Image formation in the electron microscope. I. The application of transfer theory to a consideration of elastic electron scattering

1971 J. Phys. A: Gen. Phys. 4 782

(http://iopscience.iop.org/0022-3689/4/6/005)

J. Phys. A: Gen. Phys., 1971, Vol. 4. Printed in Great Britain

Image formation in the electron microscope I. The application of transfer theory to a consideration of elastic electron scattering

D. L. MISELL
Department of Physics, Queen Elizabeth College, London W8, UK
MS. received 26*th April* 1971

Abstract. The application of transfer theory to image formation in the electron microscope is discussed. The equivalence of the transfer theory and the diffraction integral formulation of image formation is established for a coherent monochromatic incident electron beam, and for elastic electron scattering. The transfer theory is extended to include partially coherent and nonmonochromatic radiation. The use of the results of the transfer theory leads to the diffraction integral formulation, which includes the effects of chromatic aberration (thermal energy effect) and the coherence of the incident electron beam. The modification to a criterion for optimum contrast in the final image when chromatic aberration is taken into account is discussed.

1. Introduction

The purpose of this paper is to consider the application of transfer theory to a consideration of image formation in the electron microscope. In particular, the inclusion of the angular (spatial) and energy (chromatic) characteristics of the incident electron beam into the wave theory of image formation will be considered. In the case of monochromatic parallel (coherent) radiation the transfer theory leads to the diffraction integral formulation for the image intensity. The diffraction integral formulation (Scherzer 1949, Uyeda 1955, Haine 1957) has been used extensively to assess the possibility of visualizing a single atom or an array of single atoms in a conventional transmission electron microscope (Heidenreich and Hamming 1965, Eisenhandler and Siegel 1966, Reimer 1969, Niehrs 1969, 1970, Hall and Hines 1970). No detailed analysis seems to have been made on more realistic specimens including the effects of the angular and energy distributions of the incident electron beam; in particular, the chromatic aberration effect due to the thermal energy distribution of the electrons emitted from the electron gun has not been considered. Inelastic electron scattering and its role in image formation is usually considered as an incoherent background in the final image; a notable exception is the work of Haine (1957). It also seems usual to work in single scattering conditions (or in the first Born approximation) and the amplitude of the unscattered component is taken to be unity. Clearly these limitations are imposed by the rigidity of the simple form taken by the diffraction integral formulation. In the theory of image formation in the scanning transmission electron microscope (Zeitler and Thomson 1970a, 1970b) the exclusion of the thermal energy distribution of the incident electron beam and of the inelastic electron scattering is totally justified by the design features of this type of microscope (see Crewe 1970 for a review).

If a more general approach to image formation, such as the transfer theory, is used then the specific diffraction integral formulation may be derived under all conditions of illumination and the effect of chromatic aberration may be included. It is noted that the present work is specifically orientated towards a calculation of the electron microscope image and is not directed towards the far more significant and difficult problem of inferring the nature of the specimen from the image; only in the case of a linear transfer theory, as applied to a weak phase object, has success been achieved in this direction (Hanszen 1969, 1971).

In this paper, elastic electron scattering only is considered and inelastic electron scattering, which is incoherent with respect to the elastic and unscattered components of the transmitted electron beam, will be treated in a second paper. § 2 deals with the application of the transfer theory to image formation with an axially coherent and monochromatic electron beam. In §§ 3 and 4, the extension of theory to a consideration respectively of the angular divergence and energy distribution of the incident electron beam is given. § 5 considers the combined effect of the angular and energy distributions of the incident electron beam on the final image. In relation to elastic electron scattering, a criterion for optimum contrast is discussed; in addition to spherical aberration and defocusing effects, chromatic aberration is included (§ 4.2).

2. Monochromatic and spatially coherent radiation

The subject matter of this section represents the application of optical transfer theory (eg Hopkins 1953, 1955, Born and Wolf 1959) to image formation in the electron microscope; in particular, the integral diffraction equation is derived in its usual form suitable for numerical evaluation.

The incident electron is represented by a plane wave $\exp(\mathrm{i}\,\boldsymbol{K}_0 \cdot \boldsymbol{r})$ of unit amplitude. $\boldsymbol{K}_0$ is the wavevector describing the direction of the incident electron beam; $|\boldsymbol{K}_0| = 2\pi/\lambda_0$ represents the incident energy of the electron beam ($E_0 = \hbar^2 K_0{}^2/2m$). The incident electron is scattered by the specimen and the scattered wave ψ_0 immediately after the object is dependent on the conditions of specimen illumination and the electron scattering properties of the specimen. In the first case the radiation is monochromatic ($|\boldsymbol{K}_0|$ = constant) and coherent in the z direction ($\boldsymbol{K}_0$ = constant along the optic axis). If the object is such that all electrons incident on the object are transmitted (ie negligible back-scattering), then the current density $j_0(\boldsymbol{r}_0)$ immediately behind the object is

$$j_0(\boldsymbol{r}_0) = |\psi_0(\boldsymbol{r}_0)|^2 \tag{1}$$

where $\boldsymbol{r}_0$ represents coordinates (x_0, y_0) in the object plane. Inelastic electron scattering is explicitly omitted in equation (1) because of the omission in ψ_0 of any dependence on a scattered wavevector $\boldsymbol{K}$. It will be assumed that inelastic electron scattering gives rise to an incoherent background (not necessarily a constant) in the final image.

The electron microscope transmits information on $\psi_0(\boldsymbol{r}_0)$ to an image plane; the image wave $\psi_\mathrm{i}(\boldsymbol{r}_\mathrm{i})$ is defined by image coordinates $\boldsymbol{r}_\mathrm{i} = (x_\mathrm{i}, y_\mathrm{i})$. If it is assumed that noise, for example, mechanical vibrations of the electron microscope column, granularity of the recording photographic emulsion, may be neglected, then there is a unique relationship between ψ_0 and ψ_i. The linear relationship between ψ_0 and ψ_i can be represented by (Lenz 1965)

$$\psi_\mathrm{i}(\boldsymbol{r}_\mathrm{i}) = \int \psi_0(\boldsymbol{r}_0) G(\boldsymbol{r}_0, \boldsymbol{r}_\mathrm{i})\, \mathrm{d}\boldsymbol{r}_0. \tag{2}$$

$G(\boldsymbol{r}_0{}', \boldsymbol{r}_\mathrm{i})$ is the image wavefunction corresponding to a point source of electrons at $\boldsymbol{r}_0{}'$ in the object. $G(\boldsymbol{r}_0, \boldsymbol{r}_\mathrm{i})$ is determined only by the electron–optical system, namely the lens aberrations, apertures, defocusing (with reference to the objective lens system), and the specimen illumination conditions. In the isoplanatic approximation

(Born and Wolf 1959, Lenz 1965)

$$G(\boldsymbol{r}_0, \boldsymbol{r}_i) = G\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \tag{3}$$

where M is the electron–optical magnification.

