J. Phys. D: Appl. Phys., Vol. 6, 1973. Printed in Great Britain.

LETTER TO THE EDITOR

A method for the solution of the phase problem in electron microscopy

D L Misell
Physics Department, Queen Elizabeth College, Campden Hill Road, London, W8 7AH

MS received 13 November 1972

Abstract. A method is given for the evaluation, in transmission electron microscopy, of the amplitude and phase from the intensity distribution of an electron micrograph. The method requires a minimum of two micrographs taken under different defocus conditions. The iterative scheme requires only the relative defocus between micrographs, and the procedure is valid both in bright-field and dark-field microscopy for any specified coherence of the electron source. Assumptions on the scattering properties of the specimen, such as the weak-phase–weak-amplitude object, are not required. For a complete determination of the amplitude–phase distribution for electron transmission through the specimen, the electron micrograph must be corrected for the effect of lens aberrations and defocusing to give the electron wavefunction immediately after transmission; only in the case of a weak-phase object can this wavefunction be directly related to the projected potential distribution in the object.

Inelastic electron scattering is explicitly omitted from the analysis presented.

In the transmission electron microscope (conventional or scanning type) the elastic electron wave transmitted by the specimen, $\psi_0(\boldsymbol{r})$, is modified by the lens aberration function $G_1(\boldsymbol{r})$ for the objective lens to give the image wavefunction

$$\psi_1(\boldsymbol{r})=\int \psi_0(\boldsymbol{r}')G_1(\boldsymbol{r}-\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \tag{1}$$

in the isoplanatic approximation (unit magnification, $M=1$) and for coherent illumination of the specimen. The essential problem in electron microscopy is to invert equation (1) to give information on the specimen structure contained in ψ_0. Only in the case of the phase-grating approximation can ψ_0 be simply related to the projected potential distribution in the object, $\phi_0(\boldsymbol{r})$; in this approximation, $\psi_0(\boldsymbol{r})=1+\mathrm{i}\eta(\boldsymbol{r})$ and

$$\eta(\boldsymbol{r})=\sigma\phi_0(\boldsymbol{r})=-\frac{2\pi me\lambda_0}{h^2}\int_0^t V(\boldsymbol{r},z)\,\mathrm{d}z \tag{2}$$

for a specimen of thickness t.

The image resolution function $G_1(\boldsymbol{r})$ includes terms for spherical aberration (C_S), defocusing (Δf_1) and axial astigmatism (C_A), where the Fourier transform of G_1 is given by

$$F(G_1(\boldsymbol{r}))=T_1(\boldsymbol{\nu})=B(\boldsymbol{\nu})\exp\left(-\mathrm{i}K_0W_1(\boldsymbol{\nu})\right) \tag{3}$$

with

$$W_1(\boldsymbol{\nu})=\frac{C_S\nu^4\lambda_0^4}{4}+\frac{\Delta f_1\nu^2\lambda_0^2}{2}-\frac{C_A}{2}(\nu_x^2-\nu_y^2)\lambda_0^2 \tag{4}$$

for electrons of wavelength λ_0 angle of scattering θ by $\nu\lambda_0=$ give ψ_0 (which may be diffra known. Experimentally we m determined; we cannot then in not known, and it is necessary Several procedures have been involves the use of the weak-p

$$\psi_0(\boldsymbol{r})=1+\mathrm{i}\eta(\boldsymbol{r})-$$

where the phase η and absorpti restriction for a biological sec determined either by recording (Erickson and Klug 1971, Fran (Hoppe 1970, 1971, Hoppe *et* bright-field microscopy, allow scattered wave, and explicitly transform relationship betwee (Gerchberg and Saxton 1971, and image intensity has been e method appears to work in approximations of the first pro and Saxton method is the rec pattern from a selected area ar for amorphous (noncrystalline stain distribution may be non-u determine $\psi_1(\boldsymbol{r})$ from $|\psi_1(\boldsymbol{r})|$ a requiring at least two micrograp method, however, does not requ valid for dark-field microscopy are numerous (Hoppe 1970, preliminary evaluation of the p the other methods mentioned a near the optimum defocus fo structured background to the the effect of even this inelastic

