TRANSFER FUNCTIONS AND ELECTRON MICROSCOPE IMAGE FORMATION

P.W. Hawkes

Laboratoire d'Optique Electronique du C.N.R.S.
B.P. 4347, F-31055 Toulouse Cedex

1. BACKGROUND

Two types of electron microscope are capable of providing high-resolution information and it is only as the limit of resolution is approached that it is useful to discuss transfer theory. These are the conventional transmission electron microscope, in which an extended area of the specimen is illuminated with electrons, which subsequently form a magnified image of this area, and the scanning transmission electron microscope, in which a very small probe is scanned over the specimen in a regular raster and the image is formed point by point, all the incident electrons contributing to each image point (pixel). In the conventional instrument, the image contrast is generated by two mechanisms which, though not independent, may be conveniently treated separately. In order to understand these, we briefly recapitulate some practical aspects of electron microscopes. (See also the chapter by Wade in this volume.)

Commercial microscopes operate with electrons accelerated through a voltage of some 100 kV, which corresponds to a wavelength of the order of 0.04 Å (4 pm). They consist of an electron source, five or six magnetic lenses and a means of observing the image (fluorescent screen) and of recording it (photographic plate). The lenses form three groups, each with a different role. The first two act as *condensers*, and direct a fine nearly parallel beam of electrons on to the specimen with very little angular spread ; the illumination is thus spatially highly coherent. The third lens, the *objective*, is immediately downstream from the specimen, which is indeed often immersed within the objective lens. This is the only lens within which the electrons travel at a (comparatively) steep angle to the axis (a few milliradians) and the resolution of the instrument is limited by the spherical and chromatic aberrations of this lens. By light-optical standards these are huge and we are therefore obliged to work at a very small numerical aperture, with the result that the theoretical limit of resolution, suitably defined, is of the order of a few ångströms, or around a hundred wavelengths. Owing to the immense chromatic aberration, the electron accelerating voltage (i.e. wavelength) and lens currents are very highly stabilized, to better than one part in 10^5, and the illumination is thus highly coherent temporally as well as spatially. We note that, in common with all electron lenses, the focal length of the objective can be varied very

easily by altering the current in the windings. By this means, we can arrange that the plane conjugate to the (fixed) image plane coincides with the specimen or is at a certain distance from it, and through-focal series of images of the same specimen can readily be obtained.

Beyond the objective lens are formed the first intermediate image, in a plane conjugate to the specimen, and the diffraction pattern of the specimen, in a plane conjugate to the source. The remaining *intermediate and projector* lenses serve to magnify and image either of these onto the final "image" plane of the microscope, where the dimensions of the smallest structure recorded must be large enough to surpass any limitations due to the grain size of the recording medium.

The specimens used are essentially phase specimens, in the sense that virtually none of the incident electrons is halted - there is no absorption - but many electrons are deflected or "scattered" by the atoms of which the object is composed. Electrons scattered through large angles are intercepted by a small diaphragm, situated in the focal plane of the objective (the diffraction pattern plane), thus providing one of the two contrast mechanisms alluded to above: fewer electrons reach regions of the image corresponding to specimen areas containing heavy atoms, which are thus seen as dark patches in the image. This is known as diffraction contrast and is the dominant contrast mechanism except at high resolution. Electron deflection may be regarded as a phase shift of the incident electron wave function and we might therefore imagine that in the absence of some special accessory, analogous to a phase plate, no contrast would be generated at the image by electrons deflected in the specimen but not enough to strike the objective aperture. In fact, however, the combined effect of spherical aberration and of the defocus is similar to that of a phase plate of non-uniform thickness : a phase variation is added to that due to the specimen with the result that an interference pattern is generated at the image by unscattered and scattered electrons. This interference pattern is, for a certain type of specimen - weakly scattering objects- related simply to the atomic distribution in the latter and provides high resolution information, which cannot, however, be read off directly.

