

CONTRAST TRANSFER THEORY FOR NON-LINEAR IMAGING

J.E. BONEVICH and L.D. MARKS

Department of Material Science, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA

Received 27 January 1988, in final form 24 May 1988

We investigate an improved version of contrast transfer for non-linear imaging of crystals in High Resolution Electron Microscopy (HREM) that includes higher-order terms of the beam convergence. An analytical solution accommodating the second-order term of the convergence is presented. The improved form of the transmission cross-coefficient including the third-order term of the convergence is evaluated numerically for comparison with the analytical solution. Several test cases are examined that demonstrate that the conventional non-linear theory predicts erroneous results for both amplitude and phase contrast at high values of beam convergence.

1. Introduction

Central to understanding the images in High Resolution Electron Microscopy (HREM) is the process whereby the wave leaving the specimen is transferred to the final image. The nature of the transmission of the exit wave to the image has been exhaustively studied by many authors with various approximations. To summarize the literature, we will briefly review some of the most widely used terms in contrast transfer theory. Linear theory refers to the situation where the specimen acts as a weak scatterer of electrons. The image is formed by the interference of diffracted beams with the transmitted beam and the diffracted beams are considered to be of very weak intensity with respect to the transmitted beam. Consequently, the wave leaving the specimen can be expressed as:

$$\Psi(\mathbf{u}) = 1 - s(\mathbf{u}) + i\phi(\mathbf{u}), \quad (1)$$

where $s(\mathbf{u})$ is the amplitude scattering, $\phi(\mathbf{u})$ the phase scattering and $s, \phi \ll 1$. In this case we can define a contrast transfer function linking the Fourier transform of the final intensity to the Fourier transform of the wave leaving the specimen. Non-linear theory refers to the situation where the specimen strongly scatters electrons. In this case the diffracted beams have sufficient in-

tensity that in addition to interference with the transmitted beam, the diffracted beams interfere with each other.

As presented by Frank [1] and Wade and Frank [2], the linear contrast transfer theory can be written in terms of the product of the unmodified transmission cross-coefficient function and two envelope functions due to spatial and temporal partial coherence. Subsequent investigations by Wade and Jenkins [3] and Jenkins and Wade [4] developed the contrast transfer theory considering tilted and conical bright-field illumination. A closed-form representation of both amplitude and phase transmission was presented by Hawkes [5]; however, only the zeros of the transmission functions were investigated and the envelope terms were not developed.

While linear contrast transfer theory is adequate for the kinematical diffraction of electrons (i.e. a thin amorphous film), extension to strongly scattering objects (i.e. a crystalline specimen) can lead to erroneous results, and amorphous materials are in fact a special case where non-linear, dynamical solutions reduce to linear, kinematical forms, as recognized by Marks [6]. The more accurate non-linear imaging was investigated by Pulvermacher [7], Ishizuka [8] and Anstis and O'Keefe [9] using a first-order Taylor series approximation of the beam convergence (tilt). These

authors demonstrated that, with the improved analysis, some periodicities in the image arise from diffracted beam interference rather than from structural periodicities. In principle this approach can be extended to considering different beams through the crystal, for instance the multislice approach of Coene et al. [10], which is implicit in many standard programs.

One question that remains unanswered at present is non-linear and linear amplitude imaging when the higher-order convergence terms are specifically considered rather than being neglected as in refs. [7–9], i.e. extending the approach of refs. [3–5] to strongly diffracting crystals. This is the subject of this paper, and our conclusion is that these higher-order terms are important. In particular, we present an improved form of the transmission cross-coefficient that specifically includes the second-order term of the convergence with the result that larger beam convergences can be used without the deleterious effects on resolution implied by conventional non-linear theory.