The isoplanatic approximation is valid for aberrations of the objective lens that depend only on the direction of the electron trajectory (angle of scattering); this includes the third order aberrations: spherical aberration, chromatic aberration, axial astigmatism and defocusing. In the isoplanatic approximation, equation (2) becomes a two dimensional convolution integral

$$\psi_i(\boldsymbol{r}_i) = \int \psi_0(\boldsymbol{r}_0) G\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) d\boldsymbol{r}_0. \tag{4}$$

By the application of a Fourier transformation to equation (4), $\psi_i(\boldsymbol{r}_i)$ can be written as

$$\psi_i(\boldsymbol{r}_i) = \int S_0(\boldsymbol{\nu}) T(\boldsymbol{\nu}) \exp\left(-2\pi i \boldsymbol{\nu} \cdot \frac{\boldsymbol{r}_i}{M}\right) d\boldsymbol{\nu} \tag{5}$$

where the two dimensional Fourier transforms of ψ_0, ψ_i and G are defined respectively by

$$S_0(\boldsymbol{\nu}) = \int \psi_0(\boldsymbol{r}_0) \exp(2\pi i \boldsymbol{\nu} \cdot \boldsymbol{r}_0)\, d\boldsymbol{r}_0$$

$$S_i(\boldsymbol{\nu}) = \int \psi_i(\boldsymbol{r}_i) \exp\left(2\pi i \boldsymbol{\nu} \cdot \frac{\boldsymbol{r}_i}{M}\right) d\boldsymbol{r}_i \tag{6}$$

$$T(\boldsymbol{\nu}) = \int G(\boldsymbol{r}) \exp(2\pi i \boldsymbol{\nu} \cdot \boldsymbol{r})\, d\boldsymbol{r}$$

where $\boldsymbol{\nu}$ is the vector (ν_x, ν_y) and $\boldsymbol{\nu} \cdot \boldsymbol{r} = \nu_x x + \nu_y y$.

$T(\boldsymbol{\nu})$ is the amplitude transfer function (Born and Wolf 1959); $T(\boldsymbol{\nu})$ can be derived for an electron–optical system from the properties of the objective lens (Lenz 1965)

$$T(\boldsymbol{\nu}) = \frac{1}{M} \exp\{-iK_0 W(\boldsymbol{\nu})\} B(\boldsymbol{\nu}) \tag{7}$$

where W is termed the wave aberration function producing a phase shift $K_0 W$ in the scattered electron wave (the negative sign in the exponential implies a phase advance). $B(\boldsymbol{\nu})$ is an aperture function; $B(\boldsymbol{\nu}) = 1$ for the transparent parts of the aperture and $B(\boldsymbol{\nu}) = 0$ for the opaque parts.

The Kirchhoff diffraction integral, frequently used in the calculation of electron microscope images (eg Scherzer 1949, Uyeda 1955, Haine 1957), and as applied to the scattered wave in the back focal plane of the objective lens, is the equivalent of the transform from S_i to ψ_i (equation (5)). The Kirchhoff integral is over the transparent part of the objective aperture and takes account of the phase shifts introduced by aberrations and defocusing (Scherzer 1949). Each point in the back focal plane of the objective lens corresponds to a spatial frequency $\boldsymbol{\nu}$ and $S_0(\boldsymbol{\nu})$ represents the diffracted (scattered) wave in the back focal plane (Lenz 1965). It is usual to transform equation (5) to a coordinate system in real space, either $\boldsymbol{r}_B(= (x_B, y_B)$ coordinates in the back focal plane) or polar coordinates $\boldsymbol{\theta}(= (\theta, \phi)$ polar and azimuthal angles of scattering), where

$$\boldsymbol{r}_{\mathrm{B}} = f\lambda_0\boldsymbol{\nu} \quad \text{or} \quad \boldsymbol{\theta} = \lambda_0\boldsymbol{\nu}. \tag{8}$$

f is the focal length of the objective lens. $\boldsymbol{\nu}$ is equivalent to the reciprocal lattice vector $\boldsymbol{g}$ for a crystalline specimen. In either of these coordinate systems, W of equation (7) is for a lens subject to spherical aberration, defocusing and axial astigmatism respectively (eg Lenz 1965, Hanszen 1971)

$$W(\boldsymbol{r}_{\mathrm{B}}) = \frac{C_{\mathrm{s}}}{4}\frac{(x_{\mathrm{B}}^2+y_{\mathrm{B}}^2)^2}{f^4} + \frac{\Delta f}{2}\frac{(x_{\mathrm{B}}^2+y_{\mathrm{B}}^2)}{f^2} - \frac{C_{\mathrm{A}}}{2}\frac{(x_{\mathrm{B}}^2-y_{\mathrm{B}}^2)}{f^2} \tag{9}$$

or

$$\chi(\boldsymbol{\theta}) = \frac{C_{\mathrm{s}}\theta^4}{4} + \frac{\Delta f\theta^2}{2} - \frac{C_{\mathrm{A}}\theta^2}{2}\cos 2\phi. \tag{10}$$

C_{s} is the third order spherical aberration coefficient, C_{A} is the coefficient of astigmatism, Δf is the change in focal length of the objective lens ($\Delta f < 0$ corresponds to underfocusing, $\Delta f > 0$ to overfocusing). In equation (7) the negative sign implies a phase advance with respect to the electron wave travelling along the optic axis of the objective lens. The gaussian image plane is defined for an aberration free lens and $\Delta f = 0$.

The transformation of equation (5) to (θ, ϕ) coordinates is effected by

$$\begin{aligned} \theta\cos\phi &= \lambda_0\nu_x \\ \theta\sin\phi &= \lambda_0\nu_y \end{aligned} \tag{11}$$

for the small-angle approximation ($\sin\theta \simeq \theta$) which is valid in the conventional transmission electron microscope ($\theta \leqslant 0{\cdot}1$ rad).

Equation (5) becomes

$$\psi_{\mathrm{i}}(x_{\mathrm{i}}, y_{\mathrm{i}}) = \frac{K_0}{2\pi}\int_0^{2\pi}\int_0^{\pi}\Psi(\boldsymbol{\theta})H(\boldsymbol{\theta})\exp\left(-\frac{\mathrm{i}K_0}{M}(x_{\mathrm{i}}\theta\cos\theta+y_{\mathrm{i}}\theta\sin\phi)\right)\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \tag{12}$$

where

$$H(\boldsymbol{\theta}) = \frac{1}{M}\exp\{-\mathrm{i}K_0\chi(\boldsymbol{\theta})\}D(\boldsymbol{\theta}) \tag{13}$$

is the $\boldsymbol{\theta}$ equivalent of equation (7).

In bright field electron microscopy with the normal circular aperture

$$\begin{aligned} D(\boldsymbol{\theta}) &= 1 \qquad 0 \leqslant \theta \leqslant \alpha \qquad 0 \leqslant \phi \leqslant 2\pi \\ &= 0 \qquad \alpha < \theta \qquad 0 \leqslant \phi \leqslant 2\pi. \end{aligned} \tag{13a}$$

α is the semiangle subtended by the objective aperture at the specimen. Equation (12) becomes

$$\psi_{\mathrm{i}}(x_{\mathrm{i}}, y_{\mathrm{i}}) = \frac{K_0}{2\pi M}\int_0^{2\pi}\int_0^{\alpha}\Psi(\boldsymbol{\theta})\exp\{-\mathrm{i}K_0\chi(\boldsymbol{\theta})\}\exp\left(-\frac{\mathrm{i}K_0}{M}(x_{\mathrm{i}}\theta\cos\phi+y_{\mathrm{i}}\theta\sin\phi)\right) \times\theta\,\mathrm{d}\theta\,\mathrm{d}\phi. \tag{14}$$

$\Psi(\boldsymbol{\theta})$ represents the diffraction pattern in the back focal plane and includes the unscattered component (represented by $\delta(\boldsymbol{\theta})$). If $\Psi(\boldsymbol{\theta})$ is separated into the unscattered component $\beta\delta(\boldsymbol{\theta})$ (β^2 represents the fraction of the incident electron beam that is unscattered) and an elastic wave $\Psi_E(\boldsymbol{\theta})$, equation (14) becomes

$$\psi_i(x_i, y_i) = \frac{\beta}{M} + \frac{K_0}{2\pi M}\int_0^{2\pi}\int_0^{\alpha} \Psi_E(\boldsymbol{\theta}) \exp\{-iK_0\chi(\boldsymbol{\theta})\} \times \exp\left(-\frac{iK_0}{M}(x_i\theta\cos\phi + y_i\theta\sin\phi)\right)\theta\, d\theta\, d\phi. \quad (15)$$

The current density $j_i(\boldsymbol{r}_i)$ in the image is

$$j_i(\boldsymbol{r}_i) = \psi_i(x_i, y_i)\psi_i^*(x_i, y_i) = |\psi_i(x_i, y_i)|^2. \quad (16)$$

Using Parseval's theorem derived from the properties of the Fourier transform (Sneddon 1951) it can be shown that

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(|\psi_i(x_i, y_i)|^2 - \frac{\beta^2}{M^2}\right) dx_i\, dy_i = \int_0^{2\pi}\int_0^{\alpha} |\Psi_E(\theta, \phi)|^2 \theta\, d\theta\, d\phi. \quad (17)$$

Equation (17) expresses the physical situation that if the background intensity β^2/M^2 is subtracted from the image intensity j_i, then the number of electrons in $(j_i - \beta^2/M^2)$ is identical to the number scattered elastically within the objective aperture, $h_E(\alpha)$ (Reimer 1969, Crick and Misell 1971).

