Scanning Microscopy

Volume 1992
Number 6 *Signal and Image Processing in Microscopy and Microanalysis*

Article 35

1992

Image Delocalisation and High Resolution Transmission Electron Microscopic Imaging with a Field Emission Gun

W. Coene
Philips Research Laboratories, The Netherlands

A. J. E. M. Jansen
Philips Research Laboratories, The Netherlands

Follow this and additional works at: https://digitalcommons.usu.edu/microscopy

Part of the Biology Commons

Recommended Citation

Coene, W. and Jansen, A. J. E. M. (1992) "Image Delocalisation and High Resolution Transmission Electron Microscopic Imaging with a Field Emission Gun," *Scanning Microscopy*: Vol. 1992 : No. 6 , Article 35.
Available at: https://digitalcommons.usu.edu/microscopy/vol1992/iss6/35

This Article is brought to you for free and open access by the Western Dairy Center at DigitalCommons@USU. It has been accepted for inclusion in Scanning Microscopy by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.

Scanning Microscopy Supplement 6, 1992 (pages 379-403)
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

0892-953X/92$5.00+.00

IMAGE DELOCALISATION AND HIGH RESOLUTION TRANSMISSION ELECTRON MICROSCOPIC IMAGING WITH A FIELD EMISSION GUN

W. Coene* and A.J.E.M. Jansen

Philips Research Laboratories, P.O.Box 80000, 5600 JA Eindhoven, The Netherlands

Abstract

The high spatial and temporal coherence of a field emission gun (FEG) increases the information limit of high-resolution transmission electron microscopes (HRTEM), but has also its implications on the localisation of the high resolution information in the image. In this paper, we present the results of a combined theoretical and experimental study of delocalisation in HRTEM. First, we derive a spatial frequency analysis of the delocalisation for crystal defects. Next, the delocalisation is studied from a real-space point of view, in terms of the impulse-response function, for which an instructive asymptotic mathematical analysis has been set up. Finally, we present experimental HRTEM images of crystal defects and of an amorphous Ge film, which are recorded with a Philips CM20 FEG electron microscope, and which illustrate the delocalisation phenomena.

Key Words : high resolution transmission electron microscopy (HRTEM), theory of image formation, image delocalisation, field emission electron gun (FEG), spatial/temporal coherence

* Address for correspondence:
W. Coene
Philips Research Laboratories
P.O.Box 80000
5600 JA Eindhoven
The Netherlands

Phone No.: 31-40-744812
Fax. No.: 31-40-744648

Introduction

The resolution obtainable in high resolution transmission electron microscopy (HRTEM) using intermediate voltage electron microscopes (IVEM) with an accelerating voltage of 200-300 keV, is improved markedly by recent technological developments in the design of the objective lens and of new electron sources. The former developments concern a decrease in the spherical aberration coefficient C_s of the objective lens, which leads to a better point resolution ρ_s at Scherzer focus. This improvement is realised in the Philips CM20 electron microscope with the development of the UltraTWIN objective lens with a C_s value of 0.5 mm (Bakker and Asselbergs, 1990; Otten and Bakker, 1991). The developments with respect to new electron sources concern the use of a field emission gun (FEG) instead of a standard thermionic LaB_6 emitter. For the Philips CM20 FEG microscope, field emission has been achieved by using a Schottky field emitter (Mul et al., 1991; Otten and De Jong, 1991). Due to the lower energy spread of the FEG (twice better than for LaB_6), a high temporal coherence is obtained, so that the defocus spread in the HRTEM image is small. The brightness of the FEG, and therefore the spatial coherence, is also much higher than for a LaB_6 emitter, so that the angular divergence of the effective source is reduced. Typical brightness values are : 5.10^8 Amp/cm^2 srad for FEG; 10^6 Amp/cm^2 srad for LaB_6. The information resolution ρ_i, which is obtained for large underfocus conditions, is improved by the higher coherence of the FEG (in some cases even up to about half the value of the point resolution ρ_s). Thus, not only is the point resolution improved by lowering C_s, but the gap between point and information resolution is enlarged due to the improved coherence of the FEG.

At larger underfocus settings, which carry the information about the ultimate resolution, the contrast transfer function (CTF) of the microscope shows many oscillations. For a FEG, the spatial frequency information from regions where the CTF is rapidly oscillating, survives in the HRTEM image due to the high spatial coherence, as opposed to the

case of a normal thermionic LaB_6 emitter. These oscillations in the CTF not only lead to the well-known contrast reversals, but are also responsible for a delocalisation of the information in the image. These blurring effects of contrast reversal and delocalisation make the interpretation of the HRTEM images difficult, especially in the case of a FEG. Then, the interpretation of the images must be supported by image simulation. A direct solution may be obtained by means of HRTEM image reconstruction (Kirkland, 1984; Van Dyck and Op de Beeck, 1990) using focal image series, by which the effects of the CTF are eliminated, i.e. the contrast reversals are removed and the delocalised high resolution information is brought back to the position where it belongs in the object or specimen. However, the larger the delocalisation in the focal series, the smaller the field of view will become for the faithfully reconstructed image. The same argumentation holds for the digital reconstruction step in high resolution electron holography (Lichte, 1986; Lichte, 1991a).

We only consider here the delocalisation due to the transfer of the microscope. With respect to the interaction between the electron beam and the specimen, the electrons tend to channel along the atomic columns in the case of a zone-axis orientation (Marks, 1984a; Marks, 1985; Van Dyck et al., 1989), so that delocalisation due to diffraction is of minor importance. With respect to image delocalisation, much theoretical work has also been performed by Marks (1984a) and Marks (1985), but in a different framework, either using dispersive equations, or in terms of a Wannier - analysis for the HRTEM image formation.

In this work, we present an elaborate study of delocalisation in HRTEM both from a theoretical and an experimental point of view. Firstly, we analyse in reciprocal space the frequency dependence of the delocalisation at crystal defects. Then, a real space analysis of the delocalisation is presented in terms of the impulse-response function (IRF) that characterises the transfer of the microscope. By means of an asymptotic mathematical analysis, the functional behaviour of the IRF is determined. Finally, we present the results of an experimental HRTEM study which is carried out on a Philips CM20 FEG electron microscope, and which illustrates the delocalisation phenomena.

Basics of HRTEM Image Formation

In this section, we introduce the notations used further on in this paper by briefly reviewing the HRTEM image formation theory. The electron-optical imaging in HRTEM is generally described using the theory of partial coherence (Born and Wolf, 1975), in terms of a transmission-cross-coefficient (TCC; O'Keefe, 1979; Ishizuka, 1980). In the case of a FEG, the parameters which characterise the partial coherence, i.e. beam divergence α and defocus spread Δ, are much smaller than for a thermionic source. This implies that the validity of the quasi-coherent imaging approach (Frank, 1973; O'Keefe and Saxton, 1983; Coene and Van Dyck, 1988) is increased. However, since temporal coherence (due to chromatic aberration and voltage/current instabilities) is still the limiting partial coherence factor, even with the FEG, the quasi-coherent approach is not valid close to the information limit of the microscope, where non-linear imaging contributions are very important. Nevertheless, we prefer to use the quasi-coherent approach here because it is very useful for getting a qualitative intuitive idea about the phenomenon of image delocalisation, and enables us to describe image delocalisation at a crystal defect by means of the shape factor of the defect. Aspects related to non-linear imaging in terms of the TCC can be found in the discussion with the reviewers, at the end of this paper, and in Marks (1984a). Although being complete and the only way for a correct quantitative approach, the TCC description has the disadvantage of losing the intuitive insight due to the scrambling of the specimen information over Fourier space.

In the quasi-coherent approach, the Fourier optics description can be applied (Spence, 1988; Goodman, 1968), so that one can write formally for the intensity of the HRTEM image :

$$I_{im}(\mathbf{R}) = \left|\phi_{im}(\mathbf{R})\right|^2 = \left|\phi(\mathbf{R}) * t(\mathbf{R})\right|^2 = \left|FT^{-1}\left\{\hat{\phi}(\mathbf{G})\ \hat{t}(\mathbf{G})\right\}\right|^2 \qquad (1)$$

with $\mathbf{R}$=(x,y) and $\mathbf{G}$ the two-dimensional coordinate vectors in real and reciprocal space, respectively, and where * denotes a (2-dimensional) convolution. $\phi_{im}(\mathbf{R})$ represents the complex amplitude in the image plane. $\phi(\mathbf{R})$ represents the electron wave function at the exit face of the specimen foil, with $\hat{\phi}(\mathbf{G})$ as its Fourier transform (FT). $\hat{t}(\mathbf{G})$ is the contrast transfer function (CTF) of the microscope in frequency space, with the impulse response function (IRF) $t(\mathbf{R})$ as its inverse FT. It is also illustrative to consider the CTF as a complex Fourier filter applied on the object function ϕ. The CTF $\hat{t}(\mathbf{G})$ is given by (e.g. Frank, 1973; Wade and Frank, 1977) :

$$\hat{t}(\mathbf{G}) = P(\mathbf{G})\,E_{\Delta}(\mathbf{G})\,E_S(\mathbf{G}) \qquad (2)$$

with the phase transfer function (PTF) P(**G**) defined as :

$$P(\mathbf{G}) = \exp\{-\,2\pi i\,\chi(\mathbf{G})\} \tag{3}$$

in which the wave aberration function $\chi(\mathbf{G})$ accounts for the phase distortions due to spherical aberration (C_s) and defocus (Δf) :

$$\chi(\mathbf{G}) = 0.5\left(\Delta f\,\lambda\,G^2 + 0.5\,C_s\,\lambda^3 G^4\right) \tag{4}$$

with λ the electron wavelength, and with $G=|\mathbf{G}|$. The damping envelopes E_Δ and E_S in (2), respectively due to temporal and spatial coherence, are then written as (with α the half angle of beam convergence, and Δ the defocus spread) :

$$E_\Delta(\mathbf{G}) = \exp\left\{-2(\pi\Delta)^2\left(\frac{\partial\chi(\mathbf{G})}{\partial\Delta f}\right)^2\right\}, \tag{5a}$$

and

$$E_S(\mathbf{G}) = \exp\left\{-(\pi\alpha/\lambda)^2[\nabla\chi(\mathbf{G})]^2\right\}. \tag{5b}$$

HRTEM Delocalisation at Crystal Defects

Generally, image delocalisation is concerned with the extent to which the image of an atomic column is affected by its surroundings, i.e. by its neighbouring atomic columns. A perfect homogeneous crystal with a uniform thickness and without defects, has translational symmetry, and therefore the delocalisation is not directly apparent from the HRTEM image. On the other hand, image delocalisation is readily observable at any imperfection or discontinuity in the crystal which destroys locally the translational symmetry. Therefore, we first present a theoretical survey of the effect of the microscope transfer on the delocalisation at a crystal defect.

Theoretical Description

We consider here the case of a single defect between two perfect crystal matrices, as shown in Fig. 1a. The defect is assumed to be in "edge-on" orientation, that is with the defect plane parallel to the electron beam which is incident along the foil normal. Such an "edge-on" defect is an abrupt and well-localised discontinuity in the crystal, and is therefore ideal for a study of the delocalisation. Our results are however readily generalised for other types of "non-uniformities" in real samples, like precipitates, dislocations, thickness or orientation variations over a sample area (curved and/or wedge shaped samples), Fresnel-fringe-like delocalisation at the edge of a sample, etc.

If we take the defect plane d' in Fig. 1a at the origin x=0, we can formally write for the wave function at the exit face of the crystal foil :

$$\phi(\mathbf{R}) = \phi_{<1,2>}(\mathbf{R}) + \phi_{1-2}(\mathbf{R}) + \phi_{d'}(\mathbf{R}) \tag{6}$$

with for the wave function of the "average" matrix $\phi_{<1,2>}(\mathbf{R})$:

$$\phi_{<1,2>}(\mathbf{R}) = 0.5\left\{\phi_1(\mathbf{R}) + \phi_2(\mathbf{R})\right\} \tag{7}$$

and for the "difference" wave function $\phi_{1-2}(\mathbf{R})$:

$$\phi_{1-2}(\mathbf{R}) = \left\{\phi_1(\mathbf{R}) - \phi_2(\mathbf{R})\right\} s(x) \tag{8}$$

where s(x) = -0.5 sign(x) is the shape function of the defect (Fig. 1b). $\phi_1(\mathbf{R})$ and $\phi_2(\mathbf{R})$ represent the wave functions for the perfect matrix of type 1 and 2, respectively, and they take into account the dynamic electron scattering through the foil. $\phi_{d'}(\mathbf{R})$ accounts for the sideways electron scattering across the defect plane d' on propagating through the foil. In other words, $\phi_{d'}(\mathbf{R})$ represents the delocalisation of the defect due to electron scattering in the foil. For thin specimens, $\phi_{d'}(\mathbf{R})$ may be assumed to be negligible due to the electron channeling along the atomic columns (Marks, 1984b; Marks, 1985; Van Dyck et al., 1989). A similar formal description as (6) has been used by Wilson and Spargo (1982) within the context of the simulation of electron diffraction patterns.

The second term in (6), $\phi_{1-2}(\mathbf{R})$, accounts formally for a proper positioning of the defect by means of its shape function s(x). We now consider the effect of the microscope transfer on $\phi_{1-2}(\mathbf{R})$. The corresponding image amplitude in Fourier space is then obtained as

$$\hat{\phi}^{im}_{1-2}(\mathbf{G}) = \left\{\left[\hat{\phi}_1(\mathbf{G}) - \hat{\phi}_2(\mathbf{G})\right] * \hat{s}(u)\right\}\hat{t}(\mathbf{G}) \tag{9}$$

with u the conjugate coordinate of x. The shape function in Fourier space may be considered as a peaked function around the origin u=0, i.e. the information from the defect in Fourier space is present in streaks through the "Bragg spots" or spatial frequencies of the perfect matrix. Then, for the (1-dimensional) convolution in (9), one can

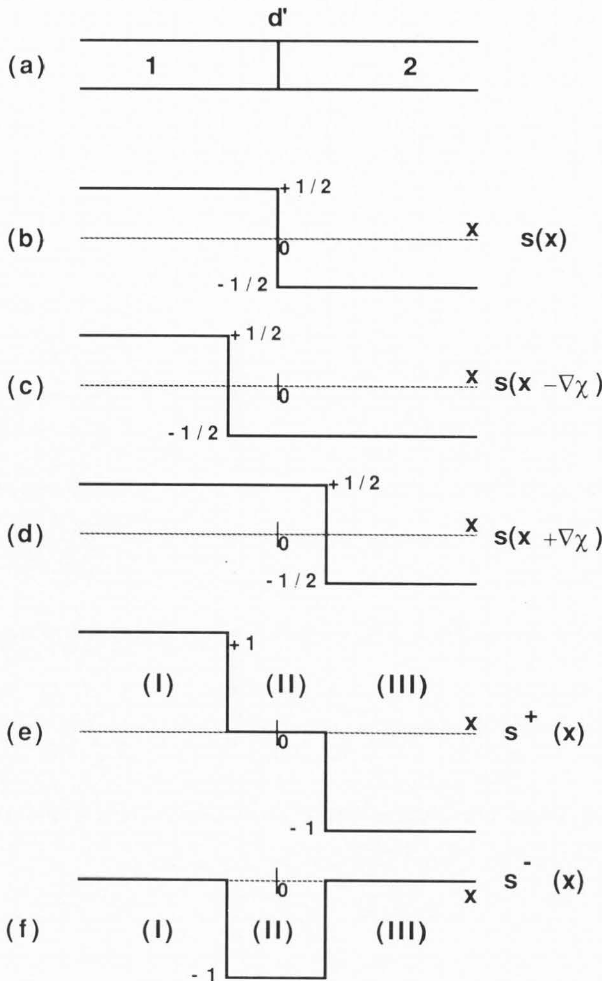

Fig. 1. (a) Defect plane d' between two perfect crystal matrices 1 and 2.
(b) Shape factor of the defect d' at the exit face of the specimen foil.
(c)-(d) Shape factors of the defect d' in the image amplitude.
(e)-(f) Shape factors $s^{\pm}$ of the defect d' in the image intensity.