2. Conventional non-linear theory

If we consider the plane wave leaving the exit surface of the specimen $\Psi(\mathbf{r})$ in real space as having a Fourier transform $\Psi(\mathbf{u})$ in reciprocal space, the image intensity can be represented in real space as

$$I(\mathbf{r}) = \text{FT}^{-1} \int \Psi(\mathbf{u}) \Psi^*(\mathbf{u} - \mathbf{v}) \times T(\mathbf{u}, \mathbf{u} - \mathbf{v}, \Delta z) d^2v, \quad (2)$$

where FT^{-1} is the inverse Fourier transform over \mathbf{u} , \mathbf{u} and \mathbf{v} the spatial frequency vectors with moduli u and v as in fig. 1 and are defined in the 2D planes perpendicular to the optic axis, $\Psi^*(\mathbf{u})$ the complex conjugate of $\Psi(\mathbf{u})$, and $T(\mathbf{u}, \mathbf{u} - \mathbf{v}, \Delta z)$ the transmission cross-coefficient [1] due to the phase shift of spatial frequencies and the lens defocus. For the purposes of this paper, we will consider only the transmission cross-coefficient represented by the phase shift and envelope terms

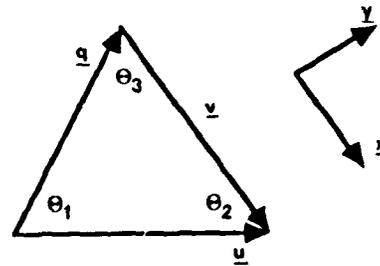


Fig. 1. Orientation of spatial frequency vectors. Note that $v = u - q$ and that $\theta_1, \theta_2, \theta_3$ are internal angles such that $\theta_1 + \theta_2 + \theta_3 = \pi$ radians. Also note the orthogonal coordinate system for the convergence x and y .

in eq. (2). Written out in detail the transmission cross-coefficient is

$$T(\mathbf{u}, \mathbf{q}, \Delta z) = \iint a(\mathbf{u}, \mathbf{q}) \exp[-i\chi(\mathbf{u}) + i\chi(\mathbf{q})] \times F(f) S(w) df dw, \quad (3)$$

where $q = u - v$, $a(\mathbf{u}, \mathbf{q})$ is the aperture function, $\chi(\mathbf{u})$ and $\chi(\mathbf{q})$ are the phase shifts in reduced units, and $F(f)$ and $S(w)$ the one- and two-dimensional distributions of the focal spread and convergence, respectively. We are implicitly neglecting variations in the diffracting condition with changes in the beam angle. The aperture function is non-zero only if both \mathbf{u} and \mathbf{q} lie within the aperture and is assumed to be unity hereafter (i.e. no objective aperture). The astigmatism is also assumed to be corrected.

In conventional non-linear theory, given by Spence [11] and Reimer [12], the phase shift of individual spatial frequencies is approximated by a Taylor series expansion, i.e.

$$\chi(\mathbf{u}, f, w) = (\pi/2) [2(\Delta z + f) |u - w|^2 + |u - w|^4] = \chi(\mathbf{u}, 0, 0) + w \cdot \nabla \chi(\mathbf{u}, 0, 0) + \pi f u^2, \quad (4)$$

where Δz is the lens defocus, f the focal spread, w the beam tilt or convergence, and $\nabla \chi$ the gradient of χ . To simplify the expressions, $\chi(\mathbf{u}, 0, 0)$ will be expressed as $\chi(\mathbf{u})$. While eq. (3) is exact, the expansion in eq. (4) relies on two assumptions: the amount of convergence is so small that only the first-order Taylor series term need be considered,

and the cross-terms between the focal spread, f , and the convergence, w , can be neglected [2].

Assuming that the focal spread and convergence have Gaussian distributions, the transmission cross-coefficient becomes

$$T(\mathbf{u}, \mathbf{q}, \Delta z) = \int \int \exp\{i[\chi(\mathbf{q}) - \chi(\mathbf{u}) + \mathbf{w} \cdot (\nabla\chi(\mathbf{u}) - \nabla\chi(\mathbf{q})) + \pi f(u^2 - q^2)]\} F(f) S(\mathbf{w}) df d\mathbf{w}, \quad (5)$$

with

$$F(f) = \sqrt{(\beta/\pi)} \exp(-\beta f^2), \quad (6)$$

$$S(\mathbf{w}) = (\alpha/\pi) \exp(-\alpha w^2), \quad (7)$$

where β and α are related to the rms widths of the focal spread and the convergence, respectively. Integrating eq. (5) results in the conventional non-linear transmission cross-coefficient

$$\exp\{i[\chi(\mathbf{q}) - \chi(\mathbf{u})]\} \times \exp(-|\nabla\chi(\mathbf{u}) - \nabla\chi(\mathbf{q})|^2/4\alpha) \times \exp[-\pi^2(u^2 - q^2)^2/4\beta], \quad (8)$$

where the exponential terms have the usual significance.