Equation (15) has been used to calculate the optimum conditions of defocusing for the resolution of a single atom or an array of atoms (eg Heidenreich and Hamming 1965, Hall and Hines 1970) with β set equal to unity. In the case of an object with a centre of symmetry, such as obtains for a single atom, equation (15) becomes (Scherzer 1949), neglecting axial astigmatism

$$\psi_i(r_i) = \frac{\beta}{M} + \frac{K_0}{M}\int_0^{\alpha} \Psi_E(\theta) \exp\{-iK_0\chi(\theta)\} J_0\left(\frac{K_0 r_i \theta}{M}\right)\theta\, d\theta. \quad (18)$$

Hence transfer theory may be used to obtain a specific relationship between ψ_0 and ψ_i in a form frequently applied in the literature. It is noted that the specific equations such as equations (15) and (18) assume a coherent incident electron wave and neglect the chromatic aberration due to the thermal energy spread of the incident beam. The chromatic aberration term for the elastically scattered electrons is at least the same order of magnitude as the defocusing and spherical aberration terms in $\chi(\theta)$ (see § 4.2 and Crick and Misell 1971). Thus the resolution criterion based on a consideration of spherical aberration, defocusing and the diffraction limit (specified in equation (18) by the Bessel function of order zero J_0 and the value for α) is not strictly valid.

In the case of scattering by a single atom, $\Psi_E(\theta) = \exp(i\pi/2)f^B(\theta)$ in the first Born approximation; the $\pi/2$ factor expresses the phase difference (delay) between the elastically scattered wave and the unscattered wave $\beta\delta(\theta)$ (eg Haine 1957). The contrast $C(r_i)$ is, from equation (18) neglecting second order terms

$$C(r_i) = \frac{j_i(r_i) - \text{background intensity}}{\text{background intensity}} \simeq 2K_0\beta^{-1}\int_0^{\alpha} f^B(\theta)\sin\{K_0\chi(\theta)\} J_0\left(\frac{K_0 r_i \theta}{M}\right)\theta\, d\theta. \quad (19)$$

More correctly the complex electron scattering factor $f(\theta)\exp\{i\eta(\theta)\}$ should be used in place of $f^B(\theta)$ (Zeitler 1966), giving $f(\theta)\sin\{K_0\chi(\theta) - \eta(\theta)\}$ in the integral.

On axis with $r_i = 0$, maximum contrast is obtained for (eg Zeitler and Thomson 1970a)

$$\text{(i) overfocus,} \quad \Delta f > 0 \qquad \Delta f = \frac{\lambda_0}{\alpha^2} - \frac{C_s\alpha^2}{2} \qquad \alpha^2 < \left(\frac{2\lambda_0}{C_s}\right)^{1/2}$$

$$\text{(ii) underfocus,} \quad \Delta f < 0 \qquad \Delta f = -\frac{C_s\alpha^2}{2} \tag{20}$$

where the overfocus condition (i) is established on the basis of the oscillatory nature of the $\sin\{K_0\chi(\theta)\}$ term in equation (19); the general condition is that the $\sin\{K_0\chi(\theta)\}$ term should be nonzero for $0 < \theta < \alpha$ and none of the spatial frequencies transmitted by the objective aperture are removed from the image (eg Hanszen 1971). The underfocus condition is based on a partial cancellation of the spherical aberration term in $\chi(\theta)$ by the defocus term. The maximum contrast is obtained for $\alpha_{opt} \simeq (4\lambda_0/C_s)^{1/4}$ and $\Delta f_{opt} = -(\lambda_0 C_s)^{1/2}$ (Zeitler and Thomson 1970a). A slightly modified criterion is obtained when chromatic aberration is included in the wave theory (see § 4.2).

3. The effect of spatial incoherence on the image for monochromatic radiation

If the wavevector $\boldsymbol{K}_0$ of the incident electron varies, then the radiation is spatially incoherent. The distribution $F(\boldsymbol{K}_0)$ of $\boldsymbol{K}_0$ represents the angular distribution of the electrons emitted from the electron gun. $\boldsymbol{K}_0$ may be considered as a wavevector with two components K_x and K_y. The scattered wave ψ_0 is now dependent on $\boldsymbol{K}_0$ and $\boldsymbol{r}_0$. $\psi_0(\boldsymbol{K}_0, \boldsymbol{r}_0)$ may be written as (Lenz 1965)

$$\psi_0(\boldsymbol{K}_0, \boldsymbol{r}_0) = \psi_0(\boldsymbol{r}_0)\exp(\mathrm{i}\boldsymbol{K}_0 \cdot \boldsymbol{r}_0) \tag{21}$$

for elastic electron scattering, provided that the angle of incidence, represented by $\boldsymbol{K}_0$, when multiplied by the specimen thickness t is less than the smallest detail to be resolved in the specimen (lateral effect, see also Crick and Misell 1971).

Initially a general distribution $F(\boldsymbol{K}_0)$ is considered (§ 3.1) and it is shown that for specific forms for $F(\boldsymbol{K}_0)$ (§ 3.2), the equations for j_i in the coherent and incoherent cases can be derived.

3.1. *General distribution*

The image wave of equation (4) may be written, for a given $\boldsymbol{K}_0$, as

$$\psi_i(\boldsymbol{K}_0, \boldsymbol{r}_i) = \int \psi_0(\boldsymbol{r}_0) G\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \exp(\mathrm{i}\boldsymbol{K}_0 \cdot \boldsymbol{r}_0)\,\mathrm{d}\boldsymbol{r}_0 \tag{22}$$

since G is not dependent on $\boldsymbol{K}_0$ for monochromatic radiation (Lenz 1965). The $|\psi_i(\boldsymbol{K}_0, \boldsymbol{r}_i)|^2$ corresponding to different $\boldsymbol{K}_0$ are superimposed incoherently, that is

$$j_i(\boldsymbol{r}_i) = \int |\psi_i(\boldsymbol{K}_0, \boldsymbol{r}_i)|^2 F(\boldsymbol{K}_0)\,\mathrm{d}\boldsymbol{K}_0 \tag{23}$$

or

$$j_i(\boldsymbol{r}_i) = \int\int\int \psi_0(\boldsymbol{r}_0)\psi_0^*(\boldsymbol{r}_0') G\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) G^*\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0'\right) F(\boldsymbol{K}_0) \times \exp\{\mathrm{i}\boldsymbol{K}_0 \cdot (\boldsymbol{r}_0 - \boldsymbol{r}_0')\}\,\mathrm{d}\boldsymbol{r}_0\,\mathrm{d}\boldsymbol{r}_0'\,\mathrm{d}\boldsymbol{K}_0 \tag{24}$$

where $F(\boldsymbol{K}_0)$ is normalized such that

$$\int F(\boldsymbol{K}_0)\,\mathrm{d}\boldsymbol{K}_0 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(K_x, K_y)\,\mathrm{d}K_x\,\mathrm{d}K_y = 1.$$