make use of the expansion :

$$\hat{t}\,(u_B + u, v_B) \cong \hat{t}\,(\mathbf{G}_B)\exp\left\{-2\pi i\, u\frac{\partial\chi(\mathbf{G}_B)}{\partial u}\right\} \tag{10}$$

closely around $\mathbf{G}_B=(u_B,v_B)$, which represent the spatial frequencies for matrices 1 and 2. Then, the image amplitude of the defect contribution $\phi_{1\text{-}2}(\mathbf{R})$ can be rewritten as :

$$\begin{aligned}\phi^{im}_{1-2}(\mathbf{R}) &= \sum_{\mathbf{G}} \hat{\phi}^{im}_{1-2}(\mathbf{G})\exp\{2\pi i\,\mathbf{G}\,\mathbf{R}\}\\ &= \sum_{\mathbf{G_B}}\int_u \left[\hat{\phi}_1(\mathbf{G}_B)-\hat{\phi}_2(\mathbf{G}_B)\right]\\ &\quad \times\ \hat{t}\left(u_B+u, v_B\right)\,\hat{s}\,(u)\\ &\quad \times\ \exp\left\{2\pi i\left[\mathbf{G}_B.\,\mathbf{R}+u\,x\right]\right\}\\ &\cong \sum_{\mathbf{G_B}}\left[\hat{\phi}_1(\mathbf{G}_B)-\hat{\phi}_2(\mathbf{G}_B)\right]\hat{t}\left(\mathbf{G}_B\right)\\ &\quad \times\ \exp\left\{2\pi i\mathbf{G}_B.\,\mathbf{R}\right\}\\ &\quad \times\ \int_u \hat{s}\,(u)\,\exp\left\{2\pi i\,u\left(x-\frac{\partial\chi(\mathbf{G}_B)}{\partial u}\right)\right\}du\end{aligned}$$

so that finally :

$$\phi^{im}_{1-2}(\mathbf{R}) \cong \sum_{\mathbf{G_B}}\left(\hat{\phi}_1-\hat{\phi}_2\right)\hat{t}\ \ s\,(x-\frac{\partial\chi}{\partial u})\ \times\ \exp\left\{2\pi i\,\mathbf{G}_B.\,\mathbf{R}\right\} \tag{11}$$

The explicit dependence of ϕ_1, ϕ_2 , $\hat{t}$ and χ on the summation vector $\mathbf{G}_B$ has been omitted in (11) for simplicity. Relation (11) is a key formula in this paper, and may be compared to the defect term $\phi_{1\text{-}2}(\mathbf{R})$ before transfer through the microscope. The shape function s(x) determines then the true position of the defect. Now, due to the microscope transfer, the shape function is shifted. The magnitude and direction of the shift depend on the spatial frequency $\mathbf{G}_B$ by means of the derivative of the wave aberration function χ. This implies that each spatial frequency sees the defect at a different position, and this position varies also with focus.

Relation (11) can be generalised for an arbitrary shape function $s(\mathbf{R})$ by replacing the s - factor in (11) by $s(\mathbf{R}-\nabla\chi(\mathbf{G}_B))$, e.g. for the case of a precipitate in a matrix, or in the case of more irregularly shaped "defects". An expression similar to (11) for the simplified case of focus transfer only ($C_s = 0$), has already been described by Van Dyck (1987). From (11), it is clear that the contribution of the defect to the image *amplitude* may be delocalised from the original defect position. For the effect on the image *intensity* , an analytical expression can be derived under the "weak-phase object" (WPO) approximation (see e.g. Cowley and Iijima, 1972; Pirouz, 1981), with for the ac components of the wave function $\phi \approx iV$, with V the potential for the electron-specimen

interaction. For the defect in Fig. 1a, this expression can be written as :

$$I_{im}(\mathbf{R}) \cong 1 - \sum_{\mathbf{G_B} \neq \mathbf{0}} \hat{t}_I \left[\left(\hat{V}_1 + \hat{V}_2 \right) + \left(\hat{V}_1 - \hat{V}_2 \right) [s(\mathbf{R} - \nabla\chi) + s(\mathbf{R} + \nabla\chi)] \right] \times \exp\left\{ 2\pi i \, \mathbf{G_B} \cdot \mathbf{R} \right\} - i \sum_{\mathbf{G_B} \neq \mathbf{0}} \hat{t}_R \left[\left(\hat{V}_1 - \hat{V}_2 \right) [s(\mathbf{R} - \nabla\chi) - s(\mathbf{R} + \nabla\chi)] \right] \times \exp\left\{ 2\pi i \, \mathbf{G_B} \cdot \mathbf{R} \right\} . \tag{12}$$

Again, the explicit dependence on $\mathbf{G_B}$ has been omitted in the summations in (12). $t_{R,I}$ represent the real and imaginary part of the transfer function, respectively. In (12), we distinguish a term modulated by t_I, which corresponds with the usual WPO expression as for perfect crystals, and an additional term, only due to the defect, and which is modulated by t_R. For these two terms, the shape function s of the defect is transformed on propagating from the object to the image, into $s^{\pm}(\mathbf{R},\mathbf{G_B})$, given by :

$$s^{\pm}\left(\mathbf{R},\mathbf{G_B}\right) = s\left(\mathbf{R} - \nabla\chi(\mathbf{G_B})\right) \pm s\left(\mathbf{R} + \nabla\chi(\mathbf{G_B})\right) \tag{13}$$

In Fig. 1e,f the shape functions for the intensity at the image plane are drawn for the case of the planar defect. The boundaries of the different regions I, II and III are dependent on the spatial frequency $\mathbf{G_B}$. Region II is centered around the original defect position, and has a width equal to $2\nabla\chi(\mathbf{G_B})$. In regions I and III, where s^- vanishes, an image of the matrices 1 (or 2) is obtained, which, for the frequency $\mathbf{G_B}$, is not distorted by the defect. In the defect region II, s^+ vanishes. There, the "average" matrix is imaged, together with the second defect term, modulated by t_I and t_R, respectively. Due to the t_R-t_I mixing, the information in the defect region is not easily interpretable. At the boundaries of region II, i.e. at $\mathbf{R} \pm \nabla\chi$, shadow images of the defect can be seen, for frequency $\mathbf{G_B}$.

The delocalisation of the information about the defect is fully incorporated by the shape factors $s^{\pm}$ in (12-13). This is easily illustrated by the following argumentation. If the dependence on $\nabla\chi$ of the defect shape function at the level of the image amplitude would be neglected, i.e. for a flat CTF $\nabla\chi = 0$, and replacing $s(\mathbf{R})$ for $s(\mathbf{R} \pm \nabla\chi)$ in (12-13),

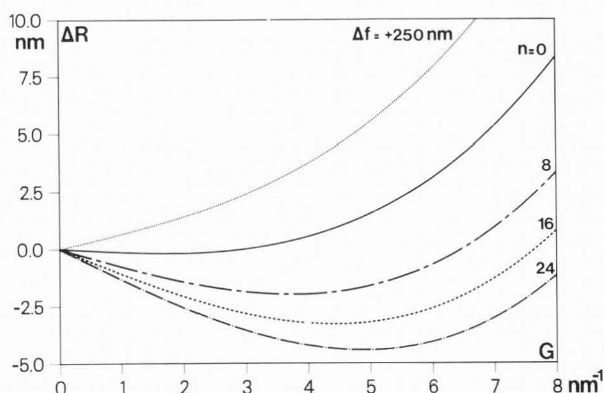

Fig. 2. Delocalisation plots as a function of spatial frequency G, representing $\Delta R = \nabla\chi(G)$, for a series of practical focus values (200 keV microscope, $C_s = 1.2$ mm). The underfocus values are chosen as the "passband" defoci $\Delta f_S(n)$ from (15), with n indicated.

then s^+ equals 2s, and s^- vanishes together with the t_R - term in the image intensity of (12). Under such an approximation, the image of the defect (12) would again be localised at its proper position, and the image would be "faithful" in the usual WPO sense.

In the case of crystal tilt, the inversion relation $\phi^*(-\mathbf{G}) = \phi(\mathbf{G})$, which has been used to derive the WPO expression (12), is violated, and then the shifted shadow images of the defect at $\mathbf{R} \pm \nabla\chi$ become asymmetric.

Focal Dependence of Image Delocalisation

From relations (11-12), we derive for the delocalisation by means of the frequency-dependent defect-shift (or information-shift) :

$$\Delta\mathbf{R} = \nabla\chi(\mathbf{G}) = \lambda\mathbf{G}\left(\Delta f + C_s \lambda^2 G^2 \right) . \tag{14}$$

Fig. 2 shows delocalisation plots as a function of spatial frequency for a series of focus values for a 200 keV microscope with $C_s = 1.2$ mm (Philips CM20-SuperTWIN). The values for the underfocus settings are particularly chosen as

$$\Delta f = \Delta f_S^{(n)} = -\sqrt{C_s \lambda}\,\sqrt{4n + \frac{3}{2}} . \tag{15}$$

These underfoci yield broad "passbands" in the imaginary part of the PTF (Spence, 1988). Scherzer or optimum focus is then obtained as the special case for n=0. The E_S envelopes (5b) corresponding to the focus values of Fig. 2, are shown in Fig. 3 ; the

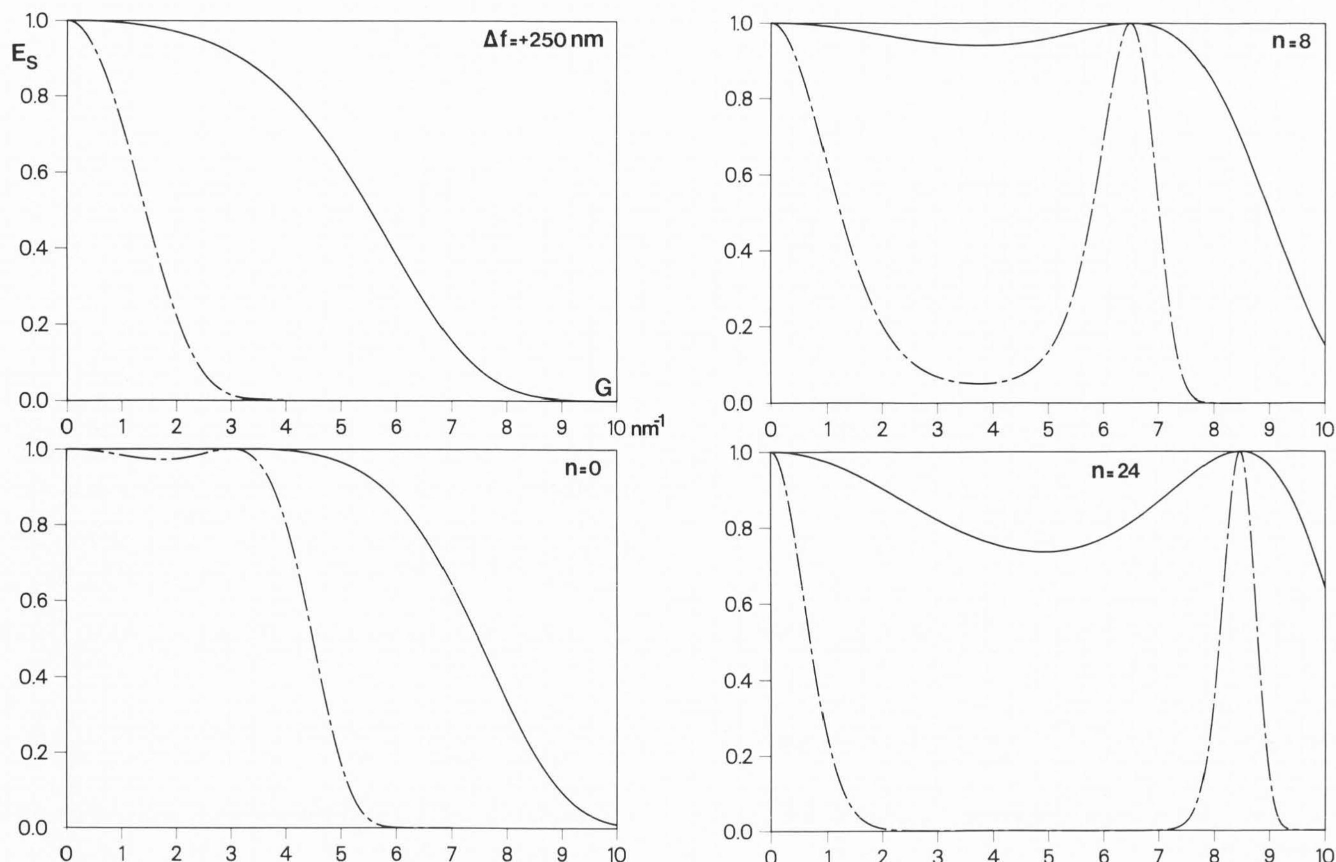

Fig. 3. Damping envelopes $E_S(G)$ due to spatial coherence, for a series of practical focus values (200 keV microscope; C_s = 1.2 mm). The underfocus values are chosen as the "passband" defoci from (15), with n indicated.

Solid line : $\alpha = 1.\ 10^{-4}$ rad (FEG);
Dashed line : $\alpha = 7.\ 10^{-4}$ rad (LaB_6).

corresponding (imaginary part of the) CTFs and IRFs are drawn in Fig. 4, and will be discussed further on. For underfocus conditions ($\Delta f<0$), a negative minimum displacement (or, the largest negative value)

$$\Delta R_{min} = \frac{2}{3}\sqrt{\frac{|\Delta f|^3}{3C_s}} \tag{16a}$$

is obtained at a spatial frequency

$$G_{min} = \sqrt{\frac{|\Delta f|}{3C_s\lambda^2}}\,. \tag{16b}$$

If $\Delta f<0$, no delocalisation ($\Delta R=0$) is found at the spatial frequencies

$$G_1=0 \quad \text{and} \quad G_2=\sqrt{\frac{|\Delta f|}{C_s\lambda^2}}\,. \tag{17}$$

Note that both G_{min} (16b) and G_2 (17) increase with the square root of the absolute value of focus $|\Delta f|$, whereas the negative minimum displacement ΔR_{min} (16a) increases with the 3/2 power of $|\Delta f|$.