For weakly scattering specimens, we can show that the image contrast is related linearly to the phase and amplitude variations impressed on the incident electron wave by the specimen, and this linear relationship is fully characterized by a knowledge of the *phase and amplitude transfer functions*, as we show in the next section. Before turning to this, however, we must explain the presence of an amplitude term, which seems to contradict our earlier assertion that electron specimens are essentially phase specimens. The paradox is, however, easily resolved. When electrons are scattered within a specimen they may be deflected with virtually no loss of energy in which case the interaction is *elastic* or they may lose energy, in which case it is said to be *inelastic*. Transfer theory is applicable to elastically scattered electrons and any others that have been inelastically scattered are effec-

tively lost from the theory. They therefore appear as an amplitude or absorption term.

2. TRANSFER THEORY (COHERENT ILLUMINATION)

The illumination in a conventional microscope is to a good approximation coherent, though departures from perfect coherence are extremely important as we shall see in the next section. In order to introduce the notion of transfer functions without unnecessary complication, we consider a parallel beam of monoenergetic electrons, corresponding to plane monochromatic waves, incident on a specimen. The latter is characterized by a complex transparency S (x, y), which we write :

$$S = (1 - s) \exp (i \phi) \tag{1}$$

The wave emerging from the specimen is assumed to be equal to the product of S and the incident wave. The propagation of the wave function through the electron lens system is well understood and it is easily shown that the wave function at the image plane, ψ_i, is related to that at the object plane, ψ_o, by the linear expression :

$$\psi_i (x,y) = \frac{1}{M} \iint K (x/M - x_o, y/M - y_o) \psi_o (x_o, y_o) dx_o dy_o \tag{2}$$

in which M denotes the magnification and certain quadratic phase factors have been omitted. Writing :

$$\tilde{\psi}_i (p,q) = \iint \psi_i (Mx, My) \exp \{- 2 \pi i (px + qy)\} dx dy$$
$$\tilde{\psi}_o (p,q) = \iint \psi_o (x_o, y_o) \exp \{- 2 \pi i (px_o + qy_o)\} dx_o dy_o$$
$$\tilde{K} (p,q) = \iint K (x, y) \exp \{- 2 \pi i (px + qy)\} dx dy \tag{3}$$

we see that :

$$\tilde{\psi}_i (p,q) = \tilde{K} (p,q) \tilde{\psi}_o (p,q)/M \tag{4}$$

It is not the image wave function or its transform that is observable, however, but the current density, which is proportional to $\psi_i \psi_i^*$. Nevertheless, for weakly scattering specimens, useful relations can be derived. For such objects, we assume that both S and ϕ are so small that we may write :

$$S = (1 - s) \exp (i \phi) \simeq 1 - s + i\phi \tag{5}$$

whereupon we find :

$$M \tilde{\psi}_i (p,q) = \tilde{K} (p,q) \{\delta(p,q) - \tilde{s}(p,q) + i \tilde{\phi}(p,q)\} \tag{6}$$

or :

$$M^2 \psi_i (Mx, My) \psi_i^* (Mx, My)$$
$$= [1 - \iint \tilde{K} (\tilde{s} - i \tilde{\phi}) \exp \{2 \pi i (px + qy)\} dp dq]$$
$$\times [1 - \iint \tilde{K}^* (\tilde{s}^* + i \tilde{\phi}^*) \exp \{- 2 \pi i (px + qy)\} dp dq] \tag{7}$$

Since $\tilde{s}^* (-p, -q) = \tilde{s} (p, q)$ and likewise for $\tilde{\phi}$, we have :

$$M^2 \psi_i (Mx, My) \psi_i^* (Mx, My) - 1$$
$$\simeq - \iint \tilde{s} (p,q) \{\tilde{K} (p,q) + \tilde{K}^*(-p, -q)\} \exp \{2 \pi i (px + pq)\} dp\, dq$$
$$+ i \iint \tilde{\phi} (p,q) \{\tilde{K} (p,q) - \tilde{K}^*(-p, -q)\} \exp \{2 \pi i (px + qy)\} dp\, dq \qquad (8)$$

Denoting the left-hand side by C (Mx, My) and setting :

$$\tilde{C} (p, q) = \iint C (Mx, My) \exp \{- 2 \pi i (px + py)\} \, dx\, dy \qquad (9)$$

we see that :

$$\tilde{C} (p, q) = \tilde{s} (p, q) B_s (p, q)$$
$$+ \tilde{\phi} (p, q) B_\phi (p, q) \qquad (10)$$

where :

$$B_s (p, q) = - \{\tilde{K} (p, q) + \tilde{K}^* (p, q)\}$$
$$B_\phi (p, q) = i \{\tilde{K} (p, q) - \tilde{K}^* (p, q)\} \qquad (11)$$