3. Improved theory

The improved version of the transmission cross-coefficient includes the higher-order terms of w . Thus, instead of the Taylor series approximation, the full expansion

$$\chi(\mathbf{u}, f, \mathbf{w}) = (\pi/2) \{2(\Delta z + f)[u^2 - 2\mathbf{u} \cdot \mathbf{w} + w^2] + [u^2 + 2u^2w^2 + w^4 - 4\mathbf{u} \cdot \mathbf{w}(u^2 + w^2) + 4|\mathbf{u} \cdot \mathbf{w}|^2]\} \quad (9)$$

is used. A similar equation can be used for $\chi(\mathbf{q}, f, \mathbf{w})$. Eq. (5) now becomes

$$T(\mathbf{u}, \mathbf{q}, \Delta z) = \int \int \exp\{(i\pi/2)\{2(\Delta z + f) \times [q^2 + 2\mathbf{w} \cdot (\mathbf{u} - \mathbf{q}) - u^2] + [q^4 - 2w^2(u^2 - q^2) - u^4 + 4\mathbf{w} \cdot (u^2\mathbf{u} - q^2\mathbf{q}) - 4(|\mathbf{u} \cdot \mathbf{w}|^2 - |\mathbf{q} \cdot \mathbf{w}|^2) + 4w^2\mathbf{w} \cdot (\mathbf{u} - \mathbf{q})]\}\} \times \sqrt{(\beta/\pi)} \exp(-\beta f^2) df (\alpha/\pi) \times \exp(-\alpha w^2) d\mathbf{w}. \quad (10)$$

If we first consider the integration over f , the envelope term for the focal spread is:

$$\exp\{-\pi^2/4\beta\}[q^4 - 2u^2q^2 + u^4 - 4\mathbf{w} \cdot (\mathbf{u} - \mathbf{q})(u^2 - q^2) + 4|\mathbf{w} \cdot (\mathbf{u} - \mathbf{q})|^2]\}. \quad (11)$$

As in conventional non-linear theory, the cross-terms between the focal spread and the convergence can be considered small. This assumption was verified by Wade and Frank [2] for the case of linear imaging and can be extended to non-linear imaging without significant error. Eq. (11) then becomes

$$\exp\{-[\pi^2(q^2 - u^2)^2/4\beta]\}. \quad (12)$$

This leaves the remaining term in eq. (5) in the form

$$(\alpha/\pi) \int \exp\{i\mathbf{w} \cdot [\nabla\chi(\mathbf{u}) - \nabla\chi(\mathbf{q})] - \alpha w^2 - \pi i[w^2(u^2 - q^2) + 2(|\mathbf{u} \cdot \mathbf{w}|^2 - |\mathbf{q} \cdot \mathbf{w}|^2) - 2w^2\mathbf{w} \cdot (\mathbf{u} - \mathbf{q})]\} d\mathbf{w}. \quad (13)$$

Using the geometry of fig. 1, we can simplify eq. (13) by choosing an orthogonal coordinate system of w based on the vector v . We write x for the component parallel to v and y for the component perpendicular to v . The integration over w can then be broken into two integrals over x (parallel)

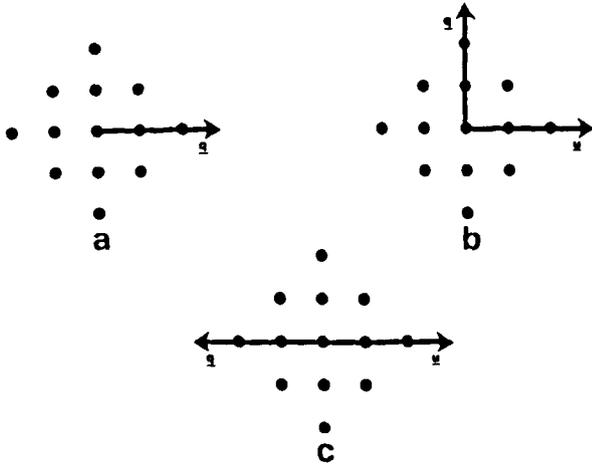


Fig. 2. The orientations of the three test cases. Interference with the transmitted beam in (a); $u=0$, $\Theta_1 = \pi/2$. Interference with g_{100} in orthogonal orientation in (b); $u = g_{100}$, $\Theta_1 = \pi/2$. Interference with g_{100} in parallel orientation in (c); $u = g_{100}$, $\Theta_1 = \pi$.