For overfocus conditions ($\Delta f>0$), the displacement increases steadily with spatial frequency G. For small overfocus Δf, and/or large frequency G, the C_s - term in (14) will be dominating; for larger overfocus and/or small frequency G, the delocalisation ΔR will increase about linearly with focus. Relation (14) can now be used to determine an optimum underfocus condition which keeps the delocalisation within certain bounds. One useful

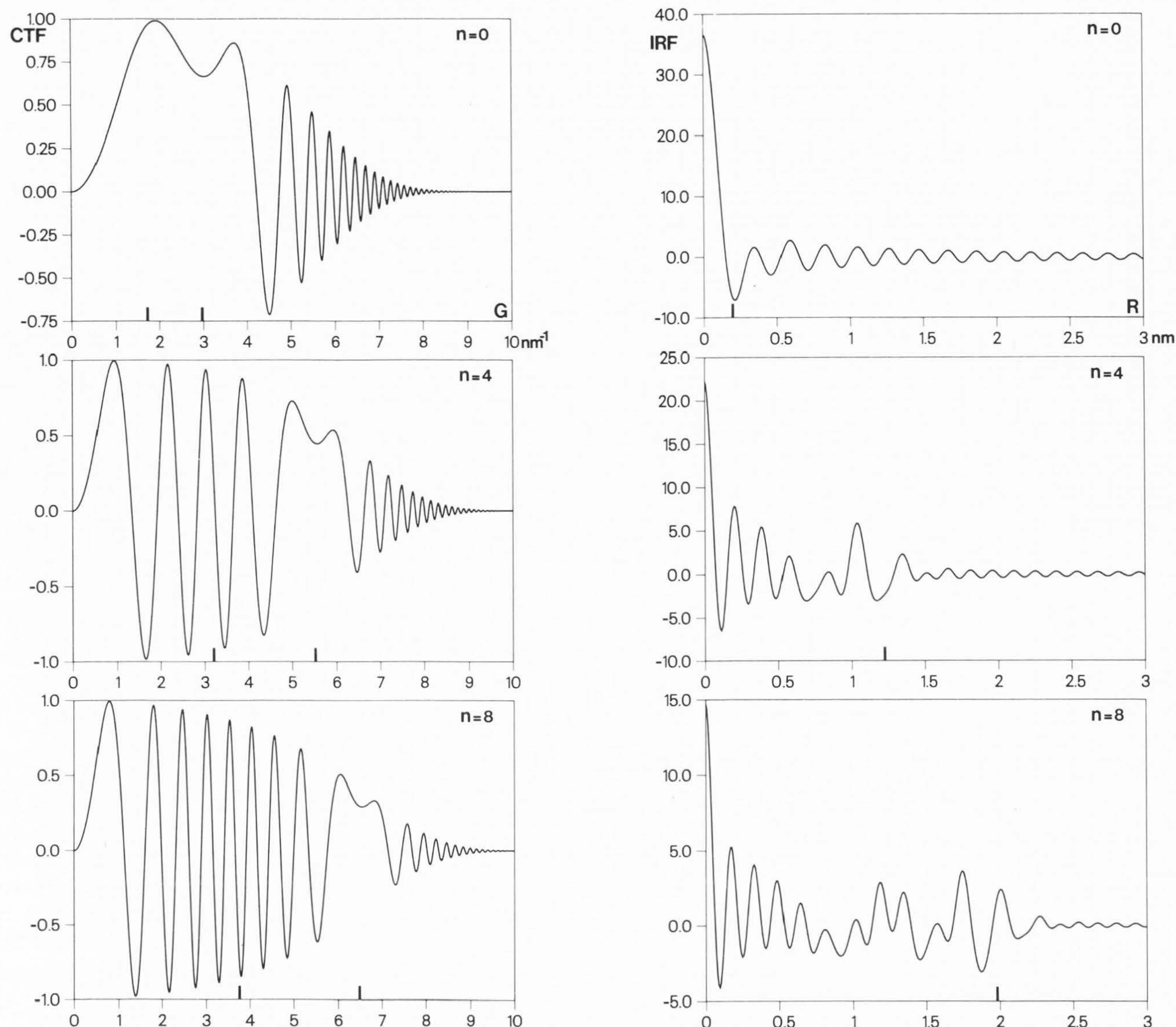

Fig. 4. Imaginary part of the impulse-response function (IRF) $t_I(R)$ and the transfer function (CTF) $\tilde{t}_I(G)$ for a series of "passband" focus values $\Delta f_S(n)$ from relation (15), with n indicated (200 keV microscope; C_s = 1.2 mm; defocus spread Δ = 4.0 nm; beam convergence $\alpha = 1.\,10^{-4}$ rad). The bar on the horizontal axis of the IRFs indicates the caustic point R_c. The bars on the horizontal axis of the CTFs indicate the frequency with maximum delocalisation and the frequency without delocalisation, respectively.

condition may be the choice that the ultra-high resolution information at a maximum frequency $G=G_{max}$ is purely localised. The optimum focus is then given by

$$\Delta f_{opt} = -C_s \lambda^2 G_{max}^2, \tag{18}$$

and the largest displacement

$$\Delta R_{min} = -\frac{2}{3\sqrt{3}} C_s \lambda^3 G_{max}^3 \tag{19}$$

is then obtained at

$$G_{min} = G_{max}/\sqrt{3}$$

Another useful condition (Lichte, 1991b) may be that the displacement is minimum over the whole frequency range $0 \le G \le G_{max}$. This leads to an optimum focus

$$\Delta f_{opt} = -\frac{3}{4} C_s \lambda^2 G_{max}^2 \tag{20}$$

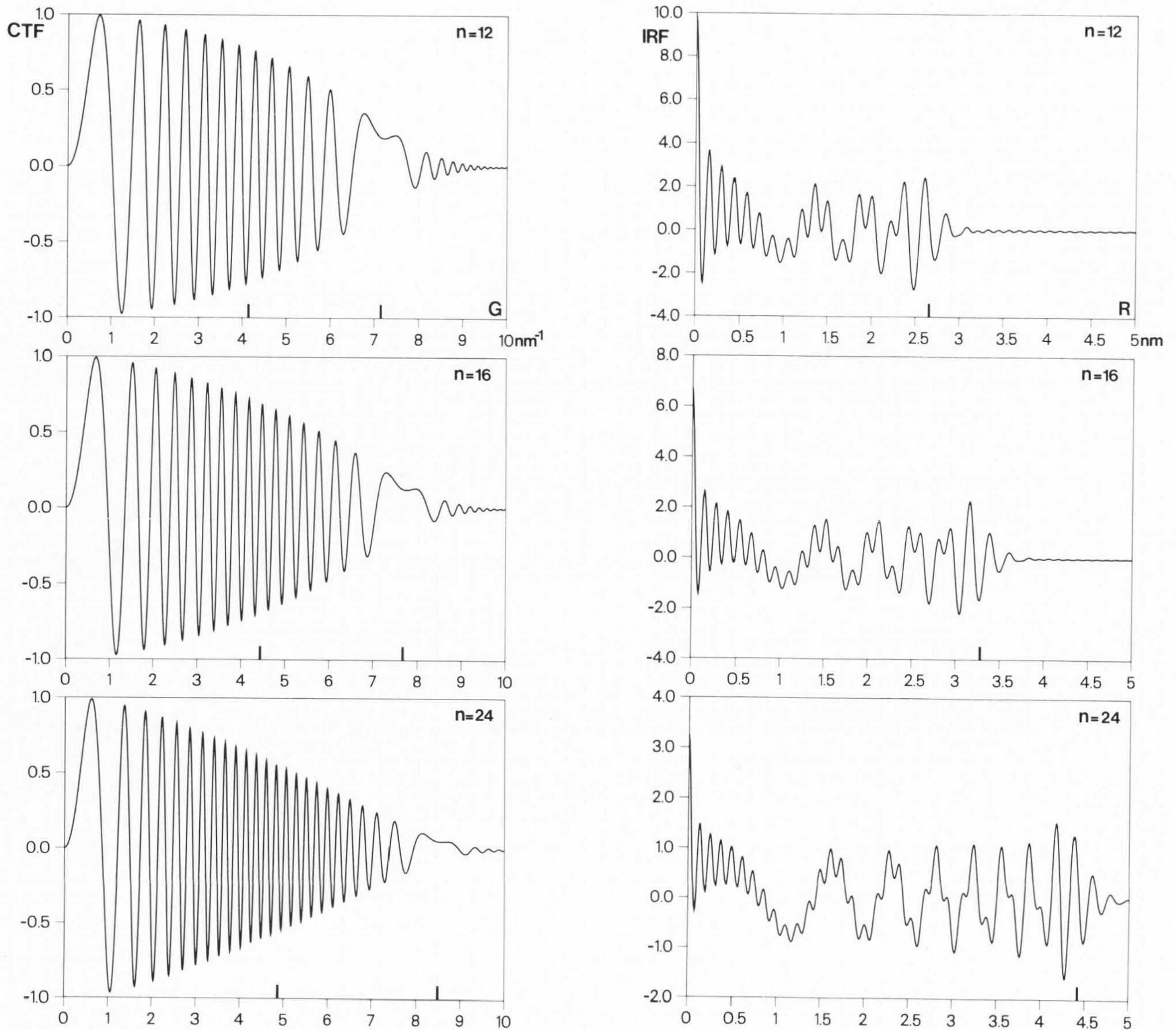

Fig. 4. Continued.

with the largest displacement

$$\Delta R_{min} = \frac{1}{4} C_s \lambda^3 G_{max}^3 \tag{21}$$

at $G=G_{max}$ and $G=G_{min}=G_{max}/2$. The latter value of ΔR_{min} is a factor of 1.5 smaller than the former value, but the first condition has the advantage that the damping due to the E_s envelope around G_{max} vanishes ($E_s=1$), which is not the case for the second condition. This consideration is of importance in the case of non-optimum signal-to-noise statistics around the ultimate resolution G_{max}.

Another important feature from both assumptions for an optimum focus is that the maximum displacement ΔR_{min} increases linearly with C_s, and with the third power of the ultimate resolution G_{max}. Note that the optimum focus itself also increases linearly with C_s. Assuming in particular that $G_{max}=\rho_i^{-1}$, then for the latter optimum focus condition we obtain

$$\Delta R_{min} = \frac{1}{4}\left(\frac{3}{2}\right)^4 \rho_i \left(\frac{\rho_s}{\rho_i}\right)^4, \tag{22}$$

with for the point resolution (Spence, 1988) :

$$\rho_s = \frac{2}{3}\left(C_s \lambda^3\right)^{0.25}.$$

Thus, the delocalisation increases with the fourth power of the ratio of the point resolution divided by the information resolution. Since the information limit ρ_i is chiefly determined by the (temporal) coherence of the electron source, it is worthwhile to indicate at this point the importance of ultra-high resolution polepieces (Bakker and Asselbergs, 1990), if one wants to limit the delocalisation effects in HRTEM. In other words, improving C_s makes the information in the image from the resolution limit ρ_i more readily interpretable. As an example, for the same information resolution ρ_i, the delocalisation with the CM20-UltraTWIN (C_s = 0.5 mm) is reduced with a factor of 4 as compared to the CM20-TWIN (C_s = 2.0 mm), and with a factor of 2.4 as compared to the CM20-SuperTWIN (C_s = 1.2 mm).

Image Delocalisation and Spatial Coherence

The importance of delocalisation in an HRTEM image is closely related to the spatial coherence of the electron source, which is taken into account via the envelope E_s in the CTF. Then, from (14) and (5b), it is obvious that the highest damping due to E_s occurs for those spatial frequencies which suffer from the highest delocalisation in the image. In this spatial frequency region, the transfer function is rapidly oscillating (Fig. 4), i.e. the value of $\nabla\chi(\mathbf{G})$ is large (Fig. 2). The amount of damping due to E_s depends on the degree of spatial coherence, indicated by the convergence angle α. In the case of HRTEM with a conventional LaB_6 electron source, practical values of α are in the range 0.5 - 1.0 mrad. For HRTEM using a FEG, the value of α may be smaller by about a factor of 10-30, yielding $\alpha = 5.\ 10^{-5}$ rad (or even lower). Note the quadratic dependence on α in the argument of the exponential in E_s (5b). In Fig. 3, the E_s envelope is drawn for the case of a Philips CM20-SuperTWIN microscope with a conventional LaB_6 source, and with a FEG, for some practical defocus values. In the case of the (relatively) limited spatial coherence of the LaB_6 source, it turns out that the highly delocalised information is severely or even completely damped by E_s, and is thus absent in the HRTEM image. On the other hand, for the high spatial coherence of the FEG, the damping due to E_s is modest, and therefore also the delocalised information is intrinsically present. In Figs. 3-4, we have taken $\alpha = 1.\ 10^{-4}$ rad for the FEG, which is somewhat larger than the value mentioned above : our results show that already for this value of α, the delocalisation effects are prominent.

Image Delocalisation and Beam Misalignment

Another point which should be mentioned in this respect is the sensitivity on an unknown residual beam misalignment or beam "tilt" **K**. This beam tilt leads to an additional phase factor in the transfer function, given by :

$$\hat{t}(\mathbf{G}+\mathbf{K}) \cong \hat{t}(\mathbf{G}) \exp[-2\pi i\, \mathbf{K}.\nabla\chi(\mathbf{G})]. \quad (23)$$

The additional phase distortion is important for the spatial frequency region where $\nabla\chi(\mathbf{G})$ is large. The information from these spatial frequencies will again only survive in the image if the spatial coherence is high (or α low), as is the case when using a FEG.

In other words, when running a FEG in a highly spatially coherent mode, delocalisation effects are much more prominent, and an accurate beam alignment or knowledge of its misalignment **K** is much more critical.

HRTEM Impulse-Response Functions and Delocalisation

In the preceeding section, we have performed a spatial frequency analysis of the delocalisation. In this section, we present a study of the delocalisation from a real-space point of view in terms of the impulse-response function (IRF) as defined in (1). The IRF represents in fact the image *amplitude* of an ideal point object at the exit face of the specimen foil, and is therefore a *direct* measure of the delocalisation. The 'real' image amplitude is then obtained by convolution (1) with the assembly of point scatterers, which is represented by the object wave ϕ.

In the ideal case of a radially symmetric transfer function (no astigmatism and no beam misalignment), we obtain for the IRF by means of the Fourier-Bessel transform :

$$t(R) = 2\pi \int_0^\infty \hat{t}(G)\, J_0(2\pi G R)\, G\, dG \quad (24)$$

with J_0 the zero-th order Bessel function. For the explicit form of t(G) in (2), the IRF can only be computed by numerical integration. On the other hand, in the case of ideal coherence (i.e. no damping envelopes in $\hat{t}$), and with C_s equal to zero, the IRF is readily computed from (22) as the Fresnel or focus propagator given by :

$$\frac{-i\pi}{\lambda \Delta f} \exp\left\{\frac{i\pi R^2}{\lambda \Delta f}\right\}. \quad (25)$$

Relation (25) also describes the usual Fresnel fringe effect for large focus values, i.e. when the effect of C_s is negligible.

However, for the general case ($C_s \neq 0$), even without damping envelopes, no analytical formula for t(R) can be given. We propose here now the results of an asymptotic analysis of (24) in the ideally coherent case, with an emphasis on the functional behaviour with respect to focus Δf and position R. This analysis is essential for the extrapolation and the interpretation of the numerically obtained IRF's. A focal series of CTF's and IRF's are shown in Fig. 4 for the "passband" defoci given by (15).

The asymptotic analysis of (22) is similar to that of the Pearcey integral (Kaminski, 1989; Paris, 1991) and is carried out by Janssen (1992). For the mathematical framework of this analysis, the interested reader is referred to the appendix (and references therein). The impulse response of the microscope depends both on C_s and Δf, but both parameters are reduced to one single essential parameter by introduction of dimensionless focus and position using the (slightly modified) Scherzer and Glaeser units :

$$\Delta f' = \Delta f \sqrt{\frac{2\pi}{C_s \lambda}}, \tag{26a}$$

and

$$R' = R \sqrt[4]{\frac{4(2\pi)^3}{C_s \lambda^3}}. \tag{26b}$$

We will restrict the analysis here to underfocus ($\Delta f<0$) since this is the most intricate case. For overfocus settings ($\Delta f>0$, large), the asymptotics of t(R') have been given in (Janssen, 1992); it turns out that no elegant extensions going beyond the Fresnel propagator (25) can be derived. From the numerically computed IRF's (including E_Δ and E_S damping envelopes) in Fig. 4, it turns out that one can distinguish several R' - ranges where t(R') exhibits (composite) oscillations with frequencies characteristic for these ranges. Especially for the larger underfocus conditions in Fig. 4, we can easily distinguish in the IRF three separate regions with different oscillation modes. The effect of the damping envelopes E_Δ and E_S should in principle be taken into account by convolution in real space with their inverse Fourier transforms. However, in order to obtain useful qualitative insight, we can account for the damping by multiplication with the values of E_Δ and E_S for the frequencies in a particular R' - range.