Equation (10) states that the Fourier transform of the image contrast is linearly related to the transforms of the specimen amplitude (s) and phase (ϕ) distributions ; the functions B_s and B_ϕ are known as the amplitude and phase transfer functions. The transform of the point-spread function, K, is readily expressed in terms of the spherical aberration coefficient, C_s and the defocus, Δ. We find :

$$B_s = - 2 a \cos \gamma (p, q)$$
$$B_\phi = 2 a \sin \gamma (p, q) \qquad (12)$$

in which :

$$\gamma = \frac{2 \pi}{\lambda} \{ \frac{1}{4} C_s \lambda^4 (p^2 + q^2)^2 - \frac{1}{2} \Delta \lambda^2 (p^2 + q^2) \} \qquad (13)$$

and a is a function equal to unity in the opening of the objective aperture and to zero elsewhere.

The fact that the transfer functions are sinusoidal creates severe problems when we wish to synthesize a faithful image from the recorded current distribution : some contrast will be inverted but worse, much information will have been irretrievably lost in the vicinity of the zeros. Several procedures have been devised and tested, for regenerating the functions s and ϕ (or their transforms) from focal series, since the zeros will occur at different spatial frequencies as Δ is varied. The amplitude term can often be neglected, since specimens thin and light enough for the condition $\phi \ll 1$ to be satisfied will cause little inelastic scattering. We then have :

$$\tilde{C} = \tilde{\phi} B_\phi = 2 a \tilde{\phi} \sin \gamma \qquad (14)$$

This offers a convenient means of studying the transfer function of actual instruments, for by taking a very thin layer of amorphous material, typically carbon, as specimen, $\tilde{\phi}$ will be a very slowly varying function and the Fourier transform (or

spatial frequency spectrum) of the image contrast will show sin γ directly, as a series of concentric rings. Various techniques for extracting C_s and Δ from these diffractograms have been devised.

3. PARTIALLY COHERENT ILLUMINATION

Although the illumination in an electron microscope is highly coherent, the effects of finite source size (partial spatial coherence) and of finite wavelength spread (partial temporal coherence) finally limit the resolution attainable. It has long been known that for weakly scattering objects, a transfer function formalism can again be derived and an equation identical with eq. (10) is obtained, although the transfer functions are of course different. A particularly interesting result, which has been established much more recently, however, reveals that for most practical situations, the transfer function corresponding to partially coherent illumination may be written as a product of the coherent transfer function (B_s or B_ϕ) and two modulating functions, one characterizing finite source size, the other finite wavelength spread. That each aspect of partial coherence can be separately represented by a modulating function is not very difficult to show - we refer to the bibliography (section 5) for details of this. It is, however, much less obvious that, if both aspects have to be considered simultaneously, as is the case in practice, it is legitimate to multiply the corresponding modulation functions. That this is indeed usually permissible has now been demonstrated, however, and the product representation has been used to obtain estimates of the wavelength spread and source size of various electron microscopes.

What effect do these modulating functions have on the transfer of information from specimen to image ? Broadly speaking, they attenuate the recorded signal at the image so that beyond a certain point, determined by the magnitude of the wavelength spread and source size, this signal cannot be distinguished from the natural granularity or noise. They thus create a real, practical limit to the attainable resolution, whereas the unmodulated sinusoidal transfer functions suggest, unrealistically, that information can be obtained in principle far beyond the "theoretical limit of resolution".

4. OTHER CASES

In the preceding sections, we have given an elementary account of the transfer function formalism for the common bright-field mode of operation of conventional electron microscopes. We now mention briefly some of the other modes used in electron microscopy and examine the applicability of transfer function theory to them.

