Proc. R. Soc. Lond. A. **329**, 327–359 (1972)
Printed in Great Britain

Electron microscopy of biological specimens by means of an electrostatic phase plate

By P. N. T. Unwin

M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge, England

(*Communicated by H. E. Huxley, F.R.S. – Received 3 March* 1972)

[Plates 23 to 29]

A new electron microscope imaging method has been developed that is especially suited to the study of thin biological materials. It involves the use of an electrostatic phase plate – a device which creates a more or less uniform difference in optical path between the unscattered and scattered waves by means of its electric field. This phase plate functions in an analogous manner to the absorbing bright contrast phase plate of light microscopy.

The contrast effects and aberrations peculiar to the method have been examined and are discussed in terms of their likely influence on the image's representation of the object structure. Analysis of electron micrographs of some biological test specimens, whose structure is relatively well known, confirms that this representation, to a resolution of *ca.* 0.85 nm, is a particularly faithful one. In the analysis the resolution limit was determined by the degree of specimen preservation, and a real limit, determined by the degree of spherical aberration in the objective lens, of *ca.* 0.5 nm is expected.

A special property of the imaging method, as distinct from the conventional bright field method, is that it emphasizes the detail within the biological material itself, but reduces the contrast from the surrounding film of stain; negative staining remains necessary only because it helps to preserve the morphology of the specimen during irradiation. Evidence is presented that this property enables the method to display information about the specimen that it would not be possible to detect with the bright field method.

1. Introduction

The phase plate (Zernike 1935) is well known as an invaluable device for enhancing detail in transparent biological specimens when examined in the light microscope. Analogous devices have been constructed to perform the same task in the electron microscope (see, for example, Kanaya & Kawakatsu 1958; Faget, Fagot & Fert 1960; Thon & Willasch 1970), but unfortunately these have so far met with only a limited degree of success. This is largely because of the greater technical difficulties involved: the electron optical equivalent is a thin film of amorphous material not more than a few hundred ångströms thick, and for a satisfactory performance this film needs to be constructed and positioned in the microscope extremely accurately (Thon & Willasch 1970).

As an alternative method for producing the phase shifts required for contrast enhancement, Boersch (1947) suggested the use of an electric field in the vicinity of the back focal plane of the objective lens. Such an electrostatic method promises to have several advantages over the thin film method; indeed, some of these were realized during preliminary experiments (Unwin 1970, 1971) with the device illustrated in figure 1, plate 23. This device – an electrostatic phase plate – consists of a small aperture with a thin poorly conducting thread spanning its diameter.

It is placed at the back focal plane of the objective lens (i.e. at the Fraunhofer diffraction plane), where a uniformly intense beam of unscattered electrons causes the central portion of the thread to become positively charged. An electric field is thus created which, because of the bounding geometry, is of a suitable shape to produce a uniform phase shift of the scattered wave over most of the region enclosed by the aperture. The required uniformity, stability and magnitude of electric potential are achieved by a mechanism involving secondary emission.

Unlike its thin film counterpart, this electrostatic phase plate leaves the amplitude of the scattered wave almost entirely undisturbed, it is relatively easy to construct and its positioning in the microscope is fairly critical in one direction only. In addition to these features, its characteristics of being able to weaken the unscattered wave (Möllenstedt 1950) and to create bright, rather than the normal dark, phase contrast appear to make it an especially suitable device for enhancing detail in the biological materials themselves, irrespective of the contrast due to the stain.

The principles involved in the design, construction and operation of this phase plate have been discussed elsewhere (Unwin 1971). The present paper is concerned with the details of the imaging process. These are of primary importance because, although the phase shifts generated have been shown previously to meet the requirements for providing a realistic image to a resolution of almost 0.5 nm, the phase plate introduces a number of aberrations and contrast effects not encountered in normal bright or dark field microscopy. Our objective is to discuss these special properties and to establish, by analysis of some typical micrographs, how faithfully this new type of image represents the original object structure.

