Phase Contrast in Scanning Transmission Electron Microscopy

By H. Rose

The Enrico Fermi Institute, The University of Chicago

Received 17 July 1973, in final form 23 July 1973

Abstract

A new phase contrast technique is proposed for the scanning transmission electron microscope (STEM) which significantly lowers the radiation damage in the case of very high resolution. The number of incident electrons per object element (dose) which is required to keep the quantum noise of the phase contrast image below a certain level is at most four times higher for the STEM than for the conventional electron microscope (CEM). The phase contrast images recorded by the STEM show fewer artifacts than those produced by a CEM using parallel illumination. Due to the conical illumination in the STEM only atoms lying in a thin layer of the specimen will show up with significant contrast. Therefore, the STEM is especially appropriate for the detection of single atoms in relatively thick molecules. To get the same images in a CEM, hollow beam illumination must be applied which requires a dose equivalent to that of the STEM.

Inhalt


The scanning transmission electron microscope as used today operates predominantly in the dark field mode using an annular detector that collects all scattered electrons falling outside the illumination cone [1]. For resolutions \( d \geq 2 \, \text{Å} \) the dark field detector yields a high collection efficiency [2] which results in a short scanning time per image element. The minimum dose is determined by the permissible noise level. The smaller the dose, the lower the radiation damage [3, 4].

Up to now phase contrast did not play a major role in scanning transmission microscopy because it was assumed that only electrons scattered elastically into the direction parallel to the optical axis could be used [4, 5, 6].

The phase contrast image formed by these electrons has more noise than the image recorded by the dark field detector because the latter detects a larger fraction of all scattered electrons. As the limit of resolution \( d \) is made smaller, more of the scattered electrons remain within the illumination cone, and fewer electrons fall onto the dark field detector. Therefore, the amount of information obtainable with the dark-field detector decreases as the resolution improves. The information carried by the scattered electrons falling inside the illumination cone seems to be lost [7].

On the other hand, the collection efficiency of the conventional electron microscope (CEM) increases with decreasing resolution limit. Especially the phase contrast image in the CEM should require a significantly smaller dose than that in the STEM, because in the former nearly all elastically scattered electrons falling inside the objective aperture can be used.

As a result of these facts, the STEM would appear to cause a significantly higher radiation damage than the CEM in the case of very high resolution (\( d < 1 \, \text{Å} \)) [8]. Therefore, the STEM would seem not to be the proper instrument for detecting single atoms in biological molecules. Until now, no exact values have been available for the ratio dose STEM/dose CEM for phase contrast, assuming an equal signal to noise ratio in the images of both microscopes. The corresponding ratio for the dark field case has been calculated previously [3].

In the following we will show that in the STEM, phase contrast images can be recorded with illumination times (dose) of the same order of magnitude.

---

---

Fig. 1. Ray diagram and the arrangement of detectors and apertures in the STEM and CEM operating at equivalent conditions.
as in the CEM. For this purpose an annular detector of appropriate geometry and an additional circular detector will be located behind the object inside the illumination cone. The arrangement is shown in Fig. 1 and corresponds to a CEM applying hollow beam bright field illumination. This mode of illumination in the CEM has certain advantages compared to the conventional parallel illumination, especially for the imaging of single atoms in thick molecules.

1. Phase Contrast and Noise

The phase contrast image results from an interference of the unscattered electron wave with the wave of the elastically scattered electrons. In a STEM or in a CEM using hollow beam illumination the incoming spherical wave can be described by plane waves convoluted over the entire illumination cone. Inelastically scattered electrons do not contribute to the phase contrast and can be filtered out by a spectrometer. In the first Born approximation the scattered wave is shifted by 90° with respect to the unscattered wave. On account of the spherical aberration and the defocus, an additional phase shift

\[ \gamma(\theta) - \gamma(\Theta) = k \sum_{v=0}^{\infty} \frac{C_{2v+1}}{2v+2} \theta^{2v+2} \]  

between an unscattered and a scattered electron occurs which creates regions of constructive and destructive interference. The angle between the optic axis and the two outgoing electrons is \( \Theta \); \( \theta \) determines the angle of incidence of the scattered electron. The index \( 2v+1 = n \) indicates the Seidel order. \( \lambda = 2\pi/k \) is the wavelength of the electrons.

When the additional phase shift is 0 the total phase shift is \( \pi/2 \), and hence there is no phase contrast. Then the scattered electrons cannot be distinguished from the unscattered ones. The phase contrast will show up strongly when the phase difference between the waves of a scattered and an unscattered electron can be shifted towards \( m\pi, m = 0, 1, \ldots \). A constant phase shift of 0 or \( \pi \) over the entire illumination cone will never be possible due to the conservation of the number of electrons. But alternating phase shifts which are 0 or \( \pi \) in certain hollow cones will be allowed. In this case there will be hollow cones of higher and lower electron density. Therefore, the optimum phase distribution would be a step function, (phase mask) so that sinya has a form similar to the dashed curve in Fig. 2. Of course the total number of electrons stays the same, and a detector covering the entire illumination cone beneath the object would be unaffected by this phase modulation. To extract the information out of the cone, two phase contrast detectors are placed beneath the object inside the illumination cone as shown in Fig. 1. In the general case the first phase contrast detector has the form of a zone plate, each of its rings covering a hollow cone of either constructive or destructive interference. The second phase contrast detector may consist of a single disc which lies in the shadow of the first detector.

