and $0.9k_o<\Delta k<1.1k_o$.

---

seven-beam eigenvalue system was used to model the diffraction behaviour. Here no further assumption was made for the initial thickness t_o, the best fit is found searching in an extended thickness intervall for the absolute minimum of the defect of the vector norm. Different initial orientations (Fig. 4a: k_0=(0.51, 0.71, 0.0), Fig. 4b: k_0=(-0.28, 1.21, 0.0)) yield very noisy results in thickness t and orientation (k_x,k_y) for the 64x64 pixels retrieved. Nevertheless, both cases show almost the same values $tp \approx 0.77\,\xi$ and $th = 0$ for the plateau of the object and the hole, respectively. As the iteration procedure seems to be amplifying the noise, the regularizations should further be enhanced. Figure 5 shows a mean value of the retrieved differences in thickness and orientation with values larger than a certain threshold omitted.

Inversion of the Discretized Diffraction Equations

The differential equations (5) allow a diffusion-like interpretation and can be discretized using standard difference algorithms [18, 19]. An algebraic equation system results, which formally reads

$$\Phi(i,j,k\text{-}1) = \mathbf{F}_1\,\{\Phi(i,j,k),\Phi(i\pm1,j,k),\Phi(i,j\pm1,k),v(i,j,k)\} \qquad (13)$$

for the complex amplitudes Φ and the elastic displacements $\mathbf{v}$ at the (xyz)-grid points (i,j,k), (i±1,j,k), (i,j±1,k) and (i,j,k±1) by which the whole object is represented.

With I,J,K denoting the maximum number of grid nodes in x, y, and z direction, respectively, the periodic boundary conditions and the initial conditions may simply read $\Phi(i,j,k) = \Phi(i+I,j,k)$, $\Phi(i,j,k) = \Phi(i,j+J,k)$ and $|\phi_g(i,j,0)|=\delta_{0g}$, $\phi_g(i,j,K) = F_g(i,j)$, respectively, with F_g derived from the wave reconstruction for a certain number of reflections. At the exit surface, an additional equation is given by applying the forward integration of Equation (13) outside the crystal to determine $\phi(i,j,K+1)$ from $\phi(i,j,K)$, where the potential is assumed to be vanishing because of the vacuum propagation. The backward integration, however, using again Equation (13) then enables the determination of $\mathbf{v}(i,j,K)$ at the exit surface. This is the numerical realization of the symbolic Equation (6), corresponding to the additional assumption of the continuity of the derivatives for the linearized equations, too.

Within the crystal the difference Equations (13) are equivalent for backward (k-1) and forward (k+1) integration with respect to the beam propagation, thus being insufficient for determining both the wave amplitudes $\Phi(i,j,k)$ and the elastic displacement field $\mathbf{v}(i,j,k)$ at the grid points (i,j,k). This becomes also obvious by simply numbering the unknowns and the equations at each node: for N beams, there are N unknown amplitudes and 3 unknown displacements, and N relations according to Equation (13), using either (k-1) or (k+1). One of the difference equations, however, can be replaced as follows: While the optical potential in the 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 $I= \Sigma\, \phi_g \phi^*_g$. The continuity equation may then

read

$$\partial I/\partial z = \Phi\nabla^2\Phi^* - \Phi^*\nabla^2\Phi + 2(\mathbf{k}+\mathbf{g})\nabla I - 2\Phi V'[e^{igv}]\mathbf{\Phi}^* \quad (14)$$

The equation of continuity can be discretized by analogy with the discretization of the differential equations above. The differential operator, however, yields mixed terms between adjacent nodes (i,j,k) and (i±1,j±1,k):

$$F_2\{\mathbf{v}(i,j),\Phi(i,j,k+1),\Phi(i,j,k), \Phi(i\pm1,j,k),\Phi(i,j\pm l,k)\} = 0 \quad (15)$$

By analogy with the Gelfand-Levitan-algorithm (see, e.g., [28]) an additional equation results by inverting the equation of continuity, which is a kind of completeness relation, yielding

$$\Sigma_g Q_g e^{2\pi igv} = 0 \quad (16)$$

for Equation (15) as well as for the additional boundary condition previously discussed. Coefficients Q_g are explicitely given in [18]. Thus, in principle, the retrieval of the displacements **v** is given by the remaining inverse problem (16), implying to find the root of a function given by an incomplete Fourier transform.

The inverse problem (16) 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). 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 being tested, viz. the Newton-Raphson algorithm itself to solve Equation (16), and genetic algorithms as well as neuronal networks. However, there is no stable solution up to now, or the algorithm is still too time consuming for application. Besides the numerical solutions, transforming Equation (16) yields an iterative form as a kind of quasi-regularization if the following relations for the arguments are used

$$\cos[2\pi(\mathbf{g}-\mathbf{h})\mathbf{v}+\arg Q_g-\arg Q_h] = \Sigma(|Q_g|^2-|Q_h|^2)/2|Q_g||Q_h| \quad (17)$$

The algorithm demands the iteration for linearly independent coefficients, thus coplanar vectors **g** leave one component unconsidered. Analytical solutions to equations (17) can be performed if four terms at a maximum are considered [19]. Furthermore, Equations (17) refer to an overdetermined system in the same manner as discussed for Equations (10-12).