and y (perpendicular). Consequently, eq. (13) becomes

$$\begin{aligned}
 & (\alpha/\pi) \int \exp\{ix \cdot [\nabla\chi(u) - \nabla\chi(q)] \\
 & \quad - x^2 [\alpha + i\pi(u^2(1 + 2 \cos^2\Theta_2) \\
 & \quad - q^2(1 + 2 \cos^2\Theta_3))]\} dx \\
 & \times \int \exp\{iy \cdot [\nabla\chi(u) - \nabla\chi(q)] \\
 & \quad - y^2 [\alpha + i\pi(u^2(1 + 2 \sin^2\Theta_2) \\
 & \quad - q^2(1 + 2 \sin^2\Theta_3))]\} dy. \quad (14)
 \end{aligned}$$

In eq. (14) the cubic term in w from eq. (13) has been neglected in order to obtain an analytical solution. The solution is straightforward using the Fourier transform of a vector quantity:

$$\begin{aligned}
 & \sqrt{(\pi/N)} \exp(-\pi^2 |M|^2/N) \\
 & = \int \exp(2\pi i K \cdot M - K^2 N) dK, \quad (15)
 \end{aligned}$$

where M is a vector in K space and N is a scalar quantity. Performing the integration in eq. (14)

and combining with eq. (12) results in the following analytical solution for $T(u, q, \Delta z)$:

$$\begin{aligned}
 T(u, q, \Delta z) & = \exp[-i\chi(u) + i\chi(q)] \\
 & \times \exp\{-[\pi^2(q^2 - u^2)^2/4\beta]\} \\
 & \times \exp\{-[\nabla\chi(u) \cos \Theta_2 - \nabla\chi(q) \cos \Theta_3]^2 \\
 & /4[\alpha + i\pi(u^2(1 + 2 \cos^2\Theta_2) \\
 & - q^2(1 + 2 \cos^2\Theta_3))]\} \\
 & \times \exp\{-[\nabla\chi(u) \sin \Theta_2 + \nabla\chi(q) \sin \Theta_3]^2 \\
 & /4[\alpha + i\pi(u^2(1 + 2 \sin^2\Theta_2) \\
 & - q^2(1 + 2 \sin^2\Theta_3))]\} \\
 & \times \alpha / \left\{ \left[\alpha + i\pi(u^2(1 + 2 \cos^2\Theta_2) \right. \right. \\
 & \quad \left. \left. - q^2(1 + 2 \cos^2\Theta_3)) \right] \right. \\
 & \quad \left. \times \left[\alpha + i\pi(u^2(1 + 2 \sin^2\Theta_2) \right. \right. \\
 & \quad \left. \left. - q^2(1 + 2 \sin^2\Theta_3)) \right] \right\}^{1/2}. \quad (16)
 \end{aligned}$$

While the envelope term for the focal spread is identical to that of conventional non-linear theory, the envelope term for the convergence is quite different. For instance, the envelope term is now complex and this makes the total envelope less effective in damping higher spatial frequencies. Consequently, higher values of the beam convergence do not have the deleterious effect on resolution expected from conventional non-linear theory.

4. Numerical analysis

In eq. (14) the cubic term of w was considered to be small enough to be neglected. This was necessary in order to obtain analytical solution of the convergence envelope terms. However, the full solution can be evaluated by iterative numerical methods. For the present case, Simpson's rule was used assuming the convergence distribution had the Gaussian form as specified. The distribution was evaluated over $\pm 5\alpha$ in steps of $\alpha/20$. (The

solution converged properly for these parameters.) The importance of evaluating the full solution can not be overlooked. We have found that the improved solution is much more sensitive to the convergence than the conventional non-linear theory. Therefore, a knowledge of the response of the full solution is a good check on the validity of the improved solution.

5. Discussion of results

A comparison between the conventional non-linear theory, the improved theory, and the full solution of the transmission cross-coefficient was undertaken using three test cases (see fig. 2). The first test was the amplitude and phase contrast response to interference of spatial frequencies with the transmitted beam (linear imaging). The second test involved interference between spatial frequencies in orthogonal orientations (non-linear imaging). The third test involved interference between spatial frequencies along parallel orientations (also non-linear imaging). The aberration parameters chosen represent typical conditions on the Hitachi H-9000 in the HREM Facility at Northwestern University, i.e. 300 keV electrons with 70 Å rms

focal spread, $C_s = 0.9$ mm, $\Delta z = -500$ Å (under-focused), and the system was a hypothetical cubic crystal with 4 Å spacings ($g_{100} = 0.25$ Å⁻¹).