Several dark-field modes are routinely available on commercial electron microscopes, but only in exceptional circumstances is a transfer function theory applicable. The reason for this is readily understood when we realise that the exis-

tence of a *linear* bright-field theory is a consequence of the fact that weak scattered wave fields interfere with a comparatively strong unscattered beam. In the dark-field modes, however, this unscattered beam is intercepted and only the scattered electrons reach the image. The possibility of establishing a linear relation between some aspect of image and object has been explored in considerable detail ; in general, no such relation exists but an exception of some interest corresponds to individual discrete specimen points such as heavy atoms bound in some characteristic geometry to a lighter molecule.

A number of other bright field modes have been attracting considerable attention recently, particularly for high resolution microscopy and for minimizing the damage caused by electron-specimen interactions. There are tilted illumination, which gives increased resolution in the tilt direction, and hollow-cone, bright-field illumination, which should give an overall increase in resolution but at the expense of reduced contrast. Since high-resolution image information is almost inevitably difficult to detect, this may be too high a price to pay, for certain types of specimen at least, but is nonetheless of considerable theoretical interest. For both types of illumination, the image-object relation can be described by a transfer function formalism but for tilted illumination, the latter is anisotropic, in the sense that the microscope response varies with the azimuth, relative to that of the tilt direction. A thorough understanding of this type of illumination is nevertheless of importance, since it should be possible to obtain three-dimensional information about easily damaged specimens with a low electron dose by using a multiple tilt scheme. For hollow-cone illumination, the transfer functions are again axially symmetric but low in absolute value which implies a diminution of contrast, as mentioned above.

To what extent can this account of the transfer theory of image formation in the conventional fixed-beam transmission microscope be transferred to the scanning instrument (STEM) ? A detailed study of the formation of the STEM image shows that the recorded current is of course affected by the properties of the probe-forming lens but that only for weak phase specimens is the relation between the signal recorded and some simple aspect of the specimen straightforward. The effect of partial coherence on STEM image formation is very different ; it can be shown that the spatial frequency spectrum of the image intensity is the same as that of a STEM with a coherently illuminated probe, modulated by the mutual intensity in the exit pupil of the probe-forming lens.

5. FURTHER READING

The texts available on image formation in the electron microscope, transfer functions and the effect of partial coherence range from non-specialist works with little mathematics to highly technical papers that pre-suppose considerable familiarity with the subject. Of the non-specialist material of a general kind, we draw

attention to a book and two other accounts by the present author (Hawkes, 1972, 1975, 1978 c) and a very readable chapter by Lenz (1971). More technical discussion is to be found in a still very accessible chapter by Frank (1973 b), a long review and a book by Misell (1973, 1978) and a particularly clear book by Saxton (1978).

Transfer functions were first introduced into electron optics by Hanszen and his colleagues and the earlier work is reviewed in Hanszen (1971). The effect of temporal partial coherence was first analysed by Hanszen and Trepte (1971) and the fact that spatial partial coherence can be expressed by means of a modulating function was shown by Frank (1973 a) ; the combined effect of the two was fully investigated by Wade and Frank (1977). A long review of partial coherence in electron microscopy by Hawkes (1978 b) gives many more references in this field. Numerous papers have been devoted to tilted and conical illumination ; for a full list, we refer to a bibliography (Hawkes, 1978 a), only mentioning the papers by Hoppe et al. (1975), Wade and Jenkins (1978) and Saxton et al. (1978). The many methods of exploiting image diffractograms to obtain information about microscope operating parameters will likewise be found in the bibliography mentioned above ; we draw attention to papers by Saxton (1977), Wade (1978) and Frank et al. (1978/9), which are specifically concerned with partial coherence. The problem of filtering electron micrographs to reduce the effect of the transfer function is discussed very thoroughly by Kübler et al. (1978) and has also been explored in practice by Saxton et al.(1977).

Finally, we draw attention to the work of Burge and Dainty (1976) on STEM imagery, including the effect of partial coherence across the probe, to the very detailed analysis by Hanszen and Ade (1977) and to a paper by Spence and Cowley (1978). The question of dark-field image formation is analysed by, for example, Ade (1975).

REFERENCES

G. Ade : Nichtlinearitätsprobleme bei Hell- und Dunkelfeldabbildungen, Optik, 42, 199-215 (1975).

R.E. Burge and J.C. Dainty : Partially coherent image formation in the scanning transmission electron microscope, Optik, 46, 229-240 (1976).