We consider a second micro

$$\psi_2(\boldsymbol{r})=\int \psi_0(\boldsymbol{r}')G$$

with $F(G_2(\boldsymbol{r}))=B(\boldsymbol{\nu})\exp(-\mathrm{i}K_0$ $\Delta f=\Delta f_2-\Delta f_1$, which can be c

$$F(G_2(\boldsymbol{r}))=B(\boldsymbol{\nu})\,e$$

Thus ψ_1 and ψ_2 are related by

$$\psi_2(\boldsymbol{r})=\int \psi_1(\boldsymbol{r}')G$$

where

$$G(\boldsymbol{r})=F^{-1}(\exp($$

for electrons of wavelength λ_0. $\nu=(\nu_x, \nu_y)$ denotes the spatial frequency, related to the angle of scattering θ by $\nu\lambda_0=\theta$; $K_0=2\pi/\lambda_0$. In principle equation (1) can be solved to give ψ_0 (which may be diffraction-limited due to the aperture function $B(\nu)$) if ψ_1 is known. Experimentally we measure the image intensity $j_1(r)=|\psi_1(r)|^2$ and only $|\psi_1|$ is determined; we cannot then invert equation (1) to give ψ_0, since the phase terms of ψ_1 are not known, and it is necessary to go back and devise a procedure to evaluate ψ_1 from $|\psi_1|$. Several procedures have been suggested for evaluating ψ_0 from $|\psi_1|$. The first method involves the use of the weak-phase–weak-amplitude object:

$$\psi_0(r)=1+i\eta(r)-\epsilon(r) \tag{5}$$

where the phase η and absorption ϵ terms are assumed to be much less than unity, a severe restriction for a biological section about 20–50 nm in thickness. Both η and ϵ can be determined either by recording two micrographs for different defocus values Δf_1 and Δf_2 (Erickson and Klug 1971, Frank 1972), or by using complementary semicircular apertures (Hoppe 1970, 1971, Hoppe *et al* 1970, Lenz 1971). Both these techniques apply only in bright-field microscopy, allowing the unscattered wave to interfere with the elastically scattered wave, and explicitly depend on η, $\epsilon \ll 1$. A second scheme uses the Fourier transform relationship between the diffracted wavefunction and the image wavefunction (Gerchberg and Saxton 1971, 1972); an iterative scheme between the diffracted intensity and image intensity has been evaluated to determine the phase terms in both planes; this method appears to work in dark-field and bright-field microscopy and without the approximations of the first procedure given above. The main limitation of the Gerchberg and Saxton method is the requirement of the correspondence between the diffraction pattern from a selected area and image area analysed; such a correspondence is relevant for amorphous (noncrystalline) specimens, particularly biological specimens where the stain distribution may be non-uniform. The procedure suggested in this letter in order to determine $\psi_1(r)$ from $|\psi_1(r)|$ avoids some of the limitations of the first scheme, although requiring at least two micrographs recorded at different defocus values Δf_1 and Δf_2; the method, however, does not require the weak-phase–weak-amplitude approximation and is valid for dark-field microscopy. The problems involved in correlating two micrographs are numerous (Hoppe 1970, 1971, Frank 1972), but the procedure is presented as a preliminary evaluation of the phase problem in the electron microscope. In common with the other methods mentioned above, we explicitly omit inelastic electron scattering which, near the optimum defocus for the elastic image, contributes only as a relatively unstructured background to the total (elastic+inelastic image); the problem of evaluating the effect of even this inelastic background on the present scheme is not considered here.

We consider a second micrograph taken at defocus Δf_2, giving the image wavefunction

$$\psi_2(r)=\int \psi_0(r')G_2(r-r')\,dr' \tag{6}$$

with $F(G_2(r))=B(\nu)\exp(-iK_0W_2(\nu))$; or in terms of $W_1(\nu)$ and the defocus difference $\Delta f=\Delta f_2-\Delta f_1$, which can be determined with precision,

$$F(G_2(r))=B(\nu)\exp(-iK_0W_1(\nu))\exp(-iK_0\Delta f\nu^2\lambda_0^2/2). \tag{7}$$

Thus ψ_1 and ψ_2 are related by the convolution integral

$$\psi_2(r)=\int \psi_1(r')G(r-r')\,dr' \tag{8}$$

where

$$G(r)=F^{-1}(\exp(-iK_0\Delta f\nu^2\lambda_0^2/2)).$$

Equation (8) forms the basis of an iterative scheme to determine ψ_1 and ψ_2 from their respective moduli $|\psi_1|$ and $|\psi_2|$.