In the first region (I), i.e. close to the origin R'=0, the IRF can be approximated by (see appendix) :

$$t^{(I)}(R') \cong \sqrt{\frac{2\pi}{C_s \lambda^3}} \times \left\{ \sqrt{\pi}\exp\left\{-\frac{i}{4}\left(\pi - \Delta f'^2\right)\right\} J_0\left(R'\sqrt{\tfrac{1}{2}|\Delta f'|}\right) - \frac{i}{\Delta f'}\exp\left\{\frac{i}{4}\frac{R'^2}{\Delta f'}\right\}\right\} \tag{27}$$

The first term in (27) is the leading term of an asymptotic series expansion, and it has a main frequency of $1/(2\pi)\sqrt{0.5|\Delta f'|}$, since for large arguments z of the Bessel function

$$J_0(z) \approx z^{-0.5}\cos\left(z - \frac{\pi}{4}\right).$$

This result fits perfectly in the frequency analysis of the preceeding section. Indeed, the frequency $1/(2\pi)\sqrt{0.5|\Delta f'|}$ is equal to G_2 in (17), using the unscaled variables G, R. At G_2, $\nabla\chi$ is zero, and we then know that the information from this frequency is localised (close to the origin R'=0) in the image. For the larger underfocus values, $1/(2\pi)\sqrt{0.5|\Delta f'|}$ shifts towards higher frequencies where the effect of the E_Δ damping envelope becomes larger. Then the second term (the Fresnel propagator) in (27) comes into play, which has a linearly increasing frequency $R'/[4\pi\Delta f']$, and is therefore much less subjected to damping. In Fig. 4f, we can indeed observe the effect of this extra term as a slow modulation super-imposed upon the faster oscillations of J_0. Also, for R' close to zero, the graph of the Bessel function has a (small) positive shift due to this extra term.

The phase factor which is multiplied with the Bessel function in the first term of $t^{(I)}(R')$ in (27) is independent of position, and has with respect to focus, a linearly increasing frequency of $|\Delta f'|/(4\pi)$. In other words, as a function of focus, the rate of image contrast reversals increases linearly with focus (for the information corresponding with R' close to the origin). For the special case of the "passband" defocus values of relation (15), the phase factor reduces to i.

Next, the second region (II) with its characteristic oscillation modes extends up to a special point, which is called the "caustic" point in the asymptotic analysis, and which is given by :

$$R'_c = \left(\frac{2}{3}|\Delta f'|\right)^{3/2}, \tag{28a}$$

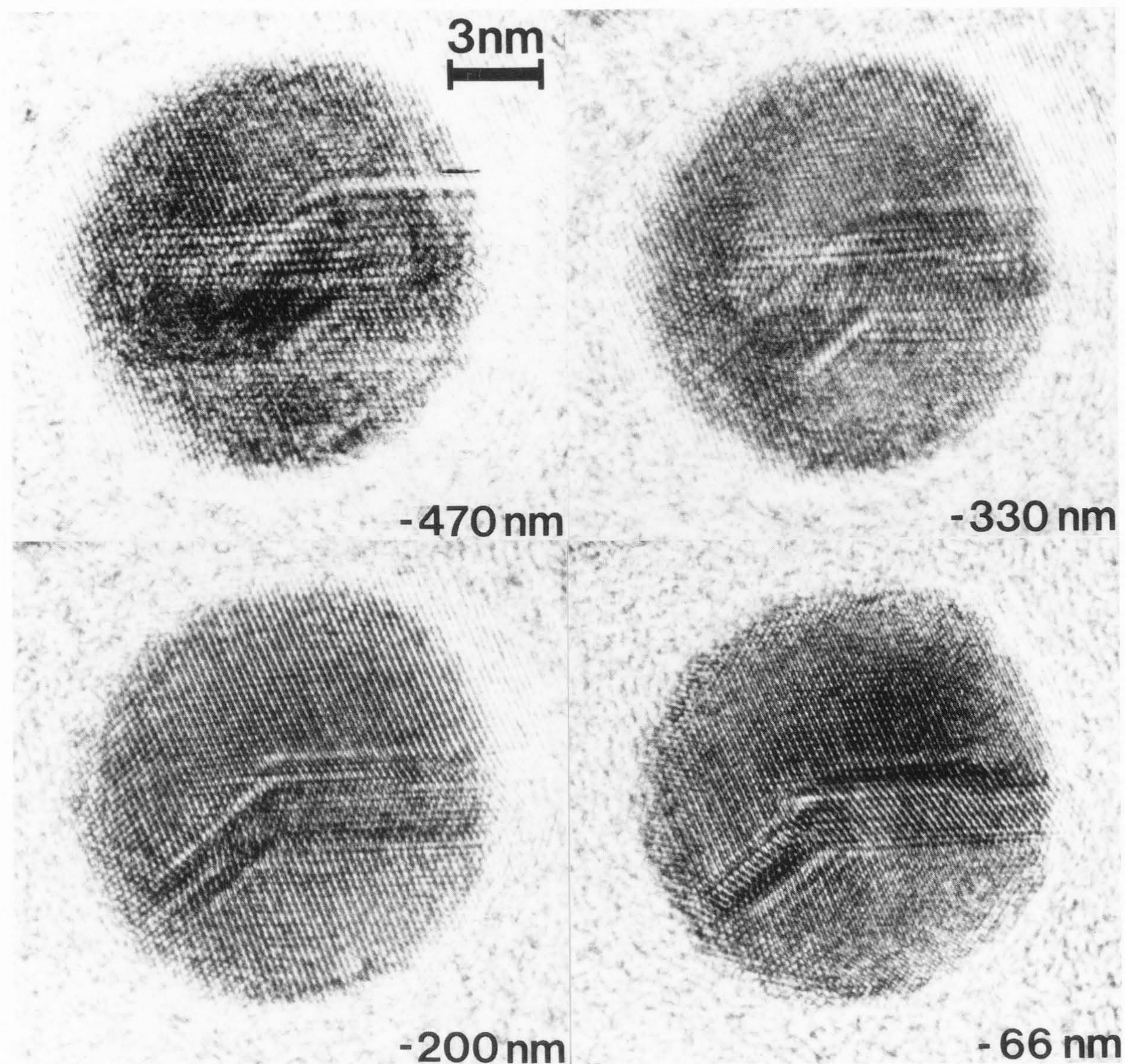

Fig. 5. Through-focus HRTEM image series of a Au - particle on a carbon film (CM20 FEG-SuperTWIN).

or, in unscaled variables :

$$R_c = \left(\frac{4\,|\Delta f|^3}{27\,C_s} \right)^{1/2} . \tag{28b}$$

Note that the value of the caustic point (28b) is identical to (16a). The caustic point is a clear boundary between two mathematical regimes in the asymptotic analysis. Beyond the caustic point, the oscillations in t(R) have a (very) small amplitude as will be shown further on. Therefore, the value of the caustic point can be considered as a fair measure of the extent of the delocalisation. A further interesting result is that according to relation (28a), the delocalisation for underfocus conditions increases with the 3/2 power of the absolute value of (scaled) focus.

On the lower side of the caustic, i.e. $R' \leq R'_c$, the

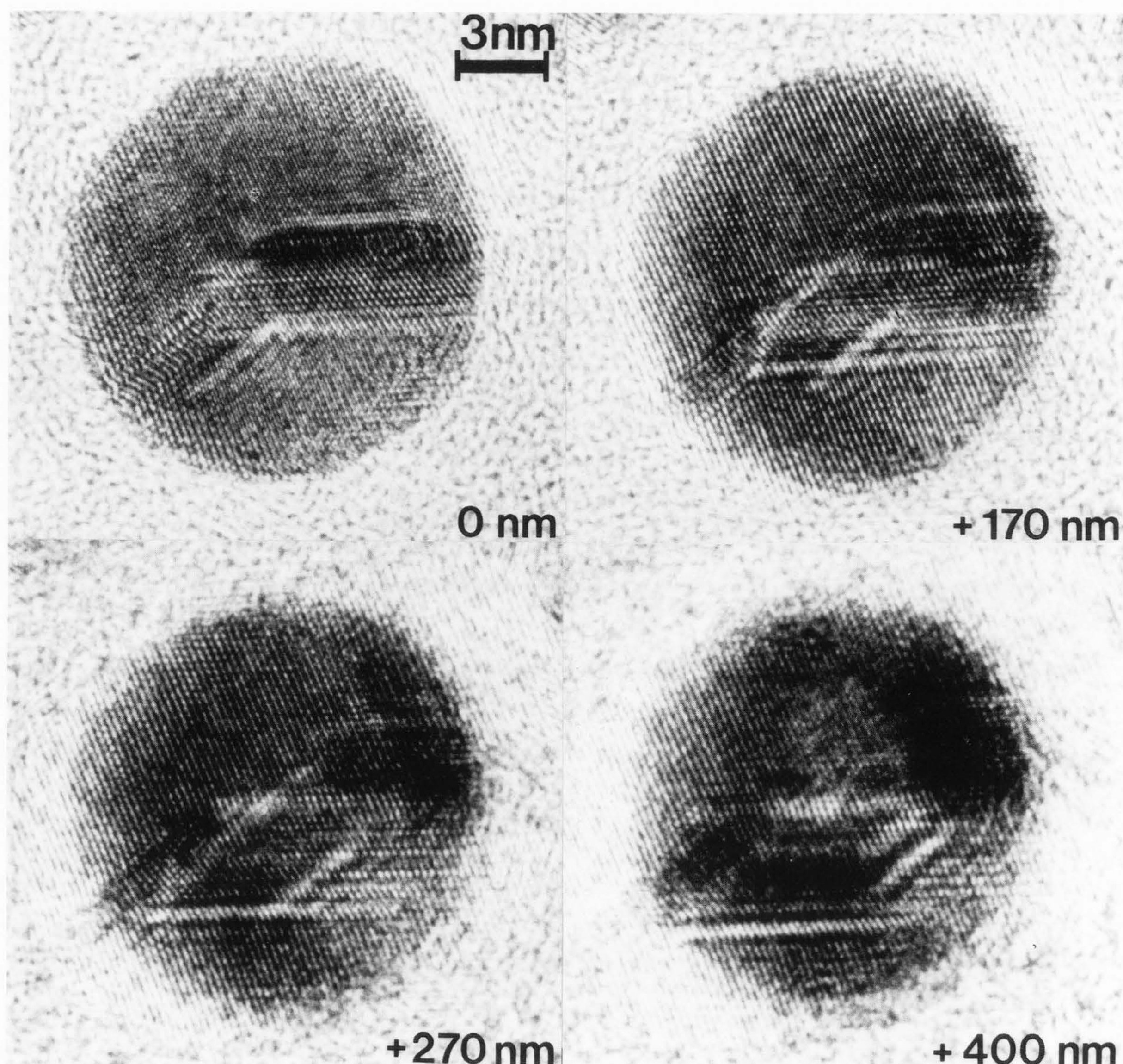

Fig. 5. Continued.

leading terms in the asymptotics are given by :

$$t^{(II)}(R') \cong \sqrt[3]{3}\sqrt{\frac{2\pi^2}{C_s\lambda^3}}\,|\Delta f|^{-2/3}$$

$$\times\ \exp\left\{i\left(\frac{\pi}{4}-\frac{1}{12}\Delta f^2\right)\right\}$$

$$\times\ \exp\left\{-i\sqrt{\frac{1}{6}|\Delta f|}\,(R'-R'_c)\right\}$$

$$\times\ \mathrm{Ai}\left(\frac{R'-R'_c}{\sqrt{2}\,(3|\Delta f|)^{1/6}}\right) \tag{29}$$

with Ai the Airy function (see appendix, Fig. A.1). The maximum amplitude of t(R') in the second region (II) amounts to about

$$\sqrt[3]{\frac{3}{|\Delta f|^2}}\sqrt{\frac{2\pi}{C_s\lambda^3}}\,.$$

Close to the caustic point, the oscillations in the IRF arise only from the exponential factor with a frequency $1/(2\pi)\sqrt{1/6\,|\Delta f'|}$, which yields in unscaled coordinates the value of G_{min} in (16b). It is instructive to make a comparison with the results of the preceeding section. At the latter frequency, $|\Delta R|$ in (14) is at maximum, and the delocalisation is identical to the caustic point (28b). Further away

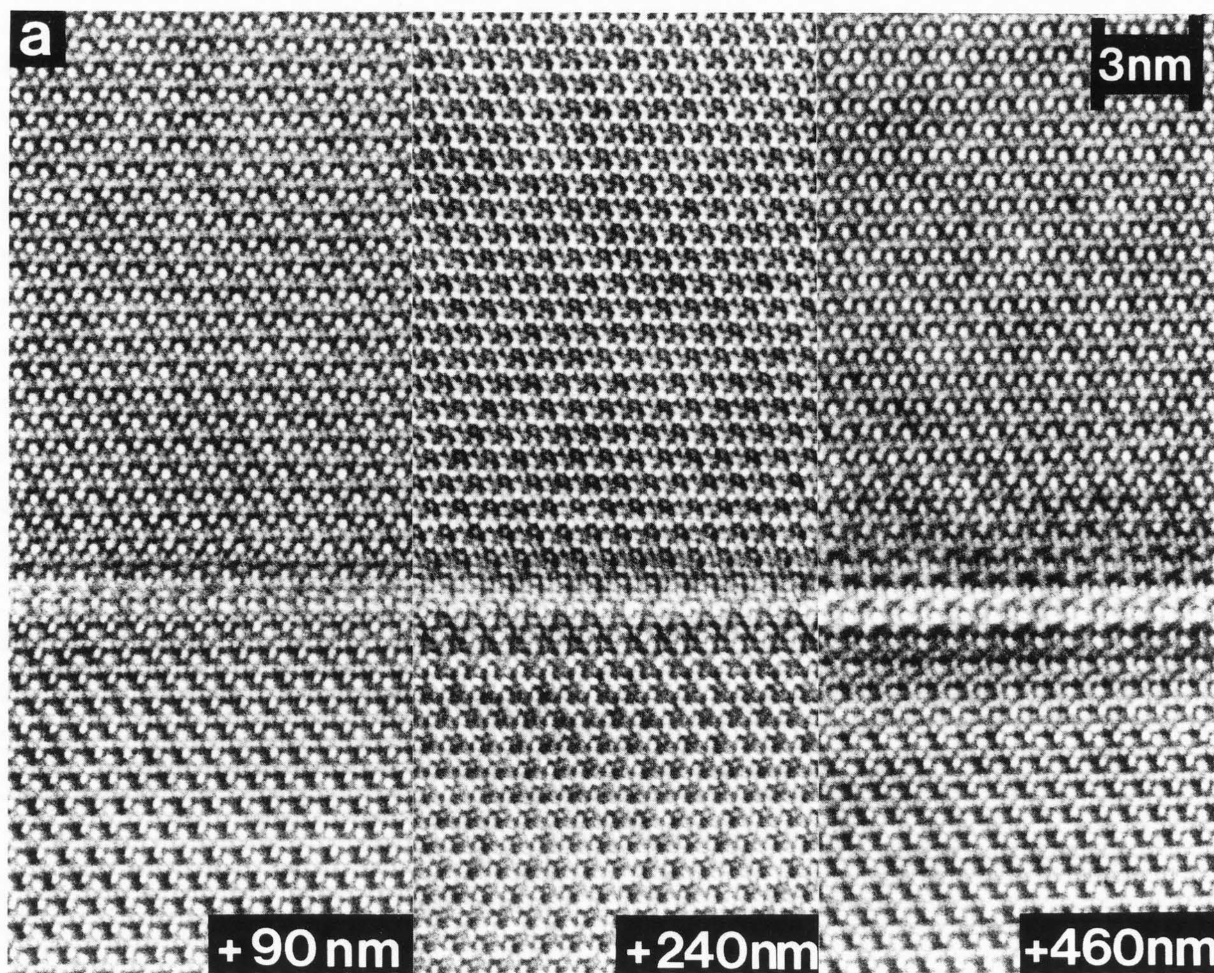

Fig. 6. Through-focus HRTEM image series of a stacking fault in [001] oriented Si_3N_4 (CM20 FEG-TWIN). (a) (large) overfocus

at the lower side from the caustic point, the oscillations of the Airy function start to interfere with those from the exponential factor, giving rise to a more complicated oscillation mode, which transforms rather smoothly into the J_o - mode of (27) close to the origin. Especially for smaller underfocus, the point of transition between regions (I) and (II) is less obvious. Since the main frequency around the caustic G_{min} equals $G_2 / \sqrt{3}$, with G_2 the main frequency around the origin $R' \approx 0$, it is obvious that the attenuation due to E_Δ is larger around the origin than at the caustic. In other words, an increase in the focus spread parameter Δ due to a worse temporal coherence, would relatively increase the amplitude at the caustic over the amplitude around the origin, and therefore, the effect of delocalisation is increased.