Conclusions

Both the direct solutions (12) and (13, 15), i.e., the explicit evaluation of thickness and orientation as well as the retrieval of the atomic displacements from a reconstructed electron wave function at the exit surface of an object, result in particular inverse problems of the first kind, viz. the analysis of object parameters from measured data. Thus, from the mathematical point of view the retrieval procedure is an ill-posed inverse problem requiring additional information, e.g., the periodicity of the object, the thickness, the orientation and the unknown reconstructed displacements in order to make the process stable and continuous, i.e., to avoid singularities, and to restrict the manifold set of solutions. 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 singular coefficients. The procedure described has transformed these difficulties to the mathematical problem of overdetermined equation systems and of determining the roots of a function with an incomplete Fourier transform. The uniqueness and the stability of the solutions are determined by the regularization methods applied, which should be discussed in more detail in a forthcoming paper.

Acknowledgements

We are grateful to all reviewers for their helpful discussions, to Dr. E. Völkl and Mrs. H. Messerschmidt for critically reading the manuscript, and to the Volkswagenstiftung for financial support.

References

1. Anstis GR (1989) Simulation techniques for reflection electron microscopy. In: Computer Simulation of Electron Microscope Diffraction and Images. Krakow W, O'Keefe M (eds). The Minerals, Metals & Materials Society, Las Vegas. pp. 229-238

2. Bertero M (1989) Linear Inverse and Ill-Posed Problems. Adv Electronics Electron Phys **75**: 1-114.

3. Coene W, Janssen G, Op de Beeck M, van Dyck D (1992) Phase retrieval through focus variation for ultra-resolution in field-emission transmission electron microscopy. Phys Rev Lett **69**: 3743-3746.

4. Gribelyuk MA (1991) Structure retrieval in HREM. Acta Crystallographica **A47**: 715-723.

5. Head AK (1969) The reconstruction of displacement fields of defects in crystals from electron micrographs: I. Analytical fields. Aust J Phys **22**: 43-50.

6. Head AK (1969) The reconstruction of displacement fields of defects in crystals from electron micrographs:

II. Discontinuous fields and many beams. Aust J Phys **22**: 345-350.

7. Howie A, Basinski ZS (1968) Approximations of the dynamical theory of diffraction contrast. Phil Mag **17**: 1039-1063.

8. Lavrentiev MM (1967) Some Improperly Posed Problems of Mathematical Physics. Springer, Berlin. pp 10-20.

9. Lichte H (1986) Electron holography approaching atomic resolution. Ultramicroscopy **20**: 293-304.

10. Lichte H (1991) Electron image plane off-axis electron holography of atomic structures. Adv Optical Electron Microsc **12**: 25-91.

11. Lichte H (1992) Holography - just another method of image processing? Scanning Microscopy Suppl **6**: 433-440.

12 Lichte H, Völkl E, Scheerschmidt K (1992) Electron holography: II. First steps of high resolution electron holography into materials science. Ultramicroscopy **47:** 231-240.

13. Lois AK (1989) Inverse und schlecht gestellte Probleme (Inverse and Ill-Posed Problems). Teubner Verlag, Stuttgart. pp 45-164.

14. Orchowski A, Lichte H (1996) High resolution electron holography of real structures at the example of a Σ=13 grain boundary in gold. Ultramicroscopy **69**: 199-209.

15. Orchowski A, Lichte H, Scheerschmidt K, Scholz R (1993) Hochauflösende Elektronenmikroskopie zur Analyse von Ó13 Tilt-Korngrenzen in Gold (High-resolution electron microscopy for the analysis of Ó13 grain boundaries in gold). Optik **94** Suppl 5: 79.

16. Orchowski A, Rau WD, Lichte H (1995) Electron holography surmounts resolution limit of electron microscopy. Phys Rev Lett **74:** 399-402.

17. Scheerschmidt K, Hillebrand R (1991) Image interpretation in HREM: Direct and indirect methods. Proc 32nd Course Int Centre of Electron Microscopy "High-Resolution Electron Microscopy - Fundamentals and Applications". Heydenreich J, Neumann W (eds). Elbe Druck, Wittenberg. pp 56-65.

18. Scheerschmidt K, Knoll F (1994) Retrieval of object information from electron diffraction, I. Theoretical preliminaries. phys stat sol (a) **146:** 491-502.

19. Scheerschmidt K, Knoll F (1995) Retrieval of atomic displacements from reconstructed electron waves as an ill-posed inverse problem. Proc Int Workshop Electron Holography. Tonomura A, Allard LF, Pozzi G, Joy DC, Ono YA (eds). Elsevier Science, Amsterdam. pp 117-124.