![Fig. 2. The components \( \sin \gamma \) and \( \cos \gamma \) of the contrast transfer function for an uncorrected STEM in the case of optimum defocus. The arrows indicate the angular extension of the first annular phase contrast detector. The dashed curve characterizes the ideal shape of \( \sin \gamma \).](image)

The desired phase distribution can be approached to a reasonable degree by choosing a proper defocus \( \Delta f \approx -C_1 = \Delta z \). It can be improved further in a corrected microscope where the coefficient \( C_2 \) of the third order spherical aberration is an additional free parameter.

In the case of a thin object the phase contrast is simply the sum of the contributions of the single atoms. When the accelerating voltage \( V \) is larger than the mean atomic potential, the first Born approximation is valid. Then the wave function \( \psi_D \) can be obtained in an arbitrary detector plane \( z_D \) which is located at a distance \( l = z_D - z_0 \) behind the object plane \( z_0 \) by applying well known calculation procedures \[10,5\]:

\[ \psi_D = \frac{1}{\theta_0^3} \int \frac{1}{\pi} \int \left[ \sum_{v=0}^{\infty} C_{2v+1} \frac{e^{i\theta}}{2v+2} \delta [\theta - \Theta] \left\{ \frac{1}{\pi} \sum_{T} e^{i\gamma_1(\theta)} e^{-ik(\theta - \Theta)} \frac{Z_1}{|\theta - \Theta|^2} d\theta \right\} \right] \]

where \( \alpha = 1/137 \) is Sommerfeld's constant. The vector \( \Theta \) and \( \theta \) are perpendicular to the optic axis; \( \gamma_1 = \gamma + (z_1 - z_0) k \theta^2/2 \).

The radius vector \( R = M \) is lying in the image plane and \( M \) times larger than the corresponding vector \( \hat{r} \) in the object plane. The center of the ith atom on the screen is indicated by \( \hat{R}_i \). The intensity of the electron beam integrated over the entire object plane

\[ I_0 = \int \int \psi_0^* \psi_0^* d^2 \hat{r} = N/vt \]

is proportional to the number \( N \) of incident electrons and inversely proportional to the electron velocity \( v = \beta c \) and to the scanning time \( t \). The two-dimensional delta function \( \delta \) describing the unscattered wave makes certain that only the region inside the illumination cone contributes to the phase contrast.
The intensity

\[ I(R) = \frac{I_0}{\pi \theta_0^2} \sum_{\mu=1}^{2n} \sum_{\nu=1}^{2n-1} \left\{ 1 + \frac{4\pi}{\beta} \sum_{\xi} \left[ J_0(k_{\xi} \theta - \tilde{\theta}) \right] \frac{\tilde{R} - R_1}{R} \right\} Z_1 - \tilde{F}_1 \left[ \sin \left[ \gamma(\theta) - \gamma(\tilde{\theta}) \right] \right] \theta d\theta d\phi \]  

(4)

recorded by the first phase contrast detector is the sum of the signals detected by its m rings. The atomic form factor \( F_1 \) is a function of \( \theta - \tilde{\theta} \), \( \phi \) is the angle subtending the two vectors. If we add the signals \( I_1 \) and \( I_2 \) the resulting signal would correspond to a detector that covers the entire illumination cone and gives the uniform background intensity \( I_0 \).

Hence the signal recorded \( I(R) \) by the second phase contrast detector is

\[ I_0(R) = I_0 - I_1(R). \]  

(5)

When no specimen is present, or if the phase shift of the lens is \( \gamma = 0 \), all electrons are homogeneously distributed inside the illumination cone. In this case the numbers of detected electrons

\[ N_1 = N \Omega_1/\pi \theta_0^2, \quad N_2 = N - N_1, \quad \Omega_1 = \pi \sum_{\mu=1}^{2n} (\theta^2_{2\mu} - \theta^2_{2\mu-1}) \]  

(6)

are proportional to the solid angles \( \Omega_1 \) and \( \Omega_2 = \pi \theta_0^2 - \Omega_1 \) covered by the two phase contrast detectors. The signals of these detectors are available simultaneously and can be combined conveniently.

We can enhance the contrast

\[ g(R) = \frac{I(\infty) - I(R)}{I(\infty)} \]  

(7)

by subtracting the two signals. In this case the intensity distribution

\[ I = I_0(R) - I_1(R) = I_0 - 2I_1(R) \]  

(8)

in the image of a single atom located at \( R_1 = 0, z_1 = 0 \) is

\[ I/I_0 = 1 - 2\Omega_1/\pi \theta_0^2 - 2z_1ZG(R)/\beta, \]

\[ G(R) = \frac{4}{\pi \theta_0^2 Z} \sum_{\mu=1}^{2n} \sum_{\nu=1}^{2n-1} \left\{ J_0(k_{\xi} \theta - \tilde{\theta}) \frac{Z - \tilde{F}}{\theta - \tilde{\theta}} \right\} Z_1 - \tilde{F}_1 \left[ \sin \left[ \gamma(\theta) - \gamma(\tilde{\theta}) \right] \right] \theta d\theta d\phi. \]  

(9)

In the limit \( R \rightarrow \infty \) the function \( G \) vanishes and we have

\[ I(\infty)/I_0 = 1 - 2\Omega_1/\pi \theta_0^2 \approx (N_1 - N_2)/N. \]  

(10)

*This is valid only within the approximation used here where the number of electrons scattered outside of the illumination cone is small compared to the number of incident electrons.

Phase Contrast in Scanning Transmission Electron Microscopy

The function \( G \) is twice as large as that of a single detector due to the fact that the images recorded separately by each of the two detectors have phase contrast opposite in sign.

Although the signals of the two detectors are subtracted, their noise fluctuations \( \Delta N_1 = \sqrt{N_1} \) and \( \Delta N_2 = \sqrt{N_2} \) will add quadratically according to Gauss.