In the third region (III), beyond the caustic point $R' \geq R'_c$, we obtain the asymptotic approximation :

$$t^{(III)}(R') \cong \sqrt{\frac{2\pi}{C_s \lambda^3 |\Delta f|^2}} \times \exp\left\{-i\left(\frac{\pi}{4} - \frac{2}{3}\Delta f^2\right)\right\} \times \exp\left\{i\sqrt{\frac{2}{3}|\Delta f|}\,(R' - R'_c)\right\}. \tag{30}$$

Hence, the main frequency beyond the caustic point is $1/(2\pi)\sqrt{2/3\,|\Delta f'|}$, which yields in unscaled coordinates a value of $2G_{min}$ (with G_{min} from (16b)). This frequency lies beyond G_2 defined in (17). For large underfocus, this means that the amplitude of the oscillations in the third region $R' \geq R'_c$ are severely damped by both E_Δ and E_S. This effect is clearly observed in the numerically obtained IRF's of Fig. 4. Apart from the damping, the maximum

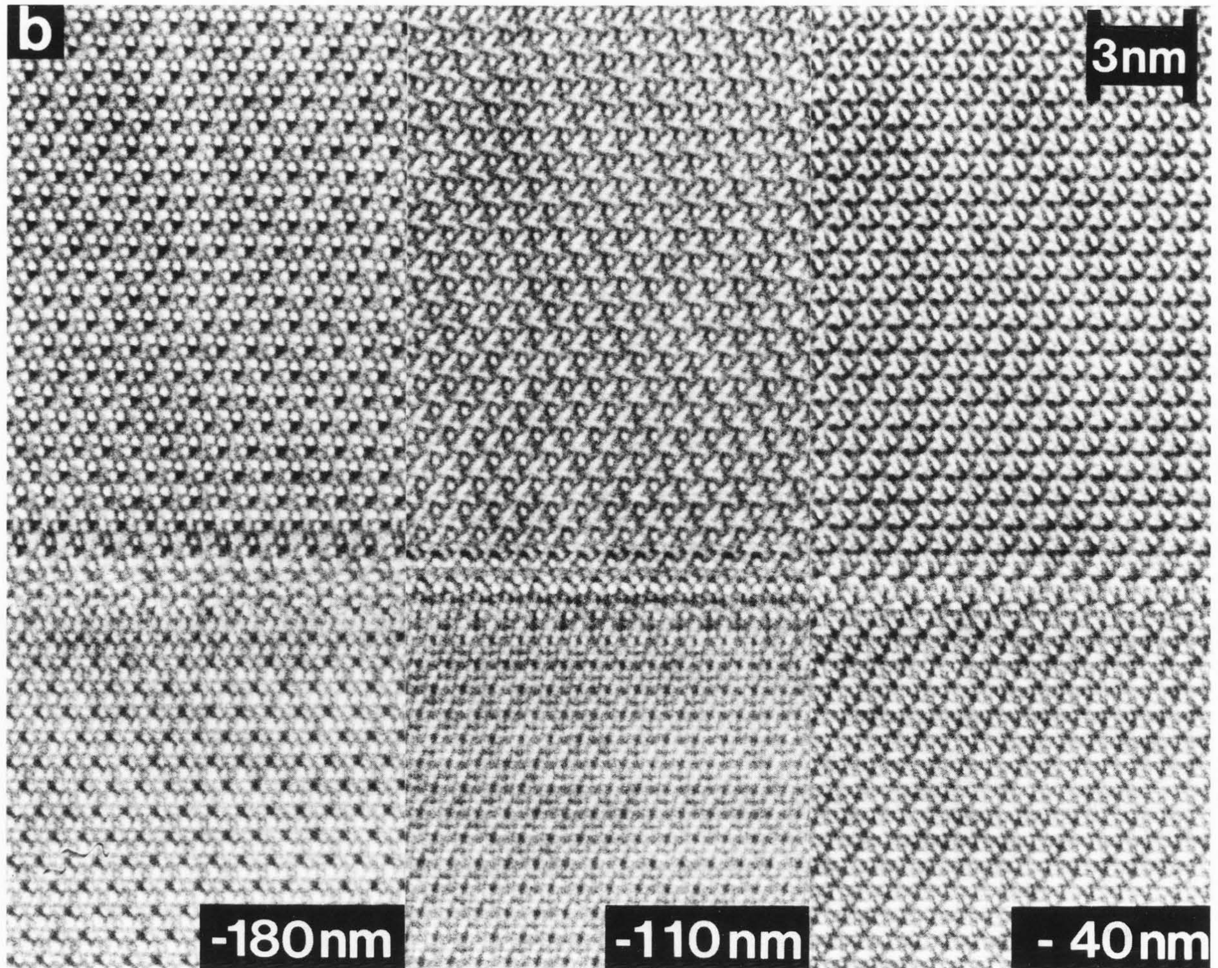

Fig. 6. Continued (b) normal underfocus

amplitude of $t^{(III)}$ amounts to

$$\frac{1}{|\Delta f|}\sqrt{\frac{2\pi}{C_s \lambda^3}},$$

which is lower than for $t^{(II)}$.
Taking these considerations into account, the caustic point in (28) turns out to be a good measure for the delocalisation as a function of defocus, yielding a 3/2 power law.

Experimental Results

HRTEM observations have been carried out for Au particles on a carbon film with a Philips CM20 FEG-SuperTWIN microscope (C_s = 1.2 mm), and with a CM20 FEG-TWIN microscope (C_s = 2.0 mm) for a Si_3N_4 crystallite, and for a thin amorphous Ge-film.

Au-particle on Carbon Film

In Fig. 5, a through-focus series recorded with a CM20 FEG-SuperTWIN of a Au particle in $[1\bar{1}0]$ orientation is shown with focus values ranging from large underfocus ($\Delta f\sim$ -470nm) to large overfocus ($\Delta f\sim$ +400nm). The Scherzer image ($\Delta f_S\sim$ -66 nm), which yields the most localised information about the structure, reveals a small twinned area extending from the bottom-left to middle-right area of the particle. The overall particle is imaged rather close to the exact $[1\bar{1}0]$ zone-axis orientation. The twinned area suffers from a small crystal tilt so that in the Scherzer image, the (002) fringes become more apparent than the $(11\bar{1})$ fringes, running inclined with respect to the (111) twin boundaries (TB). The HRTEM delocalisation becomes obvious by means of the displacements of the TB image, and by the appearance of "lattice" fringes (or better : "image" fringes) outside the particle.

According to the spatial frequency analysis of the delocalisation, each TB is split up in the image into two separate boundary planes (shadow images) with a frequency-dependent shift of $\pm\nabla\chi(\mathbf{G}_B)$ from the 'real' TB position. $\mathbf{G}_B$ is then a Bragg vector of the Au lattice, both within the particle and within the twin area, and is of the type $\pm(11\bar{1})$, $\pm(002)$. For the twin area, the (002) reflection is well excited,

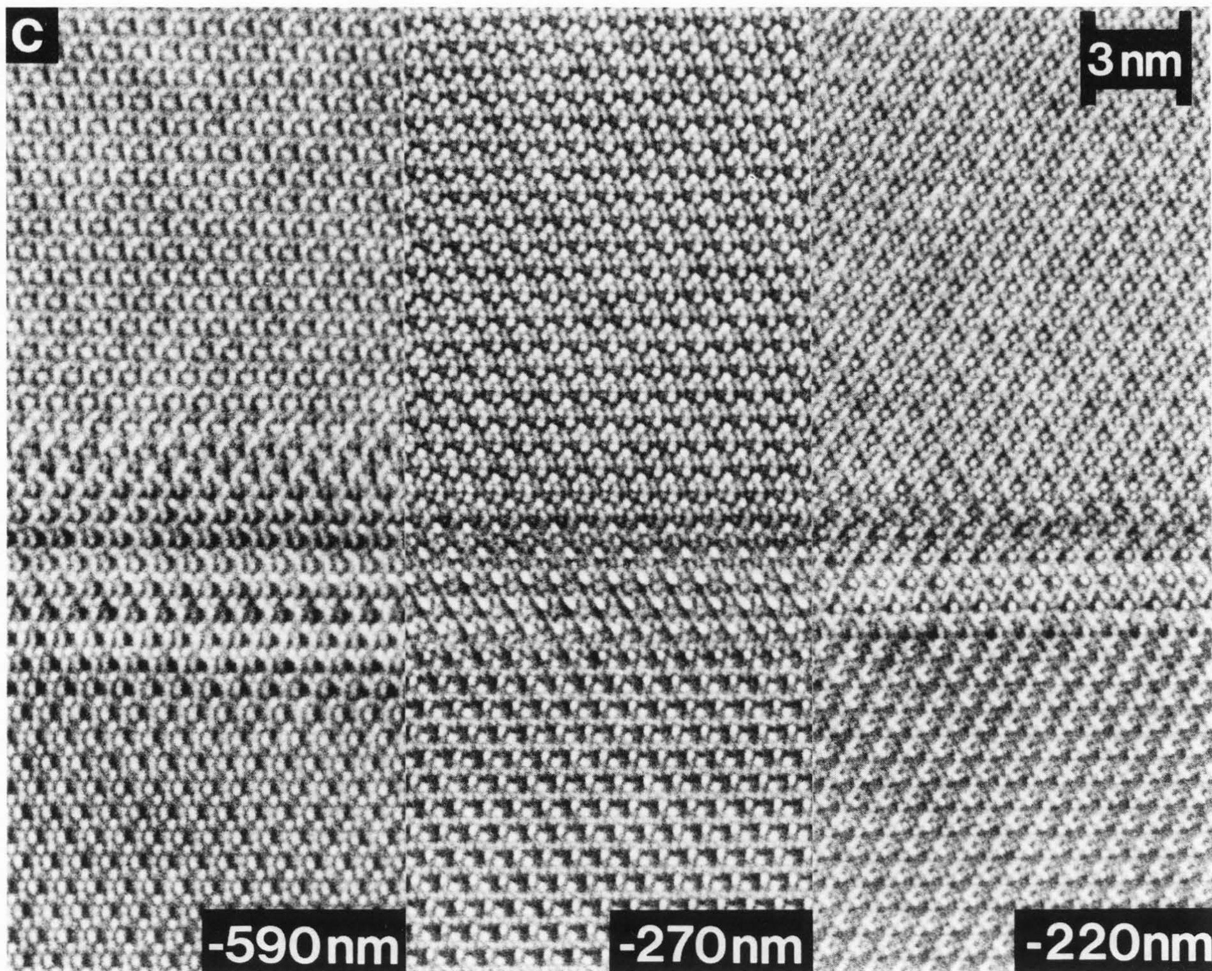

Fig. 6. Continued (c) large underfocus

and for the remaining area, the fringes corresponding to the $(11\bar{1})$ spot are somewhat emphasised, and these two vectors chiefly determine the location of the shadow images of the TBs according to relation (14). Due to the deviation from the $[1\bar{1}0]$ zone axis, the shadow images of the TBs turn out to be asymmetric in intensity, since the intensities of the inversion related ±(002) and $\pm(11\bar{1})$ diffraction spots are slightly differing. The area between the TB images is only locally imaged at Scherzer and at zero defocus. At the other focus settings, a complicated mixture of the "average" matrix (V_1+V_2) and the "difference" matrix (V_1-V_2) is imaged in the broadened defect region of each TB (cfr. relation (12)). Therefore, the contrast in this defect area exhibits typical mixture patterns, which are not easily interpretable. On the other hand, the stacking sequence of the perfect Au matrix in the remainder of the particle can still be revealed for the case of large under- or overfocus conditions as well.

Si_3N_4 Crystallite

Fig. 6 shows a through-focus series, recorded with a CM20 FEG-TWIN, of a stacking fault in a Si_3N_4 crystallite in [001] orientation. The stacking fault can be revealed more easily when looking under grazing incidence along the bottom-left-to-top-right diagonal. The image of the defect is most localised close to Scherzer focus (Δf_S~ -87 nm), that is for the focus values Δf~-110 nm and Δf~-40 nm in Fig. 6. The width of the defect area in the image obviously increases when going to high under- or overfocus settings. The defect area may be identified as the region where the contrast of the perfect matrix is disturbed. The maximum delocalisation as measured in the focal series of Fig. 6 amounts to 4 nm for Δf~-590 nm, which is in agreement with the theoretical delocalisation as indicated by the caustic point. Note also the high resolution capabilities of the FEG for the CM20-TWIN, which are illustrated by means of the high resolution contrast with fine image details that is obtained over a focus range of more than 1 μm. This is only possible due to the high coherence of the FEG. On the other hand, precise focussing with a FEG becomes more difficult if no specimen edge or defect is present within the field of view.

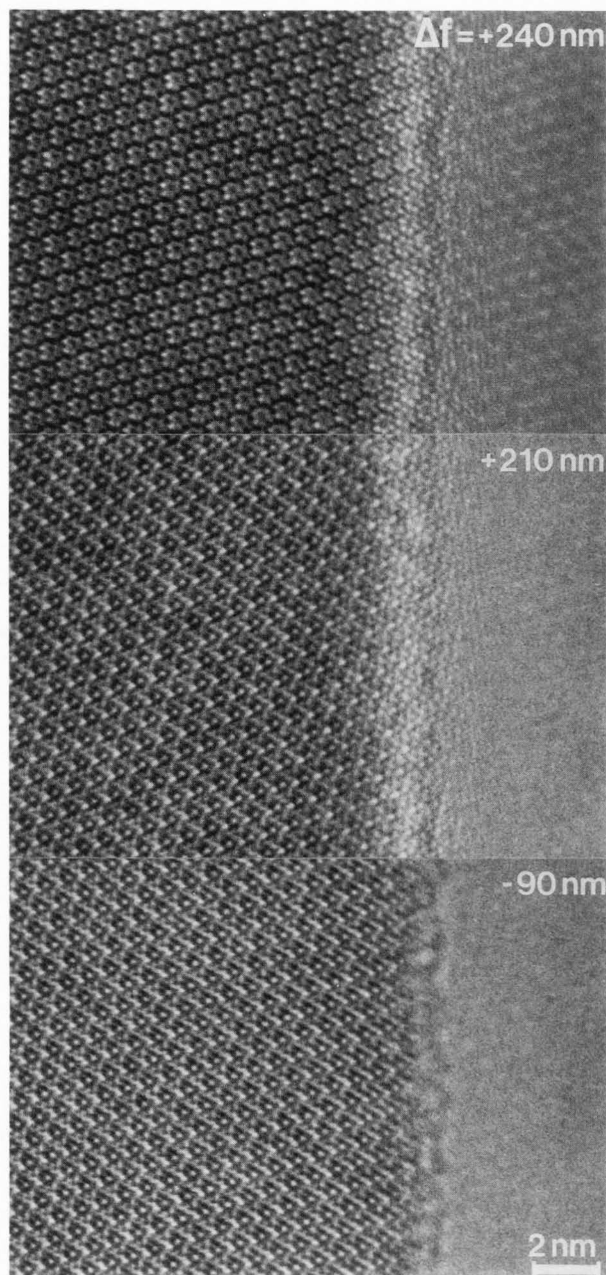

Fig. 7. HRTEM delocalisation at the edge of a sample in [001] oriented Si_3N_4 (CM20 FEG-TWIN).