The value of the rms width of the convergence was varied to elucidate the relative effects on the different theories. At low values of the convergence (1 mrad) the improved and full solutions reduce to the results of the conventional non-linear theory. However, at high values of rms convergence (3 mrad) the improved and full solutions result in an extension of the envelope. This result implies that larger values of convergence can be used (i.e. larger condensor apertures) and improve the ultimate resolution and image quality of the microscope. Fig. 3 reveals the response for the first test case (interference with transmitted beam). The rms convergence value of 3 mrad dramatically alters the response. In this case, the conventional theory predicts erroneous results for both the amplitude and phase contrast. Indeed, the conventional theory predicts substantial amplitude contrast reversal whereas the improved and full solutions do not. The extension of the envelope for the improved and full solutions is now highly evident. Thus, the improved and full solutions pass higher spatial frequencies and improve the ultimate resolution.

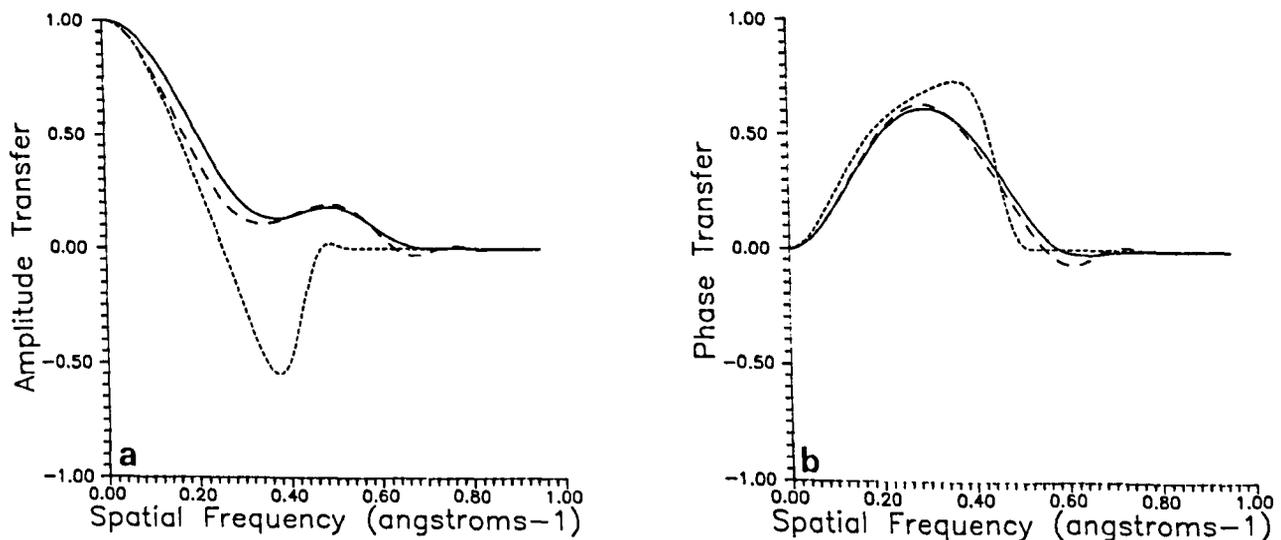


Fig. 3. Interference of spatial frequencies with the transmitted beam for conventional theory (short dashes), improved transmission function (long dashes), and the full solution (solid line). The amplitude contrast is represented in (a); phase contrast in (b). In all cases the following parameters were used: rms focal spread of 70 Å, rms convergence of 3 mrad, electron energy of 300 keV, C_s of 0.9 mm, at an underfocus of -500 Å.