The resulting noise

\[ \Delta N = \sqrt{\Delta N_1^2 + \Delta N_2^2} = \sqrt{N} \]  

(11)

is identical to the statistical fluctuation of the total number of incident electrons. When the difference of the two signals is displayed on the screen their fluctuations create a noise contrast [4]

\[ \delta = \sqrt{N}/(N_1 - N_2). \]  

(12)

The noise contrast \( \delta \) and the mean phase contrast

\[ g = \sqrt{N}/(N_1 - N_2), \]

(13)

of an atom [4] depend strongly on the ratio \( N_1/N_2 \) determined by the detector areas. Although both contrasts become infinite when \( N_1 = N_2 \), the signal to noise ratio

\[ q = \frac{\delta}{g} = \frac{4}{3} \frac{\alpha}{\beta} \frac{Z}{\sqrt{N}} G(0) \]  

(14)

which characterizes the quality of the image always remains finite. For achieving the smallest possible radiation damage we must minimize the number \( N \) of electrons which are required for a fixed signal to noise ratio \( q \).

Since the detector signals are combined electronically, one can manipulate the contrast arbitrarily. In the conventional microscope such a procedure would be connected with a significant expenditure in apparatus. Of course the signal to noise ratio \( q \) and hence the information remain in both microscopes unaffected by such a procedure.

By proper manipulation of the two signals it is easy to reverse the phase contrast in the STEM. A corresponding procedure does not exist in the CEM.

The signal to noise ratio

\[ q_1 = \frac{1}{\sqrt{x}}, \quad q_2 = \frac{1}{\sqrt{1-x}}, \quad x = (\theta^2_1 - \theta^2_2)/\theta_0^2 \]  

(15)

belonging to the images which are formed independently by each of the two detectors always remain smaller than \( q \) because in the first approximation \( q \) is proportional to \( x(1-x) \). Since the largest signal to noise ratio is always obtained with the difference signal, we will only investigate this mode of operation in the following.
2. The optimum contrast transfer function

To determine the parameters which give the minimum dose $N$ for a fixed signal to noise ratio, we have to maximize the function $G(\theta)\langle 9 \rangle$ on the right side of Eq. (14). The result will not be changed significantly if we approximate the atomic potential by a delta function. This potential characterizes a point scatterer and yields

$$\frac{Z - E}{|\theta - \bar{\theta}|^2} = \text{const.} = \gamma Z.$$  \hspace{1cm} (16)

Assuming this simplification and indicating the point scatterer by subscript $p$ we deduce from (14) that for fixed $q$ $N$ becomes smallest when the absolute value of $G_p(r = 0)$

$$G_p(r) = \frac{8\pi}{\theta_0^2} \sum_{\mu=1}^{m} \int_{\theta_0}^{\theta_2} J_\phi(k\theta r) J_\phi(k\theta r) \sin[\gamma(\theta) - \gamma(\theta)] \theta d\theta d\theta,$$  \hspace{1cm} (17)

reaches its maximum. This is the case when the phase shift $\gamma$ is a step function so that half of the incident electrons have an additional phase shift $0$ (or $\pi$) and the other half a phase shift $\pi/2$. Then the phase shift $\pi/2 + \gamma(\theta) - \gamma(\theta)$ between an unscattered electron and an arbitrary scattered electron will be 0 or $\pi/2$ for certain hollow cones and $\pi$ or $3\pi/2$ for the others. Half of the elastically scattered electrons falling into each hollow cone or into the central cone have a phase shift $\pi/2$ with respect to the unscattered electrons and do not contribute to the phase contrast. The total solid angle of the cones yielding a constructive interference is equal to that of cones with destructive interference. This ideal distribution of the phase shift can be approached in practice only to a certain amount because in an uncorrected microscope the defocus $Af \approx -C_4$ is the only free parameter available to approximate the optimum distribution. The other free parameters determine the width of the detector rings and do not affect the actual phase shift. To approach the lowest dose for a real microscope we have to maximize $G_p(\theta)$ with respect to the free parameters $\theta_0$, $\Theta_\sigma$ and $C_{2n+1}$. The subscript $\sigma$ runs from 1 to 2 m. The subscript $n$ runs from 0 to $(n-3)/2$, where $n$ is the Seidel order of that spherical aberration which cannot be chosen arbitrarily and is limiting the resolution. In an uncorrected microscope ($n = 3$) one can already obtain a reasonable approximation of the optimum phase contrast distribution although the defocus is the only free parameter. This can be seen from Fig. 2 where the components $\sin \gamma$ and $\cos \gamma$ of the contrast transfer function $\sin[\gamma(\theta) - \gamma(\theta)]$ are shown in the case of optimum defocus

$$Af \approx 1.44(C_4)\lambda^{1/2}.$$  \hspace{1cm} (18)

Then the first phase contrast detector consists of a single ring ($m = 1$) as shown in Fig. 1.

The optimum angles of this detector are

$$\Theta_1 \approx 0.87 (\lambda/C_4)^{1/4}, \quad \Theta_2 \approx 1.49 (\lambda/C_4)^{1/4}$$  \hspace{1cm} (19)

and the optimum illumination angle is

$$\theta_0 \approx 1.71 (\lambda/C_4)^{1/4}.$$  \hspace{1cm} (20)

The latter angle is 10 per cent larger than the corresponding objective aperture of a CEM operating with parallel illumination and optimum phase contrast (Scherzer focus). Also it is 21 per cent larger than the aperture angle commonly used in electron microscopy ($4\lambda/C_4)^{1/4} [10]$).

The areas of the second phase contrast detector which do not lie in the shadow of the first detector are a central disc and an outer ring connected with each other. It should be noted that the small central detector is the only component suggested in the preceding literature for achieving phase contrast in the STEM [5, 6, 8, 14].