The asterisk indicates the complex conjugate of a function. In order to transform equation (24) into a more practical form, consider a Fourier transformation of equation (22)

$$\int \psi_i(\boldsymbol{K}_0, \boldsymbol{r}_i)\exp\left(2\pi i\boldsymbol{\nu}\,.\,\frac{\boldsymbol{r}_i}{M}\right)\mathrm{d}\boldsymbol{r}_i = \int\int \psi_0(\boldsymbol{r}_0)\exp(i\boldsymbol{K}_0\,.\,\boldsymbol{r}_0 + 2\pi i\boldsymbol{\nu}\,.\,\boldsymbol{r}_0)G\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \times \exp\left\{2\pi i\boldsymbol{\nu}\,.\,\left(\frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right)\right\}\mathrm{d}\boldsymbol{r}_0\,\mathrm{d}\boldsymbol{r}_i. \tag{25}$$

From the definitions of the Fourier transforms of ψ_0 and G (equation (6)) equation (25) becomes

$$\psi_i(\boldsymbol{K}_0, \boldsymbol{r}_i) = \int S_0\left(\frac{\boldsymbol{K}_0}{2\pi} + \boldsymbol{\nu}\right) T(\boldsymbol{\nu})\exp\left(-2\pi i\boldsymbol{\nu}\,.\,\frac{\boldsymbol{r}_i}{M}\right)\mathrm{d}\boldsymbol{\nu} \tag{26}$$

and equation (24) for the image intensity is

$$j_i(\boldsymbol{r}_i) = \int\int\int S_0\left(\frac{\boldsymbol{K}_0}{2\pi} + \boldsymbol{\nu}\right)S_0^*\left(\frac{\boldsymbol{K}_0}{2\pi} + \boldsymbol{\nu}'\right)T(\boldsymbol{\nu})T^*(\boldsymbol{\nu}') \times \exp\left(-2\pi i(\boldsymbol{\nu} - \boldsymbol{\nu}')\,.\,\frac{\boldsymbol{r}_i}{M}\right)F(\boldsymbol{K}_0)\,\mathrm{d}\boldsymbol{\nu}\,\mathrm{d}\boldsymbol{\nu}'\,\mathrm{d}\boldsymbol{K}_0. \tag{27}$$

It is seen from equation (27) that the integration over $\boldsymbol{K}_0$ represents a convolution of $S_0 S_0^*$ with F. In the conventional transmission electron microscope $F(\boldsymbol{K}_0)$ represents the angular distribution of the incident beam on the specimen after focusing by the double condenser lens system. If the condenser aperture (semiangle α_c subtended at the electron source) is small, then $F(\boldsymbol{K}_0) = \delta(\boldsymbol{K}_0)$ and the integration over $\boldsymbol{K}_0$ in equation (27) gives

$$j_i(\boldsymbol{r}_i) = \left|\int S_0(\boldsymbol{\nu})T(\boldsymbol{\nu})\exp\left(-2\pi i\boldsymbol{\nu}\,.\,\frac{\boldsymbol{r}_i}{M}\right)\mathrm{d}\boldsymbol{\nu}\right|^2 \tag{28}$$

which corresponds to the spatial coherent equation (5).

Equation (27) is now transformed to polar coordinates with

$$\boldsymbol{\theta} = \lambda_0\boldsymbol{\nu} \quad \text{and} \quad \boldsymbol{\theta}_c = \lambda_0\frac{\boldsymbol{K}_0}{2\pi} = \frac{\boldsymbol{K}_0}{K_0} \tag{29}$$

where $\boldsymbol{\theta}_c$ defines the angular coordinate of the illumination.

Hence equation (27) becomes

$$j_i(\boldsymbol{r}_i) = \left(\frac{K_0}{2\pi}\right)^2\int\int\int \Psi(\boldsymbol{\theta}_c + \boldsymbol{\theta})\Psi^*(\boldsymbol{\theta}_c + \boldsymbol{\theta}')H(\boldsymbol{\theta})H^*(\boldsymbol{\theta}') \times \exp\left(-\frac{iK_0}{M}(\boldsymbol{\theta} - \boldsymbol{\theta}')\,.\,\boldsymbol{r}_i\right)I_0(\boldsymbol{\theta}_c)\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{\theta}'\,\mathrm{d}\boldsymbol{\theta}_c \tag{30}$$

where $\mathrm{d}\boldsymbol{\theta} = \theta\,\mathrm{d}\theta\,\mathrm{d}\phi$ and $\boldsymbol{\theta}\,.\,\boldsymbol{r}_i = \theta\cos\phi\,x_i + \theta\sin\phi\,y_i$.

The angular distribution of the incident electron beam $I_0(\boldsymbol{\theta}_c)$ is a measurable function (see, for example, Burge *et al.* 1970, where I_0 is measured under small-angle diffraction conditions). I_0 may be represented by a gaussian distribution of halfwidth less than 10^{-4} rad for the normal double condenser lens system (AEI EM6). It is noted that formally the integrations over $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ should be evaluated before the integration over $\boldsymbol{\theta}_c$ is performed. The convolution of $\Psi\Psi^*$ with I_0 represents a modification to the angular distribution of the scattered electron beam by the incident angular spread (see eg Crick and Misell 1971). In the case where $I_0(\boldsymbol{\theta}_c) = \delta(\boldsymbol{\theta}_c)$, equation (30) reduces to the coherent case (equation (14)).

3.2. *Specific distribution*

In order to illustrate the analysis of § 3.1, a specific form for $F(\boldsymbol{K}_0)$ is chosen, namely $I_0(\boldsymbol{\theta}_c) = 1/\pi\alpha_c^2$ for a condenser lens aperture illuminated by an electron beam of uniform intensity. Consider the integration over $\boldsymbol{K}_0$ in equation (24)

$$\Phi(\boldsymbol{r}_0-\boldsymbol{r}_0') = \int F(\boldsymbol{K}_0)\exp\{\mathrm{i}\boldsymbol{K}_0 \cdot (\boldsymbol{r}_0-\boldsymbol{r}_0')\}\,\mathrm{d}\boldsymbol{K}_0 \tag{31}$$

which represents the two dimensional Fourier transform of $F(\boldsymbol{K}_0)$. Equation (31) is transformed to $\boldsymbol{\theta}_c$ coordinates using equation (29) to give for the cylindrically symmetric $I_0(\theta_c, \phi)$

$$\Phi(\boldsymbol{r}_0-\boldsymbol{r}_0') = \frac{2\pi}{\pi\alpha_c^2}\int_0^{\alpha_c} \mathrm{J}_0(K_0\theta_c|\boldsymbol{r}_0-\boldsymbol{r}_0'|)\theta_c\,\mathrm{d}\theta_c$$

that is

$$\Phi(\boldsymbol{r}_0-\boldsymbol{r}_0') = \frac{2\mathrm{J}_1(K_0\alpha_c|\boldsymbol{r}_0-\boldsymbol{r}_0'|)}{K_0\alpha_c|\boldsymbol{r}_0-\boldsymbol{r}_0'|}. \tag{32}$$

The function Φ defines the spatial coherence of the illumination (see Hopkins 1951, for the light optical case). If $\alpha_c \to 0$, Φ has the value 1 and equation (24) for the image intensity becomes

$$j_i(\boldsymbol{r}_i) = \left|\int \psi_0(\boldsymbol{r}_0)G\left(\frac{\boldsymbol{r}_i}{M}-\boldsymbol{r}_0\right)\mathrm{d}\boldsymbol{r}_0\right|^2 \tag{33}$$

which is the coherent case (equation (4)).