Fig. 7 shows some HRTEM images from the edge of a Si_3N_4 grain. The structure of the grain is crystalline up to the very edge, i.e. almost no amorphous area due to ion-milling or contamination could be detected. Note that for large (over)focus, the fringe contrast extends clearly in the vacuum, beyond the edge of the sample. For $\Delta f\sim+240$ nm, a white dot pattern can be observed extending at least 5 nm away from the edge. The HRTEM images for both foci $\Delta f\sim+210$ nm and $\Delta f\sim-90$ nm (roughly at Scherzer focus) are "Fourier" images (Cowley and Moodie, 1957; Iijima and O'Keefe, 1979), i.e. they show an identical image contrast for the perfect Si_3N_4 matrix. However, for the Scherzer image ($\Delta f\sim-90$ nm), the image is well localised at the edge, whereas for the overfocus image ($\Delta f\sim+210$ nm), the image fringes are extending over the edge with a delocalisation of more than 2 nm. This phenomenon can also be described as a high-resolution Fresnel-fringe effect at the crystal edge, but now due to both defocus and spherical aberration. Numerical calculations of the high resolution Fresnel effect have been reported by Wilson, Bursill and Spargo (1978). Note that also in Fig. 6, the images $\Delta f\sim+90$ nm and $\Delta f\sim+460$ nm are Fourier images, with similar contrast for the perfect matrix, but with a different defect width and contrast at the stacking fault.

Thin Amorphous Ge-film

In Figs. 8-9, rotationally averaged optical diffractograms and numerical auto-correlations are shown obtained from a through-focus series of a thin amorphous Ge-film (CM20 FEG-TWIN). The defoci belong to the "passband" underfocus values, which are defined in (15). At high underfocus, the diffractograms reveal the large number of oscillations in the transfer function that can be obtained experimentally on the CM20 FEG-TWIN due to the high coherence. At Scherzer focus (n=0), and at moderate underfocus, the diffractograms also show a number of closely spaced rings beyond the passbands (e.g. more than eight CTF zero's can be observed for n=0, beyond the Scherzer plateau).

The auto-correlation functions (ACFs) of the images of the Ge - film yield a rough experimental measure of the delocalisation. Indeed, the ACF in the WPO approximation is given by :

$$\begin{aligned} ACF(\mathbf{R}) &= \int I_{im}(\mathbf{R}')\, I^*_{im}(\mathbf{R}+\mathbf{R}')\, d\mathbf{R}' \\ &= FT^{-1}\left\{ \left| \hat{I}_{im}(\mathbf{G}) \right|^2 \right\} \\ &\approx FT^{-1}\left\{ \left| \hat{V}(\mathbf{G}) \right|^2 \hat{t}_I(\mathbf{G})^2 \right\} . \end{aligned} \qquad (31)$$

As shown by Al-Ali and Frank (1980), the ACF then represents the cross-correlation of the ACF of the specimen potential V, with the ACF of the IRF t_I. $|\hat{V}(\mathbf{G})|^2$ represents the short-range correlations in the arrangement of the atoms in the amorphous Ge, and therefore, the ACF of V has a finite width. In the case of an ideal δ-function shaped object, $|\hat{V}(\mathbf{G})|^2 = 1$, the ACF of the image yields the inverse FT of the squared value of $\hat{t}_I$. On the other hand, in the case of an ideal microscope $\hat{t}_I=1$, the ACF of the image would yield the short-range correlation of the amorphous structure. The latter situation is most closely approached at Scherzer or optimum focus (n=0).

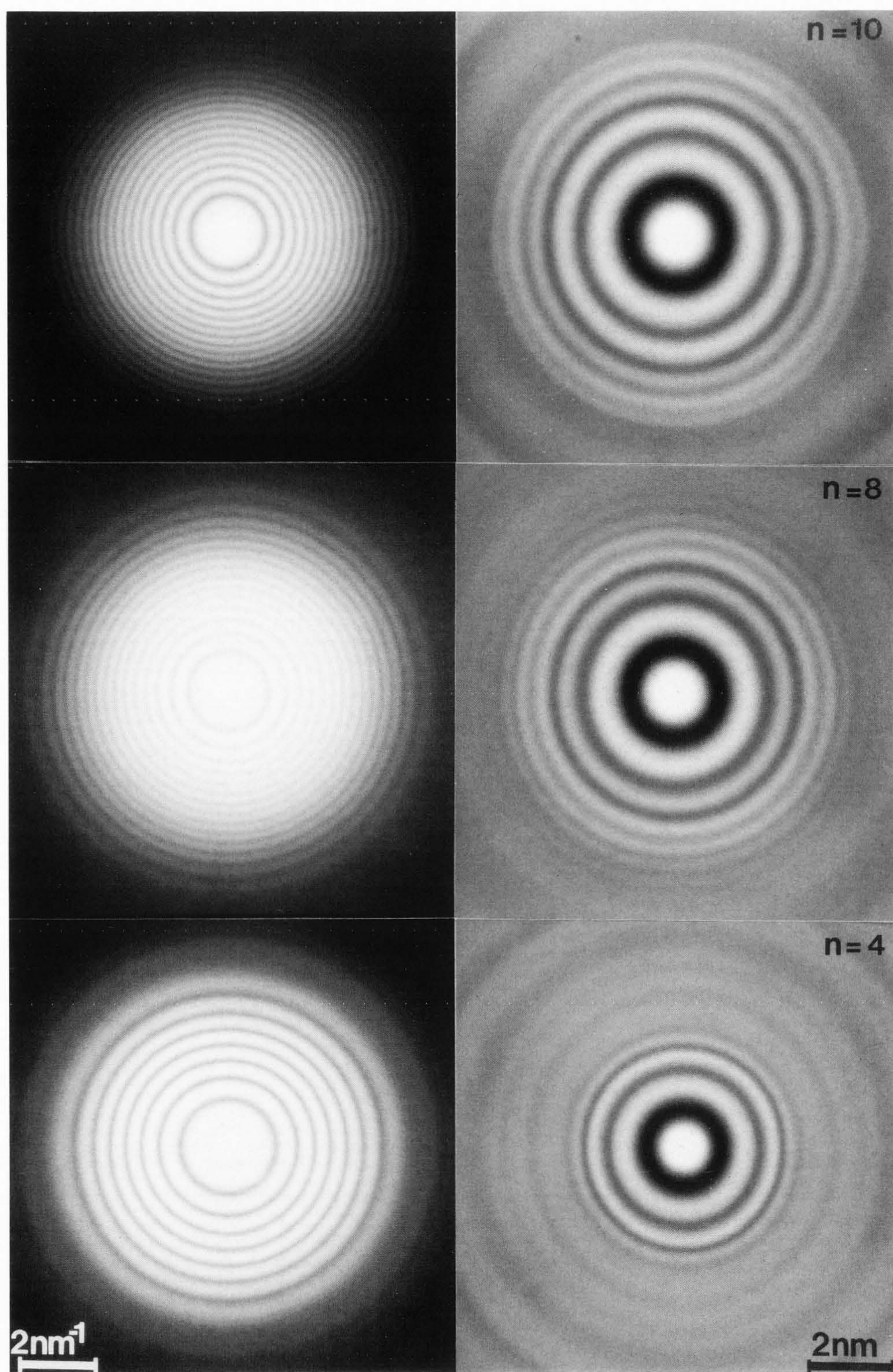

Fig. 8. Rotationally averaged optical diffractograms (a) and numerical auto-correlations (b) of a through-focus HRTEM image series of a thin amorphous Ge - film. The focus values are chosen as the "passband" focus values $\Delta f_S(n)$ from relation (15), with n indicated (CM20 FEG-TWIN).

In practice, however, it is not possible to separate microscope and specimen easily. As a function of defocus, the distance of the rings to the origin in the ACFs of Fig. 8 increases together with their amplitude relative to the central correlation peak.

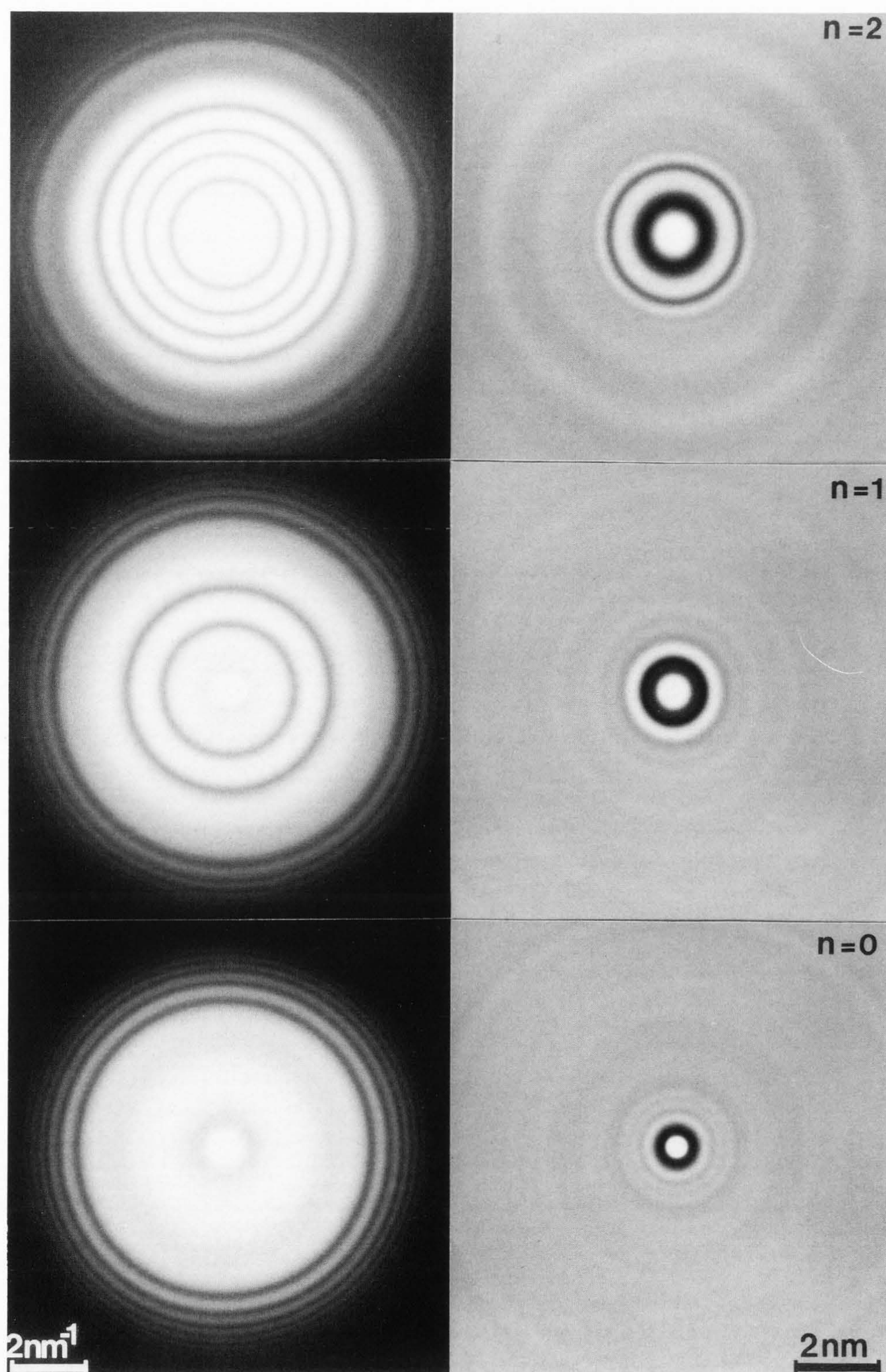

Fig. 8. Continued.

The outward shift of the correlation rings as a function of underfocus can be explained by the correlation of the common short-range structural correlation function with $FT^{-1}\{t_I^2\}$, which is roughly proportional to the real part of the IRF $t_R(R;2C_s;2\Delta f)$ for the double value of spherical aberration and defocus.

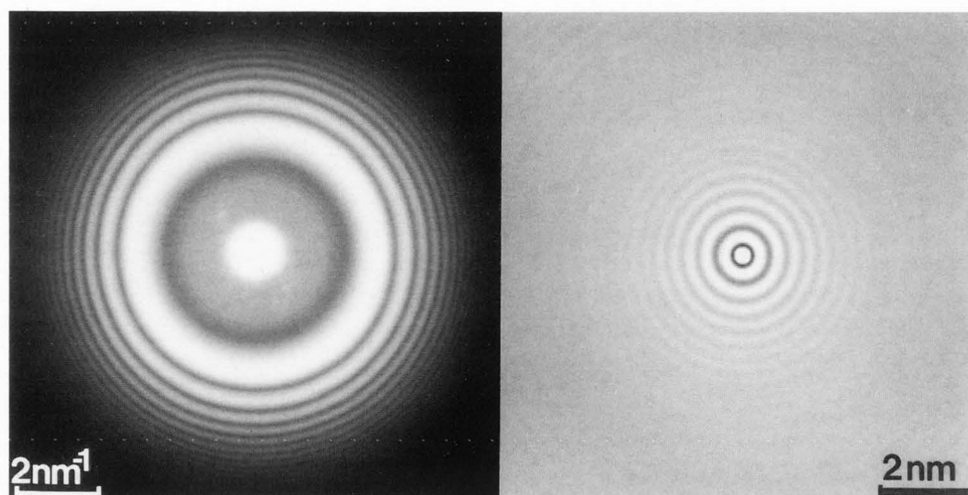

Fig. 9. Rotationally averaged optical diffractogram (a) and numerical auto-correlation (b) of an HRTEM image of a thin amorphous Ge - film taken at Gaussian focus (CM20 FEG-TWIN).

Concluding Discussion

In this paper, we have discussed the parameters which affect image (de)localisation in HRTEM with application to a field emission gun of high coherence. Due to its high spatial coherence, HRTEM lattice resolution with fine image details is observed over a (much) larger defocus range than using an LaB_6 thermionic emitter; the HRTEM images also carry information from *all* frequencies up to the information limit (determined by the temporal coherence). Indeed, the information from frequency regions where the microscope's transfer function is rapidly oscillating will still survive in the image thanks to the low divergence of the illuminating beam (or small size of the effective electron source), which can be realised due to the high brightness of the FEG. However, as a minor drawback, one has to take into account that image information may then become delocalised, with the delocalisation proportional to the oscillation speed of the transfer function in a particular frequency range. A good measure for the delocalisation in the case of underfocus is derived in terms of the caustic point. Also, the effect of a residual beam misalignment may be enlarged for a better spatial coherence. Stated in other words, the delocalisation is due to the gap between the high information limit of the FEG and the point resolution of the microscope, which cannot much further be improved due to practical technological limitations in the design of the objective lens polepieces. Image interpretation for HRTEM with a FEG, especially for images of crystal defects and "lattice discontinuities" may thus become (even) less straightforward than in the case of an LaB_6 emitter. The classical way of obtaining unambiguous structural information out of the images may then be by means of image simulation. A direct compensation for the microscope transfer may be obtained by HRTEM image reconstruction using focal image series (Kirkland, 1984; Van Dyck and Op de Beeck, 1990), or by means of electron holography (Lichte, 1986; Lichte, 1991a). For practical HRTEM with the CM20 FEG in materials science, and without relying on image simulation, care must be taken to record images close to Scherzer focus so that delocalisation effects are reduced.