To achieve phase contrast in practice two annular detectors and a disc may be placed beneath the usual dark field detector. The second of these additional detectors gives the signal of the first phase contrast detector in Fig. 1. By connecting the other annular detector with the disc the conjugate signal is obtained.

In a STEM corrected for third order spherical aberration the coefficient $C_4$ is an additional free parameter. It can be used to further improve the distribution of the phase shift $\gamma$. Optimum operating conditions are approached if we choose

$$C_1 \approx -Af = 1.91 (C_4 \lambda^{2})^{1/3}, \quad C_2 = -3.26 (C_4 \lambda^{2})^{1/3}.$$  \hspace{1cm} (21)

The components $\sin \gamma$ and $\cos \gamma$ of the contrast transfer function corresponding to these values are shown in Fig. 3. The optimum configuration of the first phase contrast detector consists of a small central disc and a broad ring ($m = 2$). Because the area of the disc is small compared to that of the ring we will neglect it in the following considerations. Then the first phase contrast detector has the same shape as in the case of an uncorrected microscope

\footnote{It should be noted that the omission of the central disc results in a significantly higher dose for very high resolution ($d < 0.5 \lambda$).}
(m = 1). Of course, the optimum illumination and detector angles have changed. These optimized angles are

$$\Theta_1 \approx 1.23 \left(\frac{\lambda}{C_4}\right)^{1/6}$$

and

$$\Theta_2 \approx 1.77 \left(\frac{\lambda}{C_4}\right)^{1/6}$$

and

$$\theta_\varnothing \approx 1.86 \left(\frac{\lambda}{C_4}\right)^{1/6}. \quad (22)$$

If one looks only at the component siny given in Fig. 2 and Fig. 3, one would initially expect certain space frequencies of the object to be lost. Fortunately, this is not true because the convolution over the illumination cone has to be taken. Furthermore the scattering amplitude in (9) and also the Fourier transform of an arbitrary grating are functions of $|\Theta \pm \theta|$ and not of $\Theta$ and $\theta$ alone. The integration over the illumination cone guarantees that most spatial frequencies are transmitted equally well in the STEM as in a CEM which uses parallel illumination and optimum defocus. The STEM can transfer more space frequencies than the CEM operating in this mode because its illumination angle is larger than the corresponding aperture angle of the CEM.

3. Resolution

The limit of resolution in an electron microscope is a quantity which determines the quality of the instrument and therefore it should not be connected with the scattering characteristic of a certain object. To determine the limit of resolution two point scatterers (18) separated by a distance $d$ are considered. According to Scherzer [10] the distance $d$ is equal to the limit of resolution when the condition

$$I(0) - I(d/2) = [I(0) - I(\infty)]/4 \quad (21)$$

is fulfilled. When the objects are imaged in phase contrast the intensity $I(R)$ of two point scatterers is given by Eq. (4) replacing the scattering characteristic of an atom by that of a point scatterer (16) and assuming that the index $i$ runs from 1 to 2. The function $G_p = G_{p|R}(\rho, \tau)$ which characterizes the intensity distribution was calculated as a function of normalized radius

$$\rho = \left(\frac{n+1}{kC_n}\right)^{a+1} \frac{kR}{M}$$

for different parameter of defocus

$$\tau = \frac{1}{2} \left(\frac{n+1}{kC_n}\right)^{a+1} kAf \quad (25)$$

in the cases of an uncorrected microscope ($n = 3$) and a corrected microscope ($n = 5$). In the latter the coefficient $C_3$ of the third order spherical aberration was chosen at its optimum in accordance with (21). Three curves are shown in Fig. 4 for an uncorrected microscope. The curve with $\tau = \tau_{op} = 3.66$ corresponds to the optimum defocus (18). Then the Fresnel fringes surround-

Fig. 4. Phase contrast image of a point scatterer recorded by an uncorrected STEM in the cases of a) optimum defocus $(\tau = \tau_{op} = 3.66)$ and b) strongest Fresnel fringes $(\tau = 4.5$ and $\tau = 6.5)$. The coordinate $\rho = (kC_n/6)^{1/4} kr$ is the normalized distance from the center of the scatterer.

Fig. 5. Phase contrast image of a point scatterer recorded by a spherically corrected STEM with $\tau$ chosen at its optimum in the cases of a) optimum defocus $(\tau = \tau_{op} = 3.66)$ and b) strongest Fresnel fringes $(\tau = 6.75)$). The coordinate $\rho = (kC_n/6)^{1/4} kr$ is the normalized distance from the center of the object.

29 Optik 39, Heft 4
mine the resolution limit by applying Eq. (24). When the microscope operates at optimum conditions (18) or (21) we find

\[ d \approx 0.36 \left( C_0 \lambda^2 \right)^{1/4} \]  

(26)

as the limit of resolution for an uncorrected microscope and

\[ d \approx 0.31 \left( C_0 \lambda^2 \right)^{1/6} \]  

(27)

for a corrected microscope. These expressions agree within a few per cent with the nominal resolution \( d = 0.61 \lambda / \theta_0 \) where \( \theta_0 \) is determined by (20) and (23).

As an example we consider the limit of resolution of a corrected STEM with \( C_0 \approx 15 \) cm. Such a microscope is under construction at the University of Chicago, and according to (27) it should yield \( d = 0.67 \) Å for 400 kV electrons.