If α_c is large, then Φ becomes $\delta(\boldsymbol{r}_0-\boldsymbol{r}_0')$ and the integral over $\boldsymbol{r}_0'$ in equation (24) may be evaluated to give

$$j_i(\boldsymbol{r}_i) = \int \left|\psi_0(\boldsymbol{r}_0)G\left(\frac{\boldsymbol{r}_i}{M}-\boldsymbol{r}_0\right)\right|^2 \mathrm{d}\boldsymbol{r}_0 = \int |\psi_0(\boldsymbol{r}_0)|^2 \left|G\left(\frac{\boldsymbol{r}_i}{M}-\boldsymbol{r}_0\right)\right|^2 \mathrm{d}\boldsymbol{r}_0 \tag{34}$$

which is the incoherent case. In general for a uniform illumination of the condenser aperture, Φ is a measure of the spatial coherence (eg Zeitler and Thomson 1970b).

4. The effect of chromatic incoherence on the image for spatially coherent radiation

The incident electron beam has an energy distribution represented by a variation in $|\boldsymbol{K}_0|$, where the energy of the incident electron beam is $\hbar^2K_0^2/2m$. In this section, spatial coherence of the incident electron beam will be assumed, that is, $\boldsymbol{K}_0$ is a constant; in § 5 spatial incoherence will be included into the theory. Initially it will be assumed that electrons with different K_0 are incoherent with respect to phase. In § 4.3, the possibility of coherence over a small interval in K_0, ΔK_0, will be considered.

The energy distribution of the incident electron beam is represented by $F(K_0)$ or $N(E)$, where E represents the energy spread about a most probable value E_0. Both $F(K_0)$ and $N(E)$ are normalized such that

$$\int_0^\infty F(K_0)\,\mathrm{d}K_0 = 1$$

and

$$\int_{-\infty}^{+\infty} N(E)\,\mathrm{d}E = 1. \tag{35}$$

Under normal operating conditions of the electron gun $N(E)$ may be represented by an exponential form (Andersen and Mol 1970), that is

$$N(E) = \frac{\beta^{\mu+1}}{\Gamma(\mu+1)}\left(\frac{\mu}{\beta}-E\right)^{\mu}\exp\left\{-\beta\left(\frac{\mu}{\beta}-E\right)\right\} \tag{36}$$

with maximum at $E = 0$ and $N(E) = 0$ at $E = \mu/\beta$. The halfwidth of $N(E)$ is determined by both μ and β; μ determines the asymetry of $N(E)$ about $E = 0$. It is noted, in order to conform with later work on electron energy loss (inelastic electron scattering), that $E > 0$ corresponds to an energy less than E_0, $E_0 - E$. The functions $F(K_0)$ and $N(E)$ are known for a given electron gun configuration or these functions may be measured (Andersen and Mol 1970).

4.1. *Chromatic coherence interval ΔK_0 is zero*

In the case of elastic electron scattering ψ_0 may be considered as independent of K_0 and equation (4) may be rewritten as

$$\psi_i(K_0, \boldsymbol{r}_i) = \int \psi_0(\boldsymbol{r}_0) G\left(K_0, \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \mathrm{d}\boldsymbol{r}_0. \tag{37}$$

G is dependent only on the modulus of $\boldsymbol{K}_0$ (Lenz 1965).

If electrons with different K_0 are incoherent, the image intensity j_i is calculated by a superposition of monochromatic electron intensities, that is

$$j_i(\boldsymbol{r}_i) = \int_0^\infty \int\int \psi_0(\boldsymbol{r}_0)\psi_0^*(\boldsymbol{r}_0')G\left(K_0, \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right)G^*\left(K_0, \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0'\right) \times F(K_0)\,\mathrm{d}\boldsymbol{r}_0\,\mathrm{d}\boldsymbol{r}_0'\,\mathrm{d}K_0. \tag{38}$$

Equation (38) becomes, on replacing ψ_0 and G by their respective Fourier transforms (equation (6))

$$j_i(\boldsymbol{r}_i) = \int_0^\infty \int\int S_0(\boldsymbol{\nu})S_0^*(\boldsymbol{\nu}')T(K_0, \boldsymbol{\nu})T^*(K_0, \boldsymbol{\nu}') \times \exp\left(-2\pi i(\boldsymbol{\nu}-\boldsymbol{\nu}')\cdot\frac{\boldsymbol{r}_i}{M}\right)F(K_0)\,\mathrm{d}\boldsymbol{\nu}\,\mathrm{d}\boldsymbol{\nu}'\,\mathrm{d}K_0 \tag{39}$$

or, on replacing K_0 by E

$$j_i(\boldsymbol{r}_i) = \int_{-\infty}^{+\infty} \int \int S_0(\boldsymbol{\nu}) S_0^*(\boldsymbol{\nu}') T(E, \boldsymbol{\nu}) T^*(E, \boldsymbol{\nu}') \times \exp\left(-2\pi i(\boldsymbol{\nu}-\boldsymbol{\nu}') \cdot \frac{\boldsymbol{r}_i}{M}\right) N(E)\, d\boldsymbol{\nu}\, d\boldsymbol{\nu}'\, dE. \tag{40}$$

$T(E, \boldsymbol{\nu})$ now includes a chromatic aberration term, that is, for a lens subject to spherical aberration, chromatic aberration and defocusing

$$T(E, \boldsymbol{\nu}) = \frac{1}{M} \exp\left(-\frac{2\pi i}{\lambda(E)} W(E, \boldsymbol{\nu})\right) B(\boldsymbol{\nu})$$

with (41)

$$W(E, \boldsymbol{\nu}) = \frac{C_s}{4} \lambda(E)^4 \nu^4 + \frac{C_c E \lambda(E)^2}{2E_0} \nu^2 + \frac{\Delta f}{2} \lambda(E)^2 \nu^2.$$

C_c is the third order chromatic aberration constant of the objective lens. $\lambda(E)$ is formally an energy dependent term, with E_0 very nearly constant, that is

$$\lambda(E) = \frac{12{\cdot}26}{(E_0 - E)^{1/2}} \{1 + 0{\cdot}978 \times 10^{-6}(E_0 - E)\}^{-1/2}\ \text{Å}$$

with E_0 and E in eV. If E_0 is 100 keV, then the thermal energy width, which is less than 2 eV, produces a deviation from λ_0 of less than 1 part in 10^5; hence in equation (41) $\lambda(E)$ may be replaced by λ_0.