The amplitude $|\psi_1(\boldsymbol{r})|$ is calculated from the image intensity $j_1(\boldsymbol{r})=|\psi_1(\boldsymbol{r})|^2$, and an initial approximation to $\psi_1(\boldsymbol{r})=|\psi_1(\boldsymbol{r})|\exp(i\eta_1(\boldsymbol{r}))$ is made using an initial phase distribution $\eta_1(\boldsymbol{r})$ (eg random phases between $\pm\pi$). We then evaluate the convolution (equation (8)) to give an approximation to $\psi_2(\boldsymbol{r})$, namely $\psi_2'(\boldsymbol{r})=|\psi_2'(\boldsymbol{r})|\exp(i\eta_2(\boldsymbol{r}))$. These $|\psi_2'|$ are compared with the actual images amplitudes $|\psi_2|$ calculated using the image intensity $j_2(\boldsymbol{r})=|\psi_2(\boldsymbol{r})|^2$; as these two results $|\psi_2|$ and $|\psi_2'|$ will in general not be similar, we retain the phase terms η_2 and form a new wavefunction from the actual amplitudes $|\psi_2|$, namely $\psi_2(\boldsymbol{r})=|\psi_2(\boldsymbol{r})|\exp(i\eta_2(\boldsymbol{r}))$. This modified ψ_2 is then used to calculate ψ_1 from the convolution $\psi_2(\boldsymbol{r})*G'(\boldsymbol{r})$, where $G'(\boldsymbol{r})=F^{-1}(\exp(iK_0\Delta f v^2\lambda_0^2/2))$, to give a new estimate for ψ_1, namely ψ_1'. Again the calculated phases η_1 are retained, and if $|\psi_1'|$ differs appreciably from $|\psi_1|$, we form the product $\psi_1(\boldsymbol{r})=|\psi_1(\boldsymbol{r})|\exp(i\eta_1(\boldsymbol{r}))$ and recalculate ψ_2 from ψ_1. The procedure given above is summarized in figure 1, where the iteration is completed when a

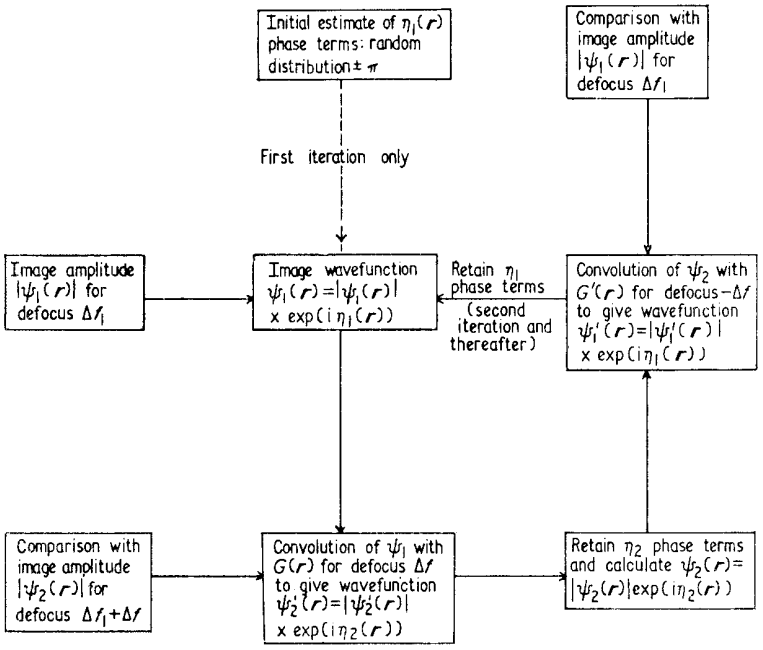


Figure 1. Flow diagram for the evaluation of the phase distribution in two electron micrographs taken at defocus values Δf_1 and $\Delta f_1+\Delta f$; the respective image intensities are $j_1(\boldsymbol{r})=|\psi_1(\boldsymbol{r})|^2$ and $j_2(\boldsymbol{r})=|\psi_2(\boldsymbol{r})|^2$.

comparison of $|\psi_1'|$ (or $|\psi_2'|$) with $|\psi_1|$ (or $|\psi_2|$), the actual amplitudes, gives only a specified small difference. The inequality of $|\psi_1'|$ and $|\psi_1|$ outside a test program arises from the error in measuring the true image intensities j_1 and j_2, due to noise and nonlinear optical densities, and the background contribution from inelastic electron scattering.

The procedure given above does not initially require the values for the spherical aberration coefficient C_S, the absolute defocus values Δf_1 and Δf_2, or the axial astigmatism coefficient C_A; only the relative defocus value between micrographs is required to determine the image wavefunction. In order to evaluate the object wavefunction ψ_0 from

either ψ_1 or ψ_2 using equatio
tion $B(\nu)$; the standard metl
fraction pattern of the electr
zeros of $\sin(K_0W_1(\nu))$ may b
lens (Thon 1971, Frank 197

Recently holographic tec
transmission microscope ima
Halioua 1972a):

$$j_1(\boldsymbol{r})=\int |\psi_0(\boldsymbol{r}'$$

and the phase information co
been evaluated in the general
1972b), but only for a low-c
bright-field microscopy. The
source coherence, if this can
front (a given $\boldsymbol{K}_0$ and K_0) eq
The image intensities j_1 and
over the coherence function