Let us now compare the resolution of an uncorrected STEM with that of an uncorrected CEM where the wavelength \( \lambda \) and \( C_0 \) are the same in both cases. When the CEM is operating with parallel illumination and optimum phase contrast its resolution distance \( d \) is approximately 1.7 times larger than that for the STEM. Due to the reciprocity theorem the resolution may become identical when hollow beam illumination is used in the CEM. This improvement in resolution is equivalent to that proposed by Hoppe [11] if we consider the hollow beam illumination as a convolution of plane waves whose propagation directions are tilted with respect to the optic axis.

In contrast to the CEM, the STEM now offers the possibility of examining a specimen in both the dark field and bright field mode simultaneously. The signal to noise ratio will be of the same order of magnitude for both modes of operation when the limit of resolution stays below 1 Å. The simultaneous use must be approached with care because the optimum defocuses and illumination angles are not the same for phase contrast and dark field operation. Applying the optimum parameters (18) and (20) for phase contrast imaging to the dark field mode we obtain for a point scatterer the distribution presented in Fig. 6. Employing this curve along with Eq. (24) we find the dark field resolution to be equal to that (26) found in the phase contrast image.

Surprisingly this dark field resolution is somewhat smaller than that obtained in standard dark field imaging \((r = 2.5, \theta_0 = (4\lambda/C_0)^{1/4})\). In this case the intensity distribution does not have the strong subsidiary maximum shown in Fig. 6. This ring has no significant influence on the resolution, but it contributes to the background and diminishes the contrast. Consequently the quality of the dark field image may be slightly reduced.

Rather than compensate the spherical aberration with a corrector, the resolution limit can be lowered with a zone plate in the STEM as well as in the CEM. The zone plate would be placed in the illumination cone ahead of the object. Here the illumination cone is subdivided into hollow cones of decreasing width. As a result the phase contrast transfer function does not change sign inside the region of a certain detector when the illumination angle \( \theta \) goes from 0 to \( \theta_0 \) because regions which give opposite sign are suppressed by the zone plate. In practice the parasitic aberrations arising from misalignment and contamination may prevent an improvement in resolution. A gain in resolution by a factor of 2 to 3 can also be achieved in correcting the spherical aberration only in two sections by means of a single octopole [12] and applying an illumination aperture of hyperbolic shape. The hole in the annular dark field detector has the same shape, while the shape of the first phase contrast detector is rather complicated.

4. Imaging of Single Atoms

It was first shown by Caceco et al [13] that single heavy atoms can be visualized in a present day STEM employing dark field detectors. The resolution limit was not better than 5 Å while the heavy atoms were separated by 12 Å in the molecules examined. Consequently it was possible to resolve these single atoms. For the detection of light atoms, such as nitrogen or carbon, a resolution of 0.5 Å or better is necessary [4]. This can be achieved only with a spherically corrected microscope or with the use of extremely high voltages. To keep the radiation damage as small as possible, high voltage microscopy should not be used4. Single atoms can also be detected in the STEM by means of the phase contrast detectors. For an approximate description of the resulting image of an atom it is sufficient to assume a Wentzel potential

\[ \phi_A = 0.85 \frac{Zhc}{2\pi r} e^{-r/a}. \]  

(28)

Sometimes it is convenient to use instead of the screening radius a the diameter

\[ d_e = 1.22 \pi a \approx 2.5 \text{ Å} Z^{-1/3} \]

In literature often the opposite prediction is found, but this statement does not hold if the signal to noise ratio which determines the information is fixed.
of the atomic electron cloud [4]. The calculation of the phase contrast image of an atom can be simplified by introducing normalized variables

$$\eta = \frac{1}{ak} \left( \frac{kCn}{n+1} \right)^{n+1} \frac{d}{2a}, \quad \rho = \left( \frac{n+1}{kCn} \right)^{n+1} kr,$$

$$u = \left( \frac{kCn}{n+1} \right)^{n+1} \frac{\Theta^2}{\sqrt{\rho^2}}, \quad w = \left( \frac{kCn}{n+1} \right)^{n+1} \frac{\Theta^2}{\sqrt{\rho^2}}.$$

When the Wentzel potential (28) is used, the azimuthal integration in Eq. (7) can be obtained in a closed form, and we find

$$G(\rho, \eta) = \frac{1.7}{w_0} \int_0^{w_0} \left[ \int_{u_1}^{\infty} \sum_{n=0}^{\infty} \frac{2}{1 + \delta_{on}} \frac{J_n(\rho \sqrt{u}) J_n(\rho \sqrt{w})}{\sqrt{\eta^2 + u + w} - 4uw} \right. \left. \left[ \frac{2}{\sqrt{uw}} \sin \left[ \frac{\gamma(w) - \gamma(u)}{4} \right] \right] \right] \, du \, dw$$

assuming a single ring detector ($m = 1$). In the limit $a \to 0 (\eta \to \infty)$ the distribution $\eta^2 G(\eta, \rho)$ approaches that of a point scatterer, and only the first summand remains of the sum. Then Eq. (30) becomes identical to the expression (17) if we set in the latter formula $\lambda = 0.85/\eta^2$. The finite extension of the atomic potential causes the additional terms ($a > 0$) in the summation. These terms do not affect the central intensity ($\eta = 0$) of the image. They influence primarily the outer part of the intensity distribution. The value of the term in brackets becomes larger when the resolution limit decreases, but it will never exceed $1$. Thus, for smaller resolutions more terms contribute to the sum (30). The intensity distribution of an atom starts to differ from the distribution of a point scatterer when the limit of resolution $d$ becomes smaller than $4a(\eta < 2)$. This deviation can be seen in comparing Fig. 5 with Fig. 7, where $d$ is identical to the screening radius ($\eta = \frac{1}{2}$). The dashed curve in Fig. 7 represents the parabolic approximation which determines the spot size of an atom. The image consists of a strong central spot with a broad base; this is surrounded by a wide weak Fresnel fringe. The corresponding image of a point scatterer (Fig. 5, $\tau_0 = -5.9$) does not have the broad base for the central spot while it does have a somewhat stronger first Fresnel fringe. When the resolution limit is larger than $d$, the structure of an atom is hidden inside the Airy distribution and the atom can be considered as a point scatterer.