In order to obtain equation (40) in a more practical form, the variable $\boldsymbol{\nu}$ is transformed to $\boldsymbol{\theta}$ space using equation (8)

$$j_i(\boldsymbol{r}_i) = \left(\frac{K_0}{2\pi}\right)^2 \int_{-\infty}^{+\infty} \int \int \Psi(\boldsymbol{\theta}) \Psi^*(\boldsymbol{\theta}') H(E, \boldsymbol{\theta}) H^*(E, \boldsymbol{\theta}') \times \exp\left(-\frac{iK_0}{M} (\boldsymbol{\theta} - \boldsymbol{\theta}') \cdot \boldsymbol{r}_i\right) N(E) d\boldsymbol{\theta}\, d\boldsymbol{\theta}'\, dE \tag{42}$$

where

$$H(E, \boldsymbol{\theta}) = \frac{1}{M} \exp\{-iK_0 \gamma(E, \theta)\} D(\boldsymbol{\theta}) \tag{43}$$

with $D(\boldsymbol{\theta})$ defined by equation (13*a*) and from equation (41)

$$\gamma(E, \theta) = \frac{C_s \theta^4}{4} + \frac{C_c E \theta^2}{2E_0} + \frac{\Delta f \theta^2}{2}. \tag{44}$$

For an object with a centre of symmetry equation (42) simplifies to (in bright field electron microscopy)

$$j_i(r_i) = \frac{K_0^2}{M^2} \int_{-\infty}^{+\infty} \int_0^{\alpha} \int_0^{\alpha} \Psi(\theta) \Psi^*(\theta') \exp\{-iK_0 \gamma(E, \theta)\} \exp\{iK_0 \gamma(E, \theta')\} \times J_0\left(\frac{K_0 r_i \theta}{M}\right) J_0\left(\frac{K_0 r_i \theta'}{M}\right) N(E) \theta\, d\theta\, \theta'\, d\theta'\, dE. \tag{45}$$

If the incident electron beam is monochromatic $N(E) = \delta(E)$ and the integrations over E of equations (42) and (45) lead to the coherent equations (15) and (18) respectively.

4.2. *The effect of chromatic aberration on image formation*

Equation (44) shows clearly that the phase shift, introduced into the scattered electron wave by chromatic aberration of the objective lens, is a least comparable with the corresponding phase shift introduced by spherical aberration. The phase shifts introduced by chromatic aberration and spherical aberration are respectively $-K_0C_cE\theta^2/2E_0$ and $-K_0C_s\theta^4/4$; the values of C_c and C_s are typically 0·2 cm for $E_0 = 20$–100 keV. If α is taken to be 0·005 rad for incident electrons of energy 100 keV and $\alpha = 0{\cdot}01$ rad for $E_0 = 20$ keV, the following phase shifts may be calculated for $\theta = \alpha$ and $E = 1$ eV:

20 keV	CA −3·66 rad	SA −3·66 rad	$\theta = 0{\cdot}01$ rad
100 keV	CA −0·42 rad	SA −0·53 rad	$\theta = 0{\cdot}005$ rad.

For $\theta < \alpha$, the chromatic aberration term is dominant because of the θ^2 dependence in contrast to the θ^4 dependence of the spherical aberration term. The only quantitative consideration of the chromatic aberration previously made appears within the framework of linear transfer theory (Hanszen and Trepte 1970, 1971). However, the transfer function T calculated by Hanszen and Trepte appears to result from an explicit assumption of chromatic coherence over the complete energy distribution.

It is noted that the angular dependence of the chromatic aberration and defocusing terms are identical. It is common practice in conventional transmission electron microscopy to underfocus ($\Delta f < 0$) the objective lens in an attempt partially to cancel the spherical aberration term. It is suggested that defocusing can be effective in a partial cancellation of the chromatic aberration term.

In order to estimate the effect of chromatic aberration on the image intensity j_i, the integration over E in equation (45) is considered, that is

$$L(\theta, \theta') = \int_{-\infty}^{+\infty} \exp\{-\mathrm{i}K_0\gamma(E, \theta)\} \exp\{\mathrm{i}K_0\gamma(E, \theta')\} N(E)\, \mathrm{d}E \tag{46}$$

or excluding terms not dependent on E

$$L(\theta, \theta') = \int_{-\infty}^{+\infty} \exp\left(-\frac{\mathrm{i}K_0C_cE}{2E_0}(\theta^2 - \theta'^2)\right) N(E)\, \mathrm{d}E \tag{47}$$

which represents the Fourier transform of $N(E)$. If as an explicit example $N(E)$ is taken to be of a gaussian form $(\beta/\pi)^{1/2} \exp(-\beta E^2)$ (an approximation to equation (36) with μ large), then

$$L(\theta, \theta') = \exp\left\{-\left(\frac{K_0C_c}{2E_0}\right)^2 \frac{(\theta^2 - \theta'^2)^2}{4\beta}\right\}$$

or

$$L(\theta) \simeq \exp\left\{-\left(\frac{K_0C_c}{2E_0}\right)^2 \frac{\theta^4}{4\beta}\right\}. \tag{48}$$

Hence the chromatic aberration effect may be considered as an attenuation term within the θ integration (equation (45)). The higher spatial frequencies ν transmitted by the objective aperture are decreased in intensity when chromatic aberration is included.

The contrast in the image at $r_i = 0$, will now depend on the factor

$$\exp\left\{-\left(\frac{K_0 C_c}{2E_0}\right)^2 \frac{\theta^4}{4\beta}\right\} \sin\{K_0\chi(\theta)\}$$

instead of the $\sin\{K_0\chi(\theta)\}$ term only (§ 2, equation (19)). In order to maintain the contrast, it is evident that α_{opt} must be decreased below the previous value (§ 2). If α_{opt} is maintained at these previous values, then the exponential factor in equation (48) has a value of 0·83 for $E_0 = 100$ keV (halfwidth of thermal distribution is 1 eV, $\alpha_{\mathrm{opt}} = 0{\cdot}0093$ rad) and a value of 0·12 for $E_0 = 20$ keV ($\alpha_{\mathrm{opt}} = 0{\cdot}0115$ rad). An increase in the thermal halfwidth to 1·5 eV (eg Boersch effect) causes still larger attenuations in the $\sin\{K_0\chi(\theta)\}$ function (eg 0·65 for $E_0 = 100$ keV). An analytic criterion for the choice of α_{opt} is not possible when the chromatic aberration term is included but numerical calculations indicate that α_{opt} should be decreased to

$$\left.\begin{aligned} \alpha_{\mathrm{opt}} &= \left\{\frac{C_s}{4\lambda_0} + \left(\frac{K_0 C_c}{2E_0}\right)^2 \frac{1}{\beta}\right\}^{-1/4} \\ \text{and} \qquad \Delta f_{\mathrm{opt}} &= -\left\{\frac{1}{\lambda_0 C_s} + \left(\frac{K_0 C_c}{E_0 C_s}\right)^2 \frac{1}{\beta}\right\}^{-1/2}. \end{aligned}\right\} \tag{49}$$

If $\beta = 2{\cdot}77\ \mathrm{eV}^{-2}$ (thermal energy halfwidth 1 eV), then $\alpha_{\mathrm{opt}} = 0{\cdot}0081$ rad ($E_0 = 100$ keV) and $\alpha_{\mathrm{opt}} = 0{\cdot}0067$ rad ($E_0 = 20$ keV); as in previous calculations $C_s = C_c = 0{\cdot}2$ cm. As expected the revised α_{opt} for $E_0 = 20$ keV is modified considerably when chromatic aberration is taken into account.

It is readily verified that if $N(E) = \delta(E)$, then equation (47) becomes $L(\theta, \theta') = 1$, and the equation for the image intensity j_i is

$$j_i(r_i) = \frac{K_0^2}{M^2} \left| \int_0^\alpha \Psi(\theta) \exp\{-iK_0\chi(\theta)\} J_0\left(\frac{K_0 r_i \theta}{M}\right) \theta \, d\theta \right|^2. \tag{50}$$

If $N(E)$ is a uniform distribution of energy, then $L(\theta, \theta') = \delta(\theta - \theta')$ and j_i is given by

$$j_i(r_i) = \frac{K_0^2}{M^2} \int_0^\alpha \left| \Psi(\theta) \exp\{-iK_0\chi(\theta)\} J_0\left(\frac{K_0 r \theta}{M}\right) \right|^2 \theta \, d\theta. \tag{51}$$

Equation (50) corresponds to chromatic and spatial coherence of the incident electron beam and equation (51) to complete chromatic incoherence and spatial coherence.