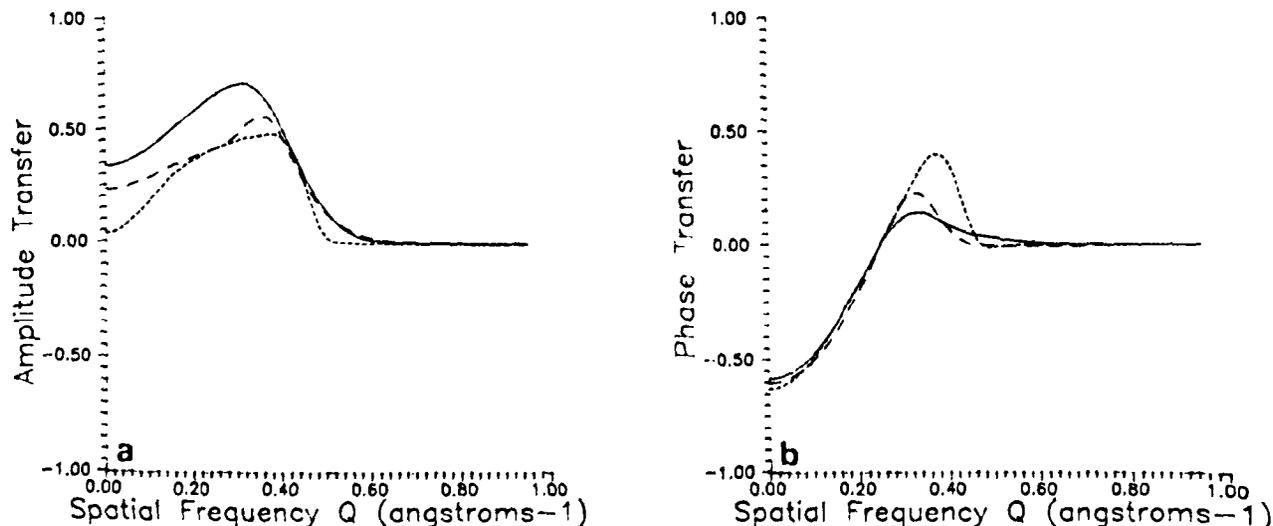


Fig. 4. Non-linear interference with g_{100} (0.25 \AA^{-1}) along orthogonal orientations for the conventional theory (short dashes), the improved transmission function (long dashes), and the full solution (solid line). The amplitude contrast is represented in (a); phase contrast in (b). Note that zero phase contrast occurs for the spatial frequency corresponding to g_{100} as expected. In all cases the following parameters were used: rms focal spread of 70 \AA , rms convergence of 3 mrad , electron energy of 300 keV , C_s of 0.9 mm , at an underfocus of -500 \AA .

The second test case (non-linear interference with $g_{100} = 0.25 \text{ \AA}^{-1}$ along orthogonal orientations) is presented in fig. 4. Increasing the convergence to 3 mrad reveals a dramatic difference between the conventional non-linear and the im-

proved theories. It is important at this time to point out that the improved solution is an approximation of the effect of the convergence. Comparison to the full solution reveals that the improved solution underestimates the response of