Although the Fresnel fringes are very weak in the STEM they never can completely vanish because the integral of the phase contrast taken over the entire image plane (screen) must be zero. In a CEM using parallel illumination

---

\*\*\* A diameter $d_0 \approx 2.2 \text{ Å} Z^{-1/4}$ would give results which agree well with those based on Hartree-Fock-Slater wave functions [2].

---

\*\*\* The thickness of the object should not exceed the electron mean free path length (Auffeildicke).
gives maximum contrast. This region is adjoined on each side by a region where the phase contrast has opposite sign. Regions of opposite phase contrast must always occur because the integral of $\varphi$ taken over the entire range of defocus vanishes. The presented curves indicate that only atoms which are located in a layer of thickness

$$D \approx 25 \frac{d^2}{\lambda}$$

appear with detectable contrast. Atoms not lying in this layer do not contribute to the phase contrast.

The contrast of the supporting film is substantially lowered when a plane of vanishing phase contrast is located inside the foil. Then the atoms lying above and beneath this plane appear with opposite phase contrast and cancel each other to a large extent. The atoms of an object will appear with contrast of the same sign as long as the entire object is lying inside of one of the three regions mentioned above. When the object covers regions of different sign the atomic images of certain layers cancel each other and only the atoms of a relatively thin layer appear with detectable contrast.

The sum over the phase contrast of numerous atoms lying far outside the optimum focal plane stays finite because the phase contrast decreases inversely proportional to the square of the defocus. This makes the STEM extremely suitable for the detection of single atoms in relatively thick specimens.

The parameter $\eta \approx d/2a$, which takes into account the finite size of the atoms, does not significantly influence the dependence of the phase contrast on defocus. The curves in Fig. 8 and Fig. 9 differ only slightly in shape when normalized. A noticeable deviation occurs only for very high resolution ($\eta = 1/3$). While the shape of the curves is relatively independent of $\eta$, their absolute values are very sensitive to this parameter. The dependence of $K(\eta, \tau_{op})$ and $\eta^2 K(\eta, \tau_{op})$ on the parameter $\eta$ is shown in Fig. 10 for a corrected microscope. Within an accuracy of 5 per cent the curves also hold for an uncorrected microscope operating at optimum defocus (18).

When the resolution limit decreases, the shape of an atom shows up and its intensity distribution becomes broader than that of a point scatterer. According to Scherzer [4] we define the radius $d_0/2$ of the atomic spot as that distance at which the amplitude of the scattered wave, approximated by a parabola, has decreased to one-fourth of its value at the center of the spot. With this definition the spot size $d_0$ was calculated from (30) for different values of $\eta$ by taking only the quadratic terms of $q$ into account and assuming

$$\tau = (kC/4)^{1/2} k\Delta f/2$$

Fig. 8. The dependence of the mean phase contrast of defocus for an uncorrected STEM. The curve $\eta = \infty$ belongs to a point scatterer, the curve $\eta = 2$ to an atom in the case $d \approx 4a$.

$$\eta K(\eta, \tau)$$

$$\tau = (kC/6)^{1/2} k\Delta f/2$$

Fig. 9. The dependence of the mean phase contrast of an atom on defocus for a spherically corrected STEM operating at different resolutions. The curve with $\eta = 1/3$ is plotted 3 times larger.

$$\frac{d_0}{d} \eta K(\eta, \tau_{op})$$

Fig. 10. The spot size $d_0/d$ and the mean phase contrast $\bar{\varphi}$ as functions of $\eta \approx d/2a$ for a spherically corrected STEM operating at optimum conditions. The function $q/q_0$ is the ratio of the signal to noise ratios of the phase contrast and dark field images. The curves are valid within 5 per cent for an uncorrected microscope.
a corrected microscope operating at optimum conditions (21). The result is shown in Fig. 10. When the resolution limit is much larger than the screening radius the spot size and the resolution limit are identical. The spot size \( d_s \) becomes larger than \( d \) for \( \eta \leq 2 \) but it always remains smaller than \( 2d \).

In the STEM additional information about the object can be obtained from the dark field image which is recorded simultaneously. Because artifacts will unlikely appear in both images, no further exposure is necessary for the determination of the image resolution. Two exposures must be taken in the CEM, but their information may already differ due to the radiation damage.

The dark field and phase contrast image of the STEM are of similar quality when both images have a signal to noise ratio of the same order of magnitude.

The signal to noise ratio in the dark field image

\[
q_D = \sqrt{n_D}
\]

is equal to the square root of the number \( n_D \) of electrons recorded by the detector. In the case of a single atom we find

\[
n_D = 4N\sigma_{el}/\pi d^2 \text{SD}.
\]

Here the scattering cross section \( \sigma_{el} \) determines the fraction of electrons scattered elastically in the object plane inside a circle of diameter \( d \text{SD} \). The spot size \( d \text{SD} \) of the dark field image of an atom referred to the object plane differs slightly from that of the phase contrast image. The ratio

\[
\frac{q}{q_D} = 0.725 \frac{d \text{SD}}{d_e} \left( \frac{\sigma_e}{\sigma_{el}} \right)^{1/2}
\]

of the signal to noise ratios of the phase contrast and dark field image is a function of the parameter \( \eta \). It depends on the resolution and the atomic number. The ratio \( \sigma_{el}/\sigma_e \) determines which fraction of the elastically scattered electrons contribute to the atomic spot of the dark field image; \( \sigma_e \) is the total elastic scattering cross section.