4.3. *Chromatic coherence interval* ΔK_0 *is finite*

Although it is improbable that the incident electron beam will exhibit phase coherence over the complete energy distribution, it is possible that coherence may be preserved in a small energy interval ΔE between E and $E + \Delta E$. The analysis of § 4.1 assumed that $\Delta E \to 0$ and in this section the analysis of § 4.1 is extended to a consideration of partial coherence in the energy distribution of the incident electron beam. In equation (40), the component wavefunctions $\psi_i(E, \boldsymbol{r}_i)$ in the image plane have been multiplied by $\psi_i^*(E, \boldsymbol{r}_i)$ before performing the integration over E. If the electrons in the energy interval E, $E + \Delta E$ exhibit coherence, then the $\psi_i(E, \boldsymbol{r}_i)$ are first summed over ΔE before multiplication by the complex conjugate. Provided

ΔE is small, equation (40) may be rewritten as

$$j_i(\boldsymbol{r}_i) = \sum_{n\Delta E=-\infty}^{n\Delta E=+\infty} \int\int S_0(\boldsymbol{\nu}) S_0^*(\boldsymbol{\nu}') T(n\Delta E, \boldsymbol{\nu}) T^*(n\Delta E, \boldsymbol{\nu}') \exp\left(-2\pi \mathrm{i} (\boldsymbol{\nu}-\boldsymbol{\nu}') \cdot \frac{\boldsymbol{r}_i}{M}\right) \times N'(n\Delta E)\, \mathrm{d}\boldsymbol{\nu}\, \mathrm{d}\boldsymbol{\nu}' \quad (52)$$

where

$$N'(n\Delta E) = \int_{n\Delta E}^{(n+1)\Delta E} N(E)\, \mathrm{d}E$$

and $T(n\Delta E, \boldsymbol{\nu}) \simeq$ constant within a given ΔE. It is seen that for $\Delta E \to 0$, equation (52) reduces to the chromatic incoherent equation (40). If ΔE becomes large, implying a chromatic coherence, $N(E) \simeq 1$ and the transfer function T contains an energy term $\bar{E}$, which reflects the average effect of chromatic aberration. Equation (52) then becomes

$$j_i(\boldsymbol{r}_i) = \left| \int S_0(\boldsymbol{\nu}) T(\bar{E}, \boldsymbol{\nu}) \exp\left(-2\pi \mathrm{i}\, \boldsymbol{\nu} \cdot \frac{\boldsymbol{r}_i}{M}\right) \mathrm{d}\boldsymbol{\nu} \right|^2. \quad (53)$$

Such an assumption has been made in the analysis of Hanszen and Trepte (1970, 1971).

As in § 4.1, equation (52) can be transformed into $\boldsymbol{\theta}$ coordinates in the back focal plane of the objective lens.

It is noted that although the analysis of §§ 4.1, 4.2 and 4.3 considered only the thermal energy distribution of the incident electron beam, other chromatic aberration effects can be included. In particular, one can consider the fluctuations in the accelerating voltage $\Delta E_0(t)$ and fluctuations in the objective lens current $\Delta I(t)$. In contrast to the analysis on the thermal energy distribution, these time dependent variations can only give rise to an incoherent superposition of image intensities; the incoherence arises because the detecting system (eg photographic plate) records, at a given instant of time, intensity. Hence the image intensity j_i is calculated from a superposition of image intensities over the period t_0 of recording, that is

$$j_i(\boldsymbol{r}_i) = \int_0^{t_0} \int\int \psi_0(\boldsymbol{r}_0) \psi_0^*(\boldsymbol{r}_0') \, G\left(\Delta E_0(t), \Delta I(t), \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \times G^*\left(\Delta E_0(t), \Delta I(t), \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0'\right) \mathrm{d}\boldsymbol{r}_0\, \mathrm{d}\boldsymbol{r}_0'\, \mathrm{d}t. \quad (54)$$

The effect of the thermal energy spread has been omitted for simplification; this corresponds to a further integration over E. As in § 4.1, ψ_0 and G may be replaced by their Fourier transforms (equation (6)) to give

$$j_i(\boldsymbol{r}_i) = \int_0^{t_0} \int\int S_0(\boldsymbol{\nu}) S_0^*(\boldsymbol{\nu}') T(\Delta E_0(t), \Delta I(t), \boldsymbol{\nu}) T^*(\Delta E_0(t), \Delta I(t), \boldsymbol{\nu}') \times \exp\left(-2\pi \mathrm{i} (\boldsymbol{\nu}-\boldsymbol{\nu}') \cdot \frac{\boldsymbol{r}_i}{M}\right) \mathrm{d}\boldsymbol{\nu}\, \mathrm{d}\boldsymbol{\nu}'\, \mathrm{d}t \quad (55)$$

or in $\boldsymbol{\theta}$ space

$$j_i(\boldsymbol{r}_i) = \left(\frac{K_0}{2\pi M}\right)^2 \int_0^{t_0} \int\int \Psi(\boldsymbol{\theta}) \Psi^*(\boldsymbol{\theta}') \exp\{-\mathrm{i}\, K_0 \rho(\Delta E_0(t), \Delta I(t), \boldsymbol{\theta})\} D(\boldsymbol{\theta}) \times \exp\{\mathrm{i}\, K_0 \rho(\Delta E_0(t), \Delta I(t), \boldsymbol{\theta}')\} D(\boldsymbol{\theta}') \exp\left(\frac{-\mathrm{i}\, K_0}{M} (\boldsymbol{\theta}-\boldsymbol{\theta}') \cdot \boldsymbol{r}_i\right) \mathrm{d}\boldsymbol{\theta}\, \mathrm{d}\boldsymbol{\theta}'\, \mathrm{d}t. \quad (56)$$

The aberration function ρ is given by

$$\rho(\Delta E(t), \Delta I(t), \boldsymbol{\theta}) = \frac{C_s\theta^4}{4} + \frac{\Delta f\theta^2}{2} + \frac{C_c\theta^2}{2}\left(\frac{\Delta E_0(t)}{E_0} - \frac{2\Delta I(t)}{I}\right). \tag{57}$$

Note that although $\Delta I > 0$ corresponds to an increase in the objective lens current, $\Delta E_0 > 0$ represents a decrease in the accelerating voltage; this is to conform with the definition of E in § 4.

The variations in E_0, I can be considered as effective changes in the focal length of the objective lens, that is (see Hanszen 1971)

$$\Delta f' = C_c\left(\frac{\Delta E_0(t)}{E_0} - \frac{2\Delta I(t)}{I}\right).$$

An obvious difficulty in assessing the effects of ΔE_0 and ΔI on the image intensity j_i is deciding on a functional form for these time variations. If an analytic form is chosen, for example, gaussian, maxwellian, sinusoidal or linear fluctuations, then the integral over t in equation (56) may be evaluated. In a similar way to that outlined in § 4.2, it is then possible to assess the effects of these chromatic defects on j_i. In high resolution electron microscopes $\Delta E_0/E_0$ and $\Delta I/I$ have been reduced to 1 part in 10^5; it is evident that stabilities of 1 part in 10^6 are required before these voltage and current fluctuations are negligible in comparison with the thermal energy spread.

Hanszen and Trepte (1970, 1971) have considered the effects of variations in E_0 and I on the transfer function. As with the analysis on the thermal energy spread, Hanszen and Trepte include only the time average effects of ΔE_0 and ΔI in the transfer function T.