20. Scheerschmidt K, Knoll F (1995) Zur Rekonstruktion von Verschiebungsfeldern aus elektronenholographisch ermittelten Objektwellen (On the reconstruction of displacements from electron-holographically obtained object waves). Optik **100** Suppl 6: 50.

21. Spence JCH, Zuo JM (1992) Electron Microdiffraction. Plenum Press, New York. pp 134-135.

22. Tichonov AN, Arsenin YY (1977) Solutions of Ill-Posed Problems. Wiley, New York. pp 1-30.

23. Van Dyck D (1985) Image calculation in high resolution electron microscopy. Problems, progress and prospects. In: Advances in Electronics and Electron Physics. Hawkes PW (ed). **65**: 295-355.

24. Van Dyck D (1989) Three-dimensional reconstruction from two-dimensional projections with unknown orientation, position and projection axis. Ultramicroscopy **30**: 435-438.

25. Van Dyck D, Coene W (1988) Direct structure retrieval from high resolution electron micrographs. Scanning Microscopy Suppl **2**: 131-137.

26. Van Dyck D, Op de Beeck M (1992) Direct methods in high resolution electron microscopy. Scanning Microsc Suppl **6**: 115-120.

27. Van Dyck D, Op de Beeck M, Coene W (1993) A new approach to object wavefunction reconstruction in electron microscopy, Optik **93**: 103-107.

28. Zakhariev BN, Suzko AA (1990) Direct and Inverse Problems. Springer, Berlin. pp 1-53.

Discussion with Reviewers

Reviewer I: How does the retrieval technique compare with the conventional image simulation methods, when applied to the same specimen?

D. van Dyck: Why did you not use a simple hypothetical structure to retrieve from? Why do you not consider the whole problem as one optimistic problem where the model is fitted to the data set without additional approximations?

Author: The approximations, i.e., perturbation or discretization, allow the direct solution of the scattering and imaging problem considerably reducing the parameter space, which has to be investigated. I hope this will overcome some ambiguities in finding optimistic values in trial and error image matching techniques. For simple models the amplitudes and phases are simulated by usual methods, with the retrieval procedure then correctly covering thickness and beam orientation. This is used to test the validity of the method and to find suitable regularizations. However, this work is not yet finished for publication.

D. van Dyck: In the literature the deformable ion model is used in connection with the column approximation. Is this also the case here?

Author: No, the differential equations are used in full generality as cited by Anstis and by Howie and Basinski. The deformable ion model allows one to describe defects by a phase factor in the potentials, the applicability of this simplification should be discussed later on.

D. van Dyck: How can the many-beam solution be decoupled into two-beam cases that can be inverted separately?
P.W. Hawkes: How realistic is the linearization by perturbation solution in application?
Author: The many-beam case is not decoupled, but the perturbation is only used to find a linearization, which is invertible around suitable starting values known a priori. The two-beam solution is studied only for comparing the perturbation with analytically known exact solutions, and to give an impression of the validity of the approximation. The validity range of the perturbation seems to be sufficiently large, nevertheless, the applicabillity is not yet clear because of the instabilities, etc., typical of inverse problems. Beyond the validity range, a new start value can be chosen in the iteration procedure.

P.W. Hawkes: Is the situation catastrophic if the problem is over- or underdetermined as you stated?
Author: No, because this is typical of inverse problems. Without noise, the system is overdetermined and cannot be inverted in a simple manner. With noise, however, the system is underdetermined, which offers the possibility of minimizing suitable norms as described. This transforms the ill-posed problem into a well-posed one, which, however, may be ill-conditioned. The instabilities are due to bad conditions and are enhanced by regularizations.

M.A. O'Keefe: Why do you use the term "direct retrieval" whereas you have to apply image processing and holographic techniques inorder to first extract the electron wave, and then to determine the object properties iteratively?
Author: There are always two steps to be solved, viz. the reconstruction of the wave function and the retrieval of the object properties. In both cases here, the trial and error matching technique usually applied is replaced by the mathematically inverse problem which corresponds to direct solutions in physical context.

M.A. O'Keefe: Any HREM or holographic image will have contrast determined not only by the interaction of the electrons with the object, but is influenced in complicated manner by the interferences in the microscope. Furthermore, if the reflections are considered separately, the information spread out due to defects will be missing from the reconstruction. In which way is this considered in the reconstruction process?
Author: Therefore the information limit is discussed not the point resolution limit, including all the aberrations of the microscope, i.e., the damping envelopes, too. Using the holographic reconstruction technique with a suitable phase plate allow one to correct these effects up to the information limit determined by the beam divergency and the microscope instabilities. The perturbation solution is solely applicable to the separated reflections of perfect structures, but the discretization proposed includes crystal defects and allows generally a finer sampling, without any restriction.

Note added in proof:

For the forthcoming paper mentioned in the text (details, numerical tests, etc.) cf. e.g., Scheerschmidt K (1998) Retrieval of object information by inverse problems in electron diffraction. J Microsc 190: 238-248.