Thus, delocalisation effects become more prominent and beam alignment is more critical if one wants to reach the improved resolution performance when running the FEG in a highly spatially coherent mode. The latter is an essential condition in holography, where the amplitude of the interference fringes is determined by the complex degree of coherence (Born and Wolf, 1975), and requires a low divergence of the illuminating beam. The delocalisation in the hologram is then corrected in the digital reconstruction step. In routine HRTEM work on the other hand, when recording a few focal images, one wants to have in the image the disposal of the information from the whole frequency range up to a maximal resolution G_{max} . Then, at larger underfocus, image details corresponding with this maximal resolution are optimally recorded, but in order to avoid severe damping of the spatial coherence envelope E_S at intermediate frequencies, it is necessary to run the FEG with a low beam divergence α.

The situation is different for HRTEM image reconstruction using focal image series with a larger number of images. Indeed, the electron dose increases with α^2, and therefore, the signal-to-noise ratio in the reconstructed image improves with increasing α. This makes it attractive to run the FEG for image reconstruction in a more focused mode (higher α) with an intentionally reduced spatial coherence. At the same time, delocalisation effects and the sensitivity to beam misalignment will be reduced. In such a case, the spatial

coherence envelope performs the role of a frequency selecting band-filter in each focal image, with the selected frequency range depending on the actual focus value.

Appendix : Asymptotic Analysis of Impulse-Response Function for Large Underfocus

In this Appendix, we present the asymptotic behaviour of the impulse-response function t(R') when $-\Delta f \to \infty$. The results we present appear to be accurate already for modest values of $-\Delta f$, to the extent that they give a useful qualitative insight even for $-\Delta f=0$.
We consider the integral

$$I(\Delta f, R') = 2\int_0^\infty \exp\left\{-i\left(v^4+\Delta f\, v^2\right)\right\} J_0(vR')\, v\, dv \tag{A.1}$$

for large values of $-\Delta f$ (the factor 2 has been included for better agreement with existing conventions). The integral (A.1) belongs to a class of Pearcey-type integrals,

$$I_\alpha(\Delta f, R') = 2\int_0^\infty \exp\left\{-i\left(v^4+\Delta f\, v^2\right)\right\} J_\alpha(vR')\, v^{\alpha+1} dv \tag{A.2}$$

with $-1<\alpha<5/2$, the asymptotics of which has been given by Janssen (1992), following the approach of Paris (1991). The case with $\alpha=-1/2$ gives the classical Pearcey integral,

$$I_{-\frac{1}{2}}(\Delta f, R') = \sqrt{\frac{2\pi}{R'}}\int_0^\infty \exp\left\{-i\left(v^4+\Delta f\, v^2\right)\right\} \cos(vR')\, dv \tag{A.3}$$

which has been studied extensively in recent years (Connor and Curtis, 1982; Kaminski, 1989; Paris, 1991). In the present context, this $I_{-1/2}$ would describe the image formation for the hypothetical case of one-dimensional electron microscopy, in very much the same way as I_0 does this for real two-dimensional and radially symmetric microscopy.
The results of Paris (1991) indicate that values of $-\Delta f \geq 4$ must already be considered large (our $-\Delta f$ ranges between 0 and 50, or more).

We shall now present the results of Janssen (1992) as far as relevant for the present paper, i.e. we take $\alpha=0$ in (A.2) and we consider the behaviour of $I(\Delta f, R')$ as $-\Delta f \to \infty$ in the cases

(a) $R' \geq 0$, fixed,

(b) $R' = \rho\left(\frac{2}{3}|\Delta f|\right)^{3/2}$ and $\rho>0$, fixed.

The results for case (a) are useful to describe the behaviour of $I(\Delta f, R')$ for R' relatively close to 0. The results for case (b) are useful to describe the behaviour of $I(\Delta f, R')$ near the caustic

$$R'_c = \left(\frac{2}{3}|\Delta f|\right)^{3/2}$$

It thus turns out that $I(\Delta f, R')$ is described accurately by the sum of a Bessel function and an exponential with quadratic phase for small R', and by the sum of a modulated Airy function and an exponential with linear phase near the caustic.
We have in case (a) the asymptotic expansion :

$$I(\Delta f, R') \approx \sqrt{\pi}\exp\left\{-\frac{1}{4}\pi i+\frac{1}{4}i(\Delta f)^2\right\} \times \sum_{m=0}^{\infty}\frac{\left(i\frac{(R')^2}{8\Delta f}\right)^m}{m!} J_{2m}\left(R'\sqrt{-\frac{1}{2}\Delta f}\right) + \frac{i}{-\Delta f}\exp\left\{i\frac{(R')^2}{4\Delta f}\right\} \times \sum_{m=0}^{\infty}\frac{(2m)!\, i^m}{m!(\Delta f)^{2m}} L_{2m}\left(-i\frac{(R')^2}{4\Delta f}\right) \tag{A.4}$$

where J_{2m} and L_{2m} are the Bessel functions and Laguerre polynomials, respectively. The leading order asymptotics is given by

$$I(\Delta f, R') \cong \sqrt{\pi}\exp\left\{-\frac{1}{4}\pi i+\frac{1}{4}i(\Delta f)^2\right\} J_0\left(R'\sqrt{-\frac{1}{2}\Delta f}\right) - \frac{i}{\Delta f'}\exp\left\{\frac{i}{4}\frac{(R')^2}{\Delta f}\right\} \tag{A.5}$$

We have found this approximation to be adequate for the purposes of this paper for $R' \leq (-8\pi\ \Delta f')^{0.5}$, so that one complete oscillation of the second term is covered. When R' is larger, or higher accuracy is required, one should incorporate more terms of the two series in (A.4). For instance, the corresponding expansion of Paris (1991) for $I_{-1/2}$ has six correct decimal places when $-\Delta f \geq 10$, $R' \leq 0.5\ |\Delta f|$, and six terms in either series are used.

We have in case (b) the leading order asymptotics :

$$I(\Delta f', R') \cong$$

$$\left(\frac{\pi}{\rho}\right)^{\frac{1}{2}} \exp\left\{\frac{1}{4}\pi i - \delta(\Delta f')^{2}\right\}$$

$$\times \left[\frac{c_0}{|\Delta f'|^{\frac{2}{3}}} Ai\left(\gamma^2 |\Delta f'|^{\frac{4}{3}}\right) + \frac{i\, c_1}{|\Delta f'|^{\frac{4}{3}}} Ai'\left(\gamma^2 |\Delta f'|^{\frac{4}{3}}\right)\right]$$

$$+ \frac{1}{i\,\Delta f'\,\rho^{\frac{1}{2}}} \exp\left\{\varepsilon(\Delta f')^{2}\right\} \tag{A.6}$$

where, with $\beta = -2/3 \ln \rho$,

$$\delta = \frac{1}{12} i - \frac{1}{6} i\beta + \frac{5}{72} i\beta^2 + O\left(\beta^3\right),$$

$$\gamma = 3^{-\frac{1}{3}} i\left(\frac{1}{2}\beta\right)^{\frac{1}{2}} + O\left(\beta^{\frac{3}{2}}\right),$$

$$\varepsilon = -\frac{2}{3} i + \frac{1}{3} i\beta - \frac{5}{36} i\beta^2 + O\left(\beta^3\right),$$

$$c_0 = 3^{\frac{1}{3}} + O(\beta),$$

$$c_1 = 3^{-\frac{1}{3}} + O\left(\beta^{\frac{1}{2}}\right). \tag{A.7}$$

We have plotted the Airy function Ai and its derivative Ai' in Fig. A.1, both for positive and negative argument (see Abramowitz and Stegun, 1970).

In (A.6) the term with Ai' is significantly smaller than both the term with Ai and the term with $\exp\{\varepsilon \Delta f'^2\}$. Ignoring this smaller term, and using the leading order approximations in (A.7), we get for small values of $|(R'-R'_c) / R'_c|$, $\rho \approx 1$, i.e. near

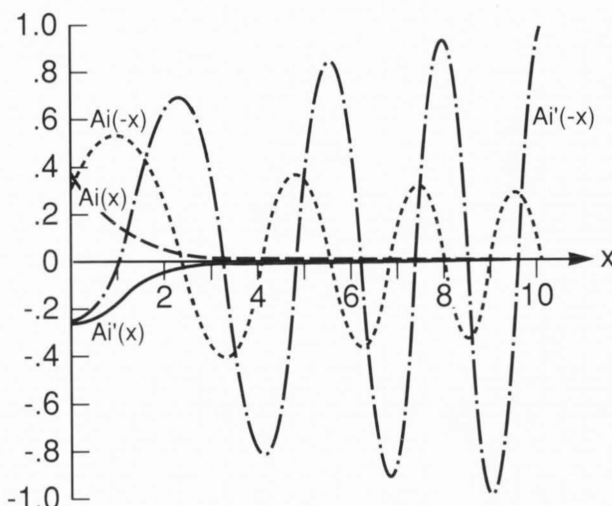

Fig. A.1. The Airy function Ai(±x) and its derivative Ai'(±x), for positive and negative argument, from Abramowitz and Stegun (1970).

the caustic :

$$I(\Delta f', R') \cong \frac{3^{1/3}\pi^{1/2}}{|\Delta f'|^{2/3}} \exp\left\{\frac{1}{4}\pi i - \frac{1}{12} i\,\Delta f'^2\right\}$$

$$\times \exp\left\{-i\sqrt{\frac{1}{6}|\Delta f'|}\ (R' - R'_c)\right\}$$

$$\times Ai\left(\frac{R' - R'_c}{\sqrt{2}\ (3|\Delta f'|)^{1/6}}\right)$$

$$-\frac{i}{|\Delta f'|} \exp\left\{\frac{2}{3} i\,\Delta f'^2\right\}$$

$$\times \exp\left\{i\sqrt{\frac{2}{3}|\Delta f'|}\ (R' - R'_c)\right\}. \tag{A.8}$$

It is clear from Fig. A.1 that the first term in the right hand-side of (A.8) is dominant for $R'<R'_c$ while the second term is dominant for $R'>R'_c$. Formula (A.8) is only accurate for $|(R'-R'_c) / R'_c|$ sufficiently small ; for a larger region of validity one should use more than only the leading terms in the expansions in (A.8) and also incorporate the terms with Ai'. However, the resulting asymptotic formula gets then quite complicated.

Acknowledgements

The authors wish to thank Dr. M.T. Otten (Philips Analytical, Electron Optics Application Laboratory) for providing the focal image series of

the Au - particle. Dr. J. Thibault-Desseaux (CEN, Grenoble) is acknowledged for providing the thin amorphous Ge sample. Dr. M.T. Otten, Dr. A.F. de Jong (Philips Research Labs.), Prof. Dr. D. Van Dyck and Drs. M. Op de Beeck (University of Antwerp, RUCA) are acknowledged for stimulating discussions and useful suggestions. The authors thank Ing. F. Hakkens (Philips Research Labs.) for preparation of the TEM samples. Thanks are also due to the Philips Electron Optics Application Laboratory for the use of the CM20 FEG-TWIN microscope.

References

Abramowitz M, Stegun IA. (1970). Handbook of Mathematical Functions, Dover, New York, 9th Printing, 446.

Al-Ali L, Frank J. (1980). Resolution Estimation in Electron Microscopy. Optik 56, 31-40.

Bakker JG, Asselbergs PES. (1990). An Ultra-High-Resolution Objective Lens for a 200 kV TEM, in Proc. of XIIth International Congress for Electron Microscopy, LD Peachley and DB Williams (eds), San Francisco Press, San Francisco, Vol. 1, 132-133.

Born M, Wolf E. (1975). Principles of Optics, Pergamon, London, Chapter X.

Coene W, Van Dyck D. (1988). New Aspects in Nonlinear Image Processing for High Resolution Electron Microscopy, in Proc. of 6th Pfefferkorn Conference, Scanning Microscopy Supplement 2, 117-129.

Cowley JM, Iijima S. (1972). Electron Microscope Image Contrast for Thin Crystals. Z. Naturforsch. 27a, 445-451.

Cowley JM, Moodie AF. (1957). Fourier Images I-The Point Source. Proc. Phys. Soc. 70, 486-496.

Connor JNL, Curtis PR. (1982). A Method for the Numerical Evaluation of the Oscillatory Integrals Associated with the Cuspoid Catastrophes : Application to Pearcey's Integral and its Derivatives. J. Phys. A15, 1179-1190.

Frank J. (1973). The Envelope of Electron Miscroscopic Transfer Functions for Partially Coherent Illumination. Optik 38, 519-536.

Goodman JW. (1968). Introduction to Fourier Optics, Mc Graw-Hill, New York, Chapters 5-6.

Iijima S, O'Keefe MA. (1979). Determination of Defocus Values using 'Fourier' Images for High Resolution Electron Microscopy. Journal of Microscopy 117, 347-354.

Ishizuka K. (1980). Contrast Transfer of Crystal Images in TEM. Ultramicroscopy 5, 55-65.

Janssen AJEM. (1992). On the Asymptotics of some Pearcey-type Integrals. Journal of Physics A : Mathematical and General, accepted for publication.

Kaminski D. (1989). Asymptotic Expansion of the Pearcey Integral near the Caustic. SIAM J. Math. Analysis 20, 987-1005.

Kirkland EJ. (1984). Improved High Resolution Image Processing of Bright Field Electron Micrographs, I. Theory. Ultramicroscopy 15, 151-172.

Lichte H. (1986). Electron Holography Approaching Atomic Resolution. Ultramicroscopy 20, 293-304.

Lichte H. (1991a). Electron Image Plane Off-axis Holography of Atomic Structures. Advances in Optical and Electron Microscopy 12, 25-91.

Lichte H. (1991b). Optimum Focus for taking Electron Holograms. Ultramicroscopy 38, 13-22.

Marks LD. (1984a). Dispersive Equations for High Resolution Imaging and Lattice Fringe Artifacts. Ultramicroscopy 12, 237-242.

Marks LD. (1984b). Direct Atomic Imaging of Solid Surfaces : I. Image Simulation and Interpretation. Surface Science 139, 281-298.

Marks LD. (1985). Image Localisation. Ultramicroscopy 18, 33-38.

Mul PM, Bormans BJ, Otten MT. (1991). Design of the CM20 FEG. Electron Optics Bulletin 130, 53-62.

O'Keefe MA. (1979). Resolution-damping Functions in non-linear Images, in Proc. of 37th Ann. Meeting of EMSA, G.W. Bailey (ed), 556-557.

O'Keefe MA, Saxton WO. (1983). The "well-known" Theory of Electron Image Formation, in Proc. of 41st Ann. Meeting of EMSA, G.W. Bailey (ed), San Francisco Press, San Francisco, 288-289.

O'Keefe MA. (1984). Electron Image Simulation : a Complementary Processing Technique. Proc. 3rd Pfefferkorn Conference on Electron Optical Systems, 209-220.