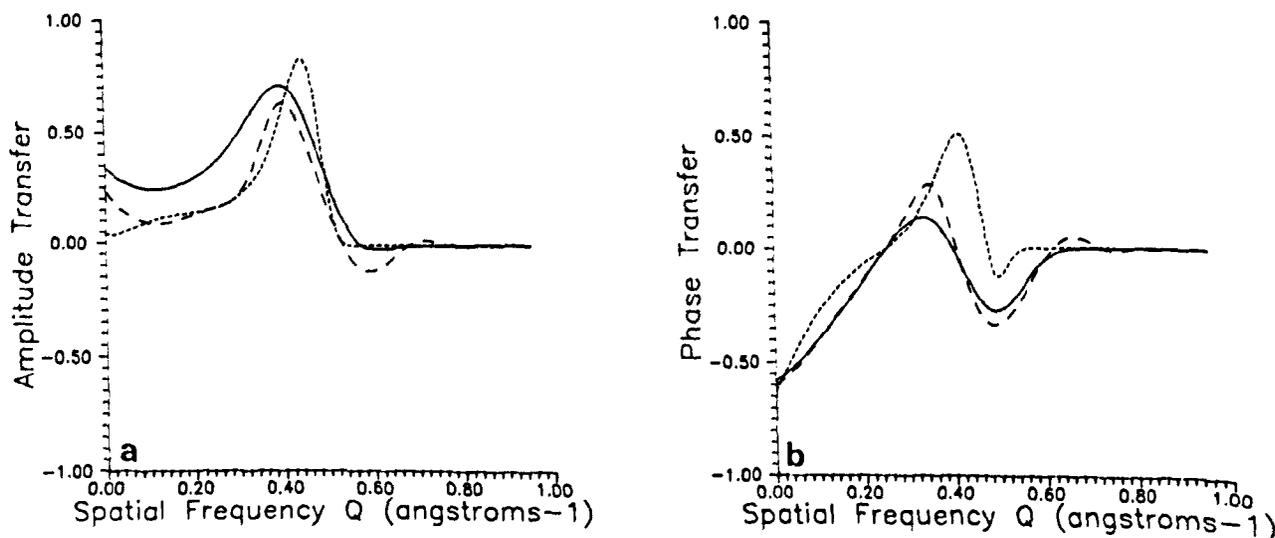


Fig. 5. Non-linear interference with g_{100} (0.25 \AA^{-1}) along parallel orientations for the conventional theory (short dashes), the improved transmission function (long dashes), and the full solution (solid line). The amplitude contrast is represented in (a); phase contrast in (b). Again note that zero phase contrast occurs for the spatial frequency corresponding to g_{100} as expected. In all cases the following parameters were used: rms focal spread of 70 \AA , rms convergence of 3 mrad , electron energy of 300 keV , C_s of 0.9 mm , at an underfocus of -500 \AA .

the transmission cross-coefficient while still being more accurate than conventional non-linear theory. An important feature in fig. 4 is the response at the spatial frequency corresponding to g_{100} (0.25 \AA^{-1}). Here the phase contrast is zero for all three theories. This is to be expected as these spatial frequencies have undergone the same phase shift χ in the microscope. Consequently, only amplitude contrast is available for imaging.

The third test case (non-linear interference with $g_{100} = 0.25 \text{ \AA}^{-1}$ along parallel orientations) is presented in fig. 5. As in the previous test cases, increasing the convergence has a dramatic effect. The conventional theory predicts substantially different behavior from the improved and full solutions. The envelope has been extended to predict higher ultimate resolution of the microscope. As in the second test case, the phase contrast corresponding to g_{100} is zero for all three theories. Again, the improved solution is an approximation of the convergence effect and the full solution represents the true response.

6. Conclusions

Conventional non-linear contrast transfer theory employs Taylor series approximations to gain an insight of the imaging process in the electron microscope. We have found that the higher-order terms of the convergence are highly significant and should not be neglected. The improved non-linear transmission cross-coefficient presented here reveals the inadequacy of conventional non-linear theory. The improved transmission cross-coefficient is an analytical solution considering second-order terms of the convergence. We have found that conventional theory predicts

erroneous results compared to the improved and full solutions for both amplitude and phase contrast at high values of convergence. In particular we have found that large convergences in the improved solution do not have the deleterious effects on resolution as implied by conventional non-linear theory. The drawback of the improved solution is that it may be overly sensitive to the convergence relative to the full solution of the transmission cross-coefficient. Despite this drawback, the improved and full solutions are still more accurate than the conventional non-linear theory.

Acknowledgement

This work was supported by the NSF on Grant number DMR 8514779.

References

- [1] J. Frank, *Optik* 38 (1973) 519.
- [2] R.H. Wade and J. Frank, *Optik* 49 (1977) 81.
- [3] R.H. Wade and W.K. Jenkins, *Optik* 50 (1978) 1.
- [4] W.K. Jenkins and R.H. Wade, in: *Developments in Electron Microscopy and Analysis*, Inst. Phys. Conf. Ser. 36, Ed. D.L. Misell (Inst. Phys., London-Bristol, 1977) p. 115.
- [5] P.W. Hawkes, *Optik* 55 (1980) 207.
- [6] L.D. Marks, *Ultramicroscopy* 25 (1988) 85.
- [7] H. Pulvermacher, *Optik* 60 (1981) 45.
- [8] K. Ishizuka, *Ultramicroscopy* 5 (1980) 55.
- [9] G.R. Anstis and M.A. O'Keefe, unpublished (1977).
- [10] W. Coene, D. Van Dyck and J. Van Landuyt, *Optik* 73 (1986) 13.
- [11] J.C.H. Spence, *Experimental High-Resolution Electron Microscopy* (Clarendon, Oxford, 1981).
- [12] L. Reimer, in: *Transmission Electron Microscopy*, Ed. D.L. MacAdam (Springer, Berlin, 1984).