References

Erickson H P and Klug A 1971 *P*
Frank J 1972 *Biophys. J.* **12** 484–5
Gerchberg R W and Saxton W O
—— 1972 *Optik* **35** 237–46
Hoppe W 1970 *Acta Crystallogr.*
—— 1971 *Phil. Trans. R. Soc.* **261**
Hoppe W, Langer R and Thon F
Lenz F 1971 *Proc. 25th Anniversa*
Stroke G W and Halioua M 1972a
—— 1972b *Optik* **35** 489–508
Thon F 1971 *Electron Microscopy*
625

either ψ_1 or ψ_2 using equation (1), we require C_S, Δf_1 (or Δf_2), C_A and the aperture function $B(\nu)$; the standard method uses an optical diffractometer, and from the optical diffraction pattern of the electron micrograph (namely the thin carbon substrate film) the zeros of sin $(K_0 W_1(\nu))$ may be determined, giving the required parameters of the objective lens (Thon 1971, Frank 1972).

Recently holographic techniques have been applied to the correction of a scanning transmission microscope image of a virus assuming incoherent illumination (Stroke and Halioua 1972a):

$$j_1(\boldsymbol{r}) = \int |\psi_0(\boldsymbol{r}')|^2 \, |G_1(\boldsymbol{r}-\boldsymbol{r}')|^2 \, \mathrm{d}\boldsymbol{r}' \tag{9}$$

and the phase information contained in ψ_0 cannot be determined, although the theory has been evaluated in the general case of partially coherent illumination (Stroke and Halioua 1972b), but only for a low-contrast object (weak-phase–weak-amplitude object) and in bright-field microscopy. The present scheme can be applied in the case of partial electron source coherence, if this can be specified; for a given point on the incident electron wavefront (a given $\boldsymbol{K}_0$ and K_0) equation (8) still describes the relationship between ψ_1 and ψ_2. The image intensities j_1 and j_2 are then calculated from an integration of $|\psi_1|^2$ and $|\psi_2|^2$ over the coherence function for the electron source.

References

Erickson H P and Klug A 1971 *Phil. Trans. R. Soc.* **261B** 105–18

Frank J 1972 *Biophys. J.* **12** 484–511

Gerchberg R W and Saxton W O 1971 *Optik* **34** 275–84

—— 1972 *Optik* **35** 237–46

Hoppe W 1970 *Acta Crystallogr.* **A26** 414–26

—— 1971 *Phil. Trans. R. Soc.* **261B** 71–94

Hoppe W, Langer R and Thon F 1970 *Optik* **30** 538–45

Lenz F 1971 *Proc. 25th Anniversary Meeting of EMAG* (London: Institute of Physics) pp 224–9

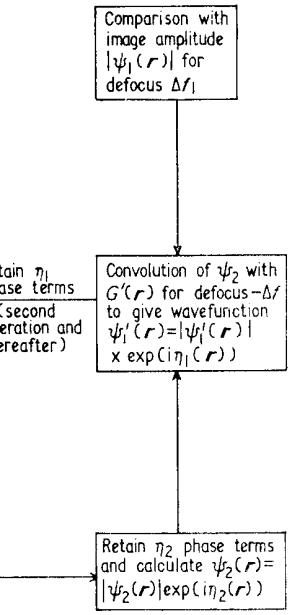
Stroke G W and Halioua M 1972a *Optik* **35** 50–65

—— 1972b *Optik* **35** 489–508

Thon F 1971 *Electron Microscopy in Material Science* ed U Valdré (New York: Academic Press) pp 570–625

etermine ψ_1 and ψ_2 from their

ntensity $j_1(\boldsymbol{r}) = |\psi_1(\boldsymbol{r})|^2$, and an
using an initial phase distribu-
uate the convolution (equation
$_2'(\boldsymbol{r})|$ exp $(i\eta_2(\boldsymbol{r}))$. These $|\psi_2'|$
ulated using the image intensity
eneral not be similar, we retain
actual amplitudes $|\psi_2|$, namely
calculate ψ_1 from the convolu-
to give a new estimate for ψ_1,
and if $|\psi_1'|$ differs appreciably
nd recalculate ψ_2 from ψ_1. The
e iteration is completed when a



e phase distribution in two electron
$+\Delta f$; the respective image intensities

actual amplitudes, gives only a
$_1|$ outside a test program arises
nd j_2, due to noise and nonlinear
inelastic electron scattering.
re the values for the spherical
nd Δf_2, or the axial astigmatism
nicrographs is required to deter-
e object wavefunction ψ_0 from