Otten MT, Bakker JG. (1991). The CM20-UltraTWIN: a 200 kV Analytical Ultra-High Resolution Microscope. Beiträge zur electronenmikroskopischen Direktabbildung und Analyse von Oberflächen 24, 115-122.

Otten MT, De Jong MJC. (1991). Advances from a Field Emission Gun TEM/STEM. EMAG Conference (London), Inst. Phys. Conf. Ser. 119, 511-514.

Paris RB. (1991). The Asymptotic Behaviour of Pearcey's Integral for Complex Variables. Proc. Roy. Soc. Lond. A432, 391-426.

Pirouz P. (1981). Thin-Crystal Approximations in Structure Imaging. Acta Cryst. A37, 465-471.

Spence JCH. (1988). Experimental High Resolution Electron Microscopy, Oxford University Press, Oxford, Chapters 3-4.

Van Dyck D. (1987). Real-Space Treatment of Electron Diffraction in Crystals. Aggregate Thesis, University of Antwerp (Belgium), Chapter 4.

Van Dyck D, Danckaert J, Coene W, Selderslaghs E, Broddin D, Van Landuyt J, Amelinckx S. (1989). The Atom Column Approximation in Dynamical Electron Diffraction Calculations, in "Computer Simulation of Electron Microscope Diffraction and Images", The Minerals, Metals and Materials Society (Warrendale, USA), W Krakow and MA O'Keefe (eds), 107-134.

Van Dyck D, Op de Beeck M. (1990). New Direct Methods for Phase and Structure Retrieval in HREM, in Proc. of XIIth International Congress for Electron Microscopy, LD Peachley and DB Williams (eds), San Francisco Press, San Francisco, Vol. 1, 26-27.

Wade RH, Frank J. (1977). Electron Microscope Transfer Functions for Partially Coherent Axial Illumination and Chromatic Defocus Spread. Optik 49, 81-92.

Wilson AR, Bursill LA, Spargo AEC. (1978). Fresnel Diffraction Effects on High-Resolution (≤3Å) Images: Effect of Spherical Aberration on the Fresnel Fringe. Optik 52, 313-336.

Wilson AR, Spargo AEC. (1982). Calculation of the Scattering from Defects using Periodic Continuation Methods. Phil. Mag. A46, 435-449.

Discussion with Reviewers

M.A. O'Keefe : Would the authors care to comment on the likely effects on image delocalisation produced by non-linear contributions from the transmission-cross-coefficient (TCC) (O'Keefe, 1979; Ishizuka, 1980) to the image ?

Authors : There is a twofold answer to your comment.

(1) in terms of the TCC. The general HRTEM imaging is indeed described in terms of the TCC T by

$$\hat{I}(\mathbf{G}) = \int \hat{\phi}(\mathbf{G}+\mathbf{G}')\,\hat{\phi}^*(\mathbf{G}')\,\hat{T}(\mathbf{G}+\mathbf{G}',\mathbf{G}')\,d\mathbf{G}' \qquad \text{(C.1)}$$

This leads in real space to (Coene and Van Dyck, 1988):

$$I(\mathbf{R}) = \int \phi(\mathbf{R}')\,\phi^*(\mathbf{R}'')\,T(\mathbf{R}-\mathbf{R}',\mathbf{R}-\mathbf{R}'')\,d\mathbf{R}'\,d\mathbf{R}'' \qquad \text{(C.2)}$$

which implies that the intensity in a point $\mathbf{R}$ of the image plane is obtained by combining all pairs of ϕ and ϕ^* in points $\mathbf{R}'$ and $\mathbf{R}''$ at the specimen exit face, with a weight factor $T(\mathbf{R}-\mathbf{R}',\mathbf{R}-\mathbf{R}'')$. The latter factor for the triple-point specimen-to-image interference in real space includes both the spreading of information due to the phase transfer function (PTF) and due to non-linear imaging effects. In the case of ideal coherence, relation (1) is retrieved from (C.1), which implies that then the spreading is only due to the (pure phase) PTF. Due to the limited coherence in practice, a further spreading is introduced. In terms of the complex image amplitude $\phi_{im}(\mathbf{R})$ in (1), one obtains from (C.2) :

$$I(\mathbf{R}) = \int \phi_{im}(\mathbf{R}')\,\phi_{im}^*(\mathbf{R}'')\,M(\mathbf{R}-\mathbf{R}',\mathbf{R}-\mathbf{R}'')\,d\mathbf{R}'\,d\mathbf{R}'' \qquad \text{(C.3)}$$

M is a peaked function around the respective origins, and reduces to a product of δ - functions in the case of ideal coherence.

(2) in terms of the shape-factor analysis. For the delocalisation at the planar defect with the electron wave represented by (6), non-linear contributions arise from interference between $\phi_{<1,2>}$ and ϕ_{1-2}, and from ϕ_{1-2} with itself. For the former interference between "average" and "difference" components of the electron wave ϕ , a similar analysis applies as for linear imaging (formula (9) and following). For the auto-interference of the "difference" component, one obtains the image contributions :

$$\left|\phi_{1-2}^{im}(\mathbf{R})\right|^2 = \sum_{\mathbf{G}_1}\sum_{\mathbf{G}_2} \hat{\phi}_{1-2}(\mathbf{G}_1)\,\hat{\phi}^*_{1-2}(\mathbf{G}_2)\,\hat{t}(\mathbf{G}_1)\,\hat{t}^*(\mathbf{G}_2)\; s\big(\mathbf{R}-\nabla\chi(\mathbf{G}_1)\big)\,s\big(\mathbf{R}-\nabla\chi(\mathbf{G}_2)\big)\; \exp\left\{2\pi i\left(\mathbf{G}_1-\mathbf{G}_2\right)\mathbf{R}\right\} \qquad \text{(C.4)}$$

The diffraction components $\mathbf{G}_1$ and $\mathbf{G}_2$ are affected by their respective delocalisation $\nabla\chi(\mathbf{G}_{1,2})$, and the net delocalisation for the non-linear interference ($\mathbf{G}_1$<->$\mathbf{G}_2$) amounts to $\Delta\mathbf{R} = \nabla\chi(\mathbf{G}_1) - \nabla\chi(\mathbf{G}_2)$. For limited spatial coherence, the non-linear interferences are damped by the TCC - envelope factor

$$\exp\left\{-(\pi\alpha/\lambda)^2\left(\nabla\chi(\mathbf{G}_1)-\nabla\chi(\mathbf{G}_2)\right)^2\right\} \qquad \text{(C.5)}$$

For a FEG, the damping effect due to (C.5) is very limited.

M.A. O'Keefe : In addition, even in the linear-image regime, there is a damping term dependent on both spatial and temporal coherence. This term arises from the mixed derivative

$$\frac{\partial^2\chi(\mathbf{G})}{\partial\,\Delta f\;\partial\mathbf{G}} \qquad \text{(C.6)}$$

O'Keefe (1979) shows that this mixed derivative gives rise to a complex spatial damping envelope in which the imaginary part is dependent on both spatial and temporal coherence. Reduced from its general (linear plus non-linear) form to a linear-image form, the spatial damping expression becomes

$$C(\mathbf{G}) = \exp\left\{-\pi^2\alpha^2\left(\Delta f + \lambda^2 C_s G^2 - i\pi\lambda\Delta^2 G^2\right)^2 G^2\right\} \qquad \text{(C.7)}$$

Classic linear-image theory predicts that transfer of the image amplitude spectrum into the image intensity spectrum is a function of the sine of the linear-image transfer function, $\sin(\chi)$ (e.g. O'Keefe, 1984). However, the complex spatial coherence envelope $C(\mathbf{G})$ allows a contribution to the linear image from the cosine $\cos(\chi)$ (O'Keefe and Buseck, 1979). Do the authors wish to comment on the possible magnitude of this effect at large values of

defocus ? (Additional reference : O'Keefe MA, Buseck, PR. (1979). Transactions ACA 15, 27-46.)
Authors : The sine-contribution of the additional phase factor due to partial coherence in (C.7) is shown in figure C.1, for the half angle of beam convergence $\alpha = 1.\ 10^{-4}$ (FEG) and $\alpha = 1.\ 10^{-3}$ (LaB_6), for a few Scherzer "passband" focus settings (as in fig.2, with n as defined in (15)). The effect is most pronounced at overfocus conditions, and at higher resolution (higher G - values). However, the sine-part of the additional phase factor is - as can be seen in fig. C.1 - very small for a FEG (value of 0.02), so that the $\cos(\chi)$ contribution to the linear image formation is almost negligible. The situation is different for a LaB_6 gun : the sine-part of the phase factor can be as high as 1, indicating that the $\cos(\chi)$ contribution is in principle important for higher frequencies G, but then, this information is completely damped by the spatial coherence envelope for linear imaging E_s (relation 5b).

K. Ishizuka : Comment on the validity of using a FEG for conventional HRTEM. In order to get an intuitive image, FEGs have only disadvantages. Thermionic guns usually give enough spatial and temporal coherence, and suppress large delocalisation. The large delocalisation means a wide point spread function, and affect HRTEM images both from crystal specimens as well as from defects.
Authors : Due to the high coherence of the FEG, the information limit is (much) better than the point resolution of the microscope, complicating the image interpretation due to the oscillating behaviour of the PTF. However, almost 20 years of image simulation supporting the interpretation of experimental HRTEM images shows that also with a thermionic source, real "intuitive" images are *not* obtainable, unless at Scherzer focus, and for thin specimens. A FEG clearly adds the effect of image delocalisation to the "usual" effect of contrast reversals, but with the benefit of a high information limit. For a LaB_6 image at larger underfocus, there are frequency regions of the specimen missing in the image, which is also not the case for a FEG. Therefore, for "intuitive" microscopy, focus can only be chosen close to Scherzer, and then the high coherence of the FEG allows to benefit from the increased contrast close to the point resolution : this is a significant improvement over LaB_6 systems. The ultra-high resolution information beyond the point resolution can be cut off by the objective aperture.

A real benefit of the high information limit of the FEG for ultra-HRTEM is obtained by on-line image processing using focal image series; then, image delocalisation and contrast reversals are automatically corrected. Image delocalisation effects can also be reduced by further decreasing the spherical aberration of the objective lens.

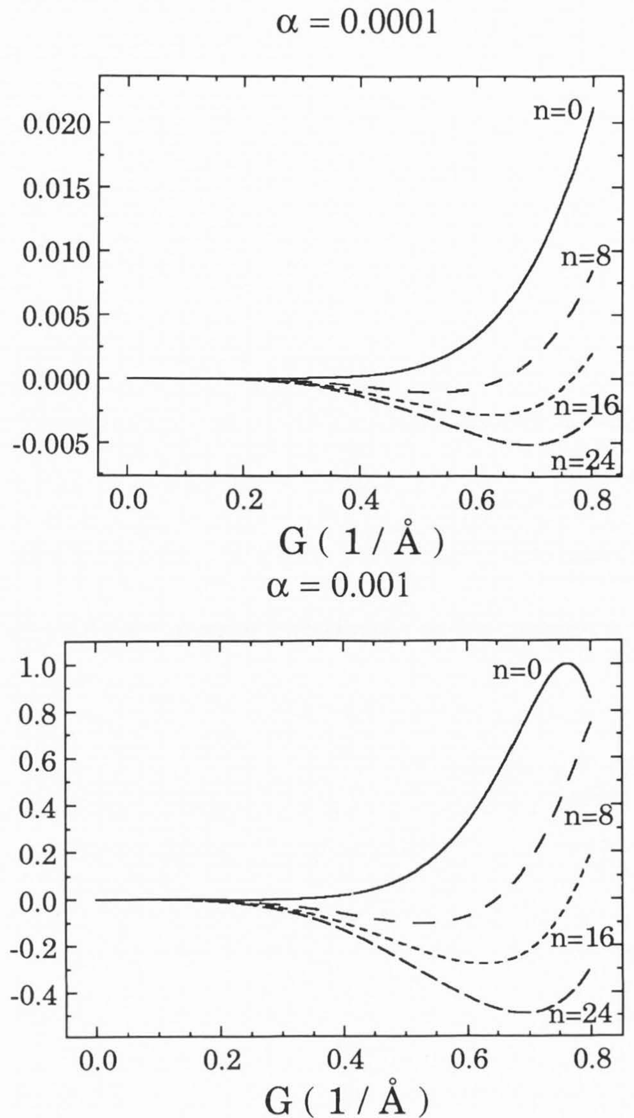

Fig. C.1 Sine-part of additional phase factor in C.7, due to combined temporal and spatial coherence. 200 keV microscope, C_s = 1.2 mm, Δ = 4.0 nm, and for a few "passband" focus settings, with n indicated.

K. Ishizuka : The delocalisation, which you describe, is just the transverse geometrical aberration, and an obvious idea. The treatment with partial coherence might be significant.
Authors : It is a soothing thought to realise that the starting point of our paper, derived on a wave-optical basis, is in agreement with simple argumentations on the basis of geometrical optics.

K. Ishizuka : Comment on the weak-phase object approximation. This approximation is essentially a projection approximation. Therefore, the inversion relation $\phi^*(-\mathbf{G}) = \phi(\mathbf{G})$ should always hold, even for the case of crystal tilt.

Authors : This is certainly true. In the discussion following (12) and (13), we have indicated in what respect crystal tilt affects the shadow images of the defect in so far that the inversion symmetry of the WPO approach is lost, due to the tilt of the Ewald sphere which makes the excitation errors for $\pm\mathbf{G}$ different. This will introduce an asymmetry in the shadow images.

K. Ishizuka : Comment on the assumption for the shape function s in Fourier space. The Fourier transform of a mathematical sign(x) function extends over into infinity. In order to the width of the phase function in Fourier space to be smaller than the reciprocal lattice of the perfect crystal region, your planar defect, i.e. the shape function in real space, should be smooth and have a width larger than the unit of the crystal.

Authors : The Fourier transform of s(x) has a 1/u behaviour (with u the frequency coordinate), which is a peaked function around the origin u=0, but with a relative large extent towards high frequencies. However, even for Ultra-HRTEM down to the Å-level, the very high-frequency behaviour is not relevant, i.e. the shape function does not need to be "mathematically sharp". The validity of this appoach is evidenced by experimental electron diffraction, where the effect of a planar defect is seen as a streak of intensity close to the Bragg spots of the perfect crystal lattice.

L.D. Marks : The authors use a weak object approximation right at the beginning of their analysis in equation (1). They argue that this is valid because the beam divergence and partial coherence are small. This is not quite correct. The weak object approximation is valid if one of two conditions is met :
(1) the scattering by the specimen is weak
(2) the incoherent convergence/chromatic aberration terms are small relative to the coherent terms.
For reference, I do not think that it is possible with current instruments to have coherent chromatic aberrations.

Authors : We are dealing here with a question of semantics. We consider the approximation in equation (1) more on the level of the electron optics, than on the level of the electron-specimen interaction. In what we call the quasi-coherent (QC) approach, the transmission-cross-coefficient is replaced by its factorised approximation (factorised with respect to both its arguments). It is not a weak object approximation, since the QC scheme also accounts for non-linear imaging effects, but its description has a (very) poor validity around the information limit of the microscope, and in the case that non-linear imaging effects are getting more important (as they do for "stronger" objects). It is beyond doubt that for a general and fully quantitative approach, the complete TCC method has to be used. However, in this paper we preferred to use the QC approach for a few reasons :
(1) it yields a simple and intuitively clear insight in the delocalisation phenomena, e.g. the shadow-image description for crystal defects
(2) delocalisation in a linear imaging model is also a matter of concern in
- phase retrieval in electron holography
- focus variation image reconstruction with the paraboloid method (Van Dyck and Op de Beeck, 1990) where chiefly the linear imaging contributions are selected for the reconstruction.