To determine which mode of operation yields the largest signal to noise ratio we have evaluated the formula (36) for the case that the dark field and phase contrast image are each taken at optimum conditions. The resulting curve \( q/q_D \) represented in Fig. 10 shows that the signal to noise ratio of the phase contrast image is lower than that of the dark field image as long as the resolution limit \( d \) is larger than 1.3 \( a \). For very high resolution the phase contrast image is superior with respect to the noise. When both images are taken simultaneously \( q/q_D \) increases for optimum phase contrast conditions and decreases for optimum dark field conditions. It should be noted that the signal to noise ratio alone is not sufficient to characterize the image quality. For this purpose the contrast must be known too.

5. Comparison of STEM and CEM

In Fig. 1 the detectors of a STEM and the illumination apertures of the corresponding CEM are sketched. If in the CEM the first illumination mode is used an image is produced which is identical to that of a STEM using only the signal of the first annular phase contrast detector. The image corresponding to the second illumination mode in the CEM is identical to that obtained with the second phase contrast detector in the STEM.

While the CEM can give only one of these images at a time, both are simultaneously available in the STEM. The efficiency of the STEM is not limited to these two signals alone. In addition, the signals of the inelastically scattered electrons and the signal of the dark field detector can be used. This is of great advantage because it gives additional information and offers the possibility of combining the signals conveniently. For example, the simultaneous use of the dark field and phase contrast signals offers the possibility of determining the phase and the amplitude of the scattered electron wave. The spectrometer for separating the inelastically scattered electrons is not shown in Fig. 1. It should be placed between the dark field and the phase contrast detectors.

Although the phase contrast images are the same in both microscopes at equivalent operating conditions, the doses required for a certain signal to noise ratio may differ. The signal to noise ratio of a CEM operating in the first illumination mode is given by

\[
q_1 = \frac{g_1}{\Omega \sin^2 \theta_1} \left[ \frac{N}{2\pi \beta Z_{1/2}} \int_{\theta_e}^{\theta_1} \int_{\phi_e}^{\phi_1} \sin[\gamma(\theta) - \gamma(\theta_1)] \Omega_1 d\theta \Omega_1 d\phi \right]^{-1/2}
\]

The meaning of the angles \( \theta \) and \( \Theta \) is reversed in going from the STEM to the CEM in Fig. 1. When the CEM operates in the second illumination mode the contrast is reversed in sign and we find for the signal to noise ratio

\[
q_2 = \frac{q_1}{x(x - 1)}, \quad x = (\Theta_2^2 - \Theta_1^2)/\Theta_1^2.
\]

The phase contrast at the center of the atomic image is

\[
g_1(0) = \frac{3}{2} q_1 \Omega = \frac{3}{2} q_1 N^{-1/2} = \frac{2Z}{\beta x} G(0)
\]

In the limit \( \Theta_1 = \Theta_2 \rightarrow 0 \) this equation coincides with the well known formula for parallel illumination first derived by Scherzer [10].

For the CEM the signal to noise ratio \( q_1 \) is at its maximum with parallel illumination. Then the required dose is a minimum resulting in the lowest radiation damage. The STEM reaches lowest radiation damage when the solid angles of both detectors become identical. \(^3\)

To make a statement on the relative quality of the STEM and the CEM we have to compare the doses \( N_{STEM} \) and \( N_{CEM} \) which are required for a

\(^3\) This statement is valid only when the phase shift can be chosen arbitrarily. In a real microscope the optimum solid angles may slightly differ.
fixed signal to noise ratio \( q_{STEM} = q_{CEM} \). Applying the formulas (12), (13) and (37) we find the following ratio of the required doses

\[
\frac{N_{CEM}}{N_{STEM}} = 4x^2G_{STEM}(0)/G_{CEM}(0). \tag{40}
\]

Here we have assumed that in the CEM the first illumination mode is used and the difference signal is taken in the STEM. When the CEM operates in the second illumination mode we have to replace in Eq. (40) \( x \) by \( 1 - x \).

It should be mentioned that we compare only the phase contrast images of both microscopes. The additional use of the dark field signal in the STEM will further reduce the required dose rate. When the STEM and the CEM are operating with equivalent conditions we find \( G_{CEM} = G_{STEM} \) in comparing (9) with (37) and (39). Then the ratio of the doses depends only on \( x \) and they become equal for \( x \approx 1/2 \) where the STEM gives lowest radiation damage. Now we want to compare the microscopes when both of them are operating at optimum conditions*. In this case \( G_{CEM} \) and \( G_{STEM} \) are not identical and have to be calculated from (9) and (37). When the illumination angle \( \theta_0 \) is larger than \( 1/ka \) the free parameters \( C_0, \gamma = 1, 3, \cdots n \) must be chosen for the STEM in such a way that the phase mask \( (\gamma) \) has \((n + 3)/2 \) zones of alternating phases. The corresponding first phase contrast detector consists of \([n+3]/4 \) rings, where the brackets indicate the integer value. As a result we find

\[
\frac{N_{STEM}}{N_{CEM}} = 3.2 \quad \text{for} \quad n = 3, \quad \frac{N_{STEM}}{N_{CEM}} \approx 4 \quad \text{for} \quad n \geq 5. \tag{41}
\]

Here \( n \) describes the Seidel order of the spherical aberration that limits the resolution. These given ratios are upper limits and show that even in the case of highest resolution \( (d \ll 1 \text{Å}) \) the required dose of the STEM is at the most four times higher than that of a CEM. This result does not agree with previous estimations [4, 8]. It should be noted that the conditions for optimum signal to noise used to arrive at Eq. (41) result in a resolution distance \( d_{CEM} \approx 1.7 \, d_{STEM} \).