5. The effect of spatial and chromatic incoherence on the image

The practical case in which the incident electron beam has both an angular and energy distribution is now considered. The analysis is a simple extension of the content of §§ 3 and 4. The general case of an angular-energy distribution $F(\boldsymbol{K}_0, K_0)$ is considered and it is assumed that electrons with different $\boldsymbol{K}_0$ and K_0 are incoherent. The image wave is then, from equations (22) and (37)

$$\psi_i(\boldsymbol{K}_0, K_0, \boldsymbol{r}_i) = \int \psi_0(\boldsymbol{r}_0)\, G\left(K_0, \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \exp(\mathrm{i}\,\boldsymbol{K}_0 \cdot \boldsymbol{r}_0)\, \mathrm{d}\boldsymbol{r}_0 \tag{58}$$

and the image intensity is given by

$$j_i(\boldsymbol{r}_i) = \int_0^\infty \iiint \psi_0(\boldsymbol{r}_0) G\left(K_0, \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0\right) \psi_0^*(\boldsymbol{r}_0') G^*\left(K_0, \frac{\boldsymbol{r}_i}{M} - \boldsymbol{r}_0'\right) \times \exp\{\mathrm{i}\boldsymbol{K}_0 \cdot (\boldsymbol{r}_0 - \boldsymbol{r}_0')\} F(\boldsymbol{K}_0, K_0)\, \mathrm{d}\boldsymbol{r}_0\, \mathrm{d}\boldsymbol{r}_0'\, \mathrm{d}\boldsymbol{K}_0\, \mathrm{d}K_0. \tag{59}$$

$F(\boldsymbol{K}_0, K_0)$ is normalized such that

$$\int_0^\infty \int F(\boldsymbol{K}_0, K_0)\, \mathrm{d}\boldsymbol{K}_0\, \mathrm{d}K_0 = 1.$$

For the incident electron beam $F(\boldsymbol{K}_0, K_0)$ is a separable function of $\boldsymbol{K}_0$ and K_0. From the general equation (59), the equations of § 2 (chromatic and spatial coherence), § 3 (chromatic coherence and spatial incoherence) and § 4 (chromatic incoherence and spatial coherence) can be derived.

In order to have equation (57) in the form given for previous cases, the equation is transformed to $\boldsymbol{\theta}$ coordinates

$$j_{\mathrm{i}}(\boldsymbol{r}_{\mathrm{i}}) = \left(\frac{K_0}{2\pi}\right)^2 \int_{-\infty}^{+\infty} \iiint \Psi(\boldsymbol{\theta}_{\mathrm{c}}+\boldsymbol{\theta})\Psi^*(\boldsymbol{\theta}_{\mathrm{c}}+\boldsymbol{\theta}')H(E,\boldsymbol{\theta})H^*(E,\boldsymbol{\theta}') \times \exp\left(-\frac{\mathrm{i}\,K_0}{M}(\boldsymbol{\theta}-\boldsymbol{\theta}')\,.\,\boldsymbol{r}_{\mathrm{i}}\right) I_0(\boldsymbol{\theta}_{\mathrm{c}})N(E)\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{\theta}'\,\mathrm{d}\boldsymbol{\theta}_{\mathrm{c}}\,\mathrm{d}E. \quad (60)$$

$H(E, \boldsymbol{\theta})$ is the wave aberration function with the chromatic aberration term included (equations (43) and (44)).

As mentioned in § 3, the integration over $\boldsymbol{\theta}_{\mathrm{c}}$ in equation (60) is evaluated after performing the $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ integrations.

6. Conclusion

It has been demonstrated that optical transfer theory (Hopkins 1953, 1955) can be applied to image formation in the transmission electron microscope. The main third order aberrations, namely spherical aberration, chromatic aberration, axial astigmatism and defocusing, can be included in the calculation of the image intensity. In particular, two effects of importance in the transmission electron microscope, namely the spatial coherence and chromatic coherence of the incident electron beam, have been included in the calculation of j_{i}. The final expressions for j_{i} are suited to numerical evaluation.

An approximation made in the present work is the neglect of Fresnel diffraction: this corresponds to the omission of obliquity factors such as $\exp(\mathrm{i}K_0 r_0^2/2z)$ (z is the distance between the electron source and the specimen) in the Fourier transform of the object wavefunction $\psi_0(\boldsymbol{r}_0)$ (equation (6)). In practice this approximation is valid provided that the object does not exhibit any sharp discontinuities in structure, such as an edge. Provided that no phase shifts are introduced into the diffracted wave by, for example, defocusing the objective lens, then Fresnel diffraction does not sensibly affect the image. However, Fresnel fringes due to an edge or hole in a specimen film are a dominant feature of the image for $\Delta f = \pm 2000$ Å. As with the Fraunhofer diffraction pattern, Fresnel fringes are adversely affected by the angular and energy spread of the incident electron beam, by the lens aberrations and by instabilities in the accelerating voltage and objective lens current. It is noted that although Fresnel fringes are a useful guide to the spatial and chromatic coherence of the electron source and to the objective lens defects, Fresnel fringes give very little information on the actual specimen structure.

Acknowledgments

The author is pleased to acknowledge the valuable assistance given by Professor F. Lenz in the form of a copy of his lecture notes on 'Transfer of information in the electron microscope'. The author is grateful to Professor R. E. Burge for discussion and also to the University of London for the provision of excellent computing facilities.

Refererences

Andersen, W. H. J., and Mol, A., 1970, *J. Phys.* D: *Appl. Phys.*, **3**, 965–79.
Born, M., and Wolf, E., 1959, *Principles of Optics* (London: Pergamon).
Burge, R. E., Misell, D. L., and Smart, J. W., 1970, *J. Phys.* C: *Solid St. Phys.*, **3**, 1661–72.

Crewe, A. V., 1970, *Q. Rev. Biophys.*, **3,** 137–75.
Crick, R. A., and Misell, D. L., 1971, *J. Phys.* D: *Appl. Phys.*, **4,** 1–20.
Eisenhandler, C. B., and Siegel, B. M., 1966, *J. appl. Phys.*, **37,** 1613–20.
Haine, M. E., 1957, *J. Sci. Instrum.*, **34,** 9–15.
Hall, C. R., and Hines, R. L., 1970, *Phil. Mag.*, **21,** 1175–86.
Hanszen, K.-J., 1969, *Z. angew. Phys.*, **27,** 125–31.
—— 1971, *Advances in Optical Electron Microscopy*, Vol. 4, eds R. Barer and V. E. Cosslett (New York: Academic Press), Pp. 1–84.
Hanszen, K.-J., and Trepte, L., 1970, *Proc. 7th Int. Conf. Electron Microscopy, Grenoble*, Vol. 1 (Paris: Societé Français de Microscopie Electronique), Pp. 45–6.
—— 1971, *Optik*, **32,** 519–38.
Heidenreich, R. D., and Hamming, R. W., 1965, *Bell System Tech. J.*, **44,** 207–33.
Hopkins, H. H., 1951, *Proc. R. Soc.* A, **208,** 263–77.
—— 1953, *Proc. R. Soc.* A, **217,** 408–32.
—— 1955, *Proc. R. Soc.* A, **231,** 91–103.
Lenz, F., 1965, *Lab. Invest.*, **6,** 808–18.
Niehrs, H., 1969, *Optik*, **30,** 273–93.
—— 1970, *Optik*, **31,** 51–71.
Reimer, L., 1969, *Z. Naturf.*, **24a,** 377–89.
Scherzer, O., 1949, *J. appl. Phys.*, **21,** 20–8.
Sneddon, I. N., 1951, *Fourier Transforms* (New York: McGraw-Hill).
Uyeda, R., 1955, *J. Phys. Soc. Japan*, **10,** 256–64.
Zeitler, E., 1966, *Proc. 6th Int. Conf. Electron Microscopy, Kyoto*, Vol. 1 (Tokyo: Maruzen), Pp. 43–4.
Zeitler, E. and Thomson, M. G. R., 1970a, *Optik*, **31,** 258–80.
—— 1970b, *Optik*, **31,** 359–66.