J. Frank : The envelope of electron microscopic transfer functions for partially coherent illumination, Optik, 38, 519-536 (1973 a).

J. Frank : Computer processing of electron micrographs, in Advanced Techniques in Electron Microscopy, ed. by J.K. Koehler (Springer, Berlin and New York, 1973 b), pp. 215-274.

J. Frank, S.C. McFarlane and K.H. Downing : A note on the effect of illumination aperture and defocus spread in brightfield electron microscopy, Optik, 52, 49-60 (1978/9).

K.-J. Hanszen : The optical transfer theory of the electron microscope : fundamental principles and applications, Adv. Opt. Electron Micr., 4, 1-84 (1971).

K.-J. Hanszen and G. Ade : A consistent Fourier optical representation of image formation in the conventional fixed beam electron microscope, in the scanning transmission electron microscope and of holographic reconstruction, PTB-Bericht APh-11, 31 pp (1977).

K.-J. Hanszen and L. Trepte : Der Einfluss von Strom- und Spannungsschwankungen, sowie der Energiebreite der Strahlelektronen, auf Kontrastübertragung und Auflösung des Elektronenmikroskops, Optik 32, 519-538 (1971)

P.W. Hawkes : Electron Optics and Electron Microscopy (Taylor and Francis, London, 1972).

P.W. Hawkes : Computer processing of electron micrographs. Int. Rev. Cytol. 42, 103-126 (1975).

P.W. Hawkes : Electron image processing : a survey, Comp. Graph. Im. Proc. 8, 406-446 (1978 a).

P.W. Hawkes : Coherence in electron optics, Adv. Opt. Electron Micr. 7, 101-184 (1978 b).

P.W. Hawkes : Computer processing of electron micrographs, in Principles and Techniques of Electron Microscopy, ed. by M.A. Hayat (Van Nostrand-Reinhold, New York and London, 1978 c), vol. 8, pp. 262-306.

W. Hoppe, D. Köstler, D. Typke and N. Hunsmann : Kontrastübertragung für die Hellfeld-Bildrekonstruktion mit gekippter Beleuchtung in der Elektronenmikroskopie, Optik, 42, 43-56 (1975).

O. Kübler, M. Hahn and J. Seredynski : Optical and digital spatial frequency filtering of electron micrographs, Optik 51, 171-188 and 235-256 (1978).

F. Lenz : Transfer of image information in the electron microscope, in Electron Microscopy in Material Science, ed. by U. Valdrè (Academic Press, New York and London, 1971), pp. 540-569.

D.L. Misell : Image formation in the electron microscope with particular reference to the defects in electron-optical images, Adv. Electron. Electron Phys. 32, 63-191 (1973).

D.L. Misell : Image Analysis, Enhancement and Interpretation, (North Holland, Amsterdam, New York and Oxford, 1978).

W.O. Saxton : Spatial coherence in axial high resolution conventional electron microscopy, Optik 49, 51-62 (1977).

W.O. Saxton : Computer techniques for image processing in electron microscopy, Adv. Electron. Electron Phys., Suppl. 10, 289 pp (1978).

W.O. Saxton, A. Howie, A. Mistry and A. Pitt : Fact and artefact in high resolution microscopy, in Developments in Electron Microscopy and Analysis, 1977, ed. by D.L. Misell (Institute of Physics, Bristol, 1977) pp. 119-122.

W.O. Saxton, W.K. Jenkins, L.A. Freeman and D.J. Smith : TEM observations using bright field hollow cone illumination, Optik, 49, 505-510 (1978).

J.C.H. Spence and J.M. Cowley : Lattice imaging in STEM. Optik, 50, 129-142 (1978).

R.H. Wade : The phase contrast characteristics in bright field electron microscopy, Ultramicroscopy, 3, 329-334 (1978).

R.H. Wade and J. Frank : Electron microscope transfer functions for partially coherent axial illumination and chromatic defocus spread, Optik, 49, 81-92 (1977).

R.H. Wade and K.H. Jenkins : Tilted beam electron microscopy : the effective coherent aperture, Optik, 50, 1-17 (1978).