6. Stability Criterion

The phase contrast is extremely sensitive to a fluctuation \( \delta f \) of the objective focal length caused by the instabilities of the lens fields. In addition to this, the unavoidable energy spread \( \Delta E \) of the cathode and the fluctuation \( \Delta V \) of the accelerating voltage \( V \) randomly change the electron wavelength, further reducing the phase contrast. To prevent a deterioration of the phase contrast image each of these fluctuations must be kept small. The fluctuations do not significantly deteriorate the image so long as the phase contrast for the center of the atom does not fall below 80 per cent. This is true when the following stability criterion is fulfilled:

\[
C_0^2 \left( 1 + \varepsilon V \right)^2 \left( \frac{\Delta E \sin \theta}{\epsilon} + \frac{(\Delta V)^2}{V^2} \right) + (\delta f)^2 \leq 2 \, \frac{d^4 i_s}{\lambda^2}. \tag{42}
\]

* The parameters for the CEM operating at optimum conditions (parallel illumination) are given in Refs. [10, 4].

To get an estimate of the order of magnitude of the permissible fluctuations, we apply the stability criterion to a spherically corrected microscope \( (C_3 \approx 0, C_0 \approx 1 \text{mm}) \) using a field emission gun with a source of energy spread \( \Delta E \approx \pm 0.15 \text{eV} \). The accelerating voltage is assumed to be 100 kV. When this voltage and the currents in the objective lens can be stabilized to better than 1/2 p.p.m., the second and third term on the left side of Eq. (42) can be neglected. In this case the phase contrast will not be significantly weakened by the velocity spread of the electrons as long as the resolution limit is larger than 0.8 Å. If one wants to go to still smaller resolutions, the additional correction of the chromatic aberration is unavoidable.

In principle one can avoid the chromatic correction in going to higher accelerating voltages. Unfortunately, this is not presently possible because of technological difficulties in keeping the relative fluctuations of the high voltage beneath 1 p.p.m.

7. Conclusions

In the preceding chapters we have shown that single atoms can be detected by means of phase contrast in the STEM as well as in the CEM. Certainly, the STEM requires a four times higher dose than the CEM operating with parallel illumination. This disadvantage of the STEM is counterbalanced by an image containing few artifacts and smaller resolved object details. An image corresponding to that of a STEM can be recorded by the CEM using hollow beam illumination, but then the dose is the same as in the STEM. In the latter, additional use can be made of the electrons collected by the dark field detector and of the inelastically scattered electrons. These can be separated by a spectrometer and used to form an image containing additional information. Also the electrons which have excited K electrons may be used for the determination of the mean atomic number [15]. In a CEM no inelastically scattered electrons can be used, even when the objective lens is chromatically corrected in first degree. The electrons which have undergone only a small energy loss will be imaged perfectly, but they carry no high resolution information and contribute predominantly to the background [9]. The electrons which are inelastically scattered in the inner shells of the atoms carry high resolution information. Unfortunately, they have suffered such a high energy loss that it exceeds the range of the chromatic correction. Therefore all inelastically scattered electrons should be filtered out in a CEM.

The described phase contrast technique may become important in both the STEM and the CEM for the detection of single atoms in thick molecules. To achieve the resolution which is necessary for the visualization of light atoms the spherical and the chromatic aberration of the objective lenses must be corrected in both microscopes.
Acknowledgements

I wish to thank Ted Roth, who did most of the numerical computations, and Professor A. V. Crewe and Professor E. Zeiiller for the opportunity of performing this work during my sabbatical year at the Enrico Fermi Institute of the University of Chicago. The valuable discussions which I had with them, the members of their groups and especially with Dr. R. W. Moses, Jr. had helped to clear many questions connected with this work.

References


Experimental Studies of Some Dampings of Electron Microscope Phase Contrast Transfer Function

By A. Beorochia and P. Bonhomme

Laboratoire de Microscopie et Spectroscopie Electroniques, Faculté des Sciences, Reims, France

E.R. A No 376 (C.N. RS)

Received 26 July 1973

Abstract

We prove experimentally by diffractogram tests of electron micrograph, theoretical studies about phase contrast transfer function envelope owing to chromatic aberrations and partial coherence of the illuminating beam.

Inhalt

Experimentelle Untersuchung einiger Maxima der Phasenkontrastübertragungsfunktionen eines Elektronenmikroskopes. Durch lichtoptische Fourieranalyse von elektronenmikroskopischen Aufnahmen bestätigen wir experimentell die gerechneten Lagen einiger Maxima der Phasenkontrastübertragungsfunktionen, die sich aus Farbfehlerkoefizienten und Beleuchtungseffekten ergeben.

I. Introduction

It has been shown [9, 11, 12, 13] that spherical aberration and defocusing of objective can be corrected inside or outside of the microscope. If we assume these corrections done, the size of the smallest phase object details transferred is limited by chromatic aberration or electron source size effect. Transfer function envelop have already been theoretically studied in some typical case [6, 7]. Our aim is to show here some experimental checkings of these theoretical results.

II. Experimental Method

We apply known defects to high voltage or objective current, or we work with different apertures illuminating beam. We use an Hitachi HU II C electron microscope equipped with a superconducting objective [4] (Cn = 1.8 mm, Cz = 1.9 mm). Direct inspection of amorphous carbon micrograph diffractograms or photometric recordings of their diametral intensity are used for determination of the smallest details transferred in each case.