The results of the latter analysis indicate that the phase plate image is a particularly realistic one. They also suggest that it can be used to obtain more information from biological specimens than could be obtained from a corresponding bright field image. The reason for this advantage stems from differences in the state of preservation and scattering properties of the specimen and the stain, and the ability of the phase plate to affect the contrast of these two materials in different ways.

2. The object and method of observation

The regularly repeating, rod shaped, stacked disk aggregate of tobacco mosaic virus protein was used as the test object for the present investigation. This material makes a particularly suitable test object because it has a well known structure and it creates characteristic and distinct diffraction maxima over a wide range of spatial frequencies when imaged in an electron beam. It is also sufficiently thin that, even when encased in a thin film of heavy metal stain, the contrast in its image can be realistically discussed by assuming single electron scattering events only are involved.

The 'stacked disk' specimens are built up from large numbers of protein rings, each divided into 17 identical subunits (Finch, Leberman, Chang & Klug 1966).

The rings themselves are arranged in pairs of 34 subunits; these constitute the true repeating unit, the disk. Although there may exist a certain degree of azimuthal disorder, each successive disk is in general rotated with respect to the preceding one by 3/10 of the subunit angle ($2\pi/17$ rad), leading to an axial repeat every ten disks (Finch & Klug 1971). The observed dimensions of this structure may vary considerably, depending on the preparation conditions, but typically the component protein rings each have an inside diameter of 4 nm and an outside diameter of 17 nm, and repeat every 2.5 nm.

The specimens were prepared for electron microscopy by supporting them on very thin carbon films (probably < 5 nm thick in most cases) and negatively staining them with uranyl acetate or formate. The purpose of the staining was not to provide contrast, as it is in bright field imaging – the biological material appears in good contrast by itself when the phase plate is used – but to help preserve the original morphology. Our experience and that of others (e.g. Williams & Fisher 1970) indicates that unless enclosed by a film of stain, the specimen almost instantaneously collapses or disintegrates in the electron beam.

In imaging these specimens, the position of exact focus and the correctness of the electrical conditions generated by the phase plate were judged by the appearance of the fine phase detail observable in the carbon support film at high magnifications. The procedure employed was first to focus the object with the beam of unscattered electrons passing through the region between the thread and the aperture edge, then move the thread into position and wait about five seconds for the potential at the centre of the thread to stabilize. During this latter period, providing the electrical conditions had been correctly set, the overall contrast and sharpness of the image would improve very obviously to stabilize at a distinct maximum. The appearance of this maximum was taken to indicate that the desired recording conditions had been achieved, its coincidence with conditions of optimum contrast transfer (see §4) having been established previously (Unwin 1971).

All the electron micrographs were taken with a Philips EM300 operating at 100 kV and having an objective lens with a focal length (f) of 1.6 mm and a spherical aberration coefficient of 1.6 mm. The dimensions of the phase plates used were as shown in figure 1, but with some small variation in thread diameter (generally between 0.3 and 0.5 μm); a range of second condenser apertures (diameters between 50 and 100 μm) was used with these, together with conditions of critical illumination,† to obtain a suitable variation in the proportion of the unscattered electrons contributing to the image.

† I.e. with the beam focused on the object plane. The unscattered electrons then form a demagnified image of the second condenser aperture at the plane of the phase plate which, since this aperture is more or less uniformly illuminated, is a uniformly intense disk. This uniform intensity is needed for the phase plate to generate the correct form of electric field.

3. Aberrations and artefacts in the phase plate image

There are two types of artefact peculiar to the phase plate image: those that arise from its incorrect adjustment – errors of positioning and of charge – and those that are an inherent feature of the different imaging system, e.g. the diffraction effect due to the thread. Both types are discussed below, while in §5.3 a comparison is made between a typical phase plate image containing artefacts with one that has been reconstructed to give a true representation of the object structure.

We note here that the phase plate method of observation is capable of providing such a realistic image only if, within the resolution considered, no imaging elements (Fourier components) are prevented from contributing to the image, and only if the interference pattern due to the scattered wave is formed against a uniform coherent background wave of unscattered electrons whose amplitude is at least sufficient that the summed amplitude for the total wave is always positive.† These conditions are met in practice by having the thread of the phase plate so thin that a part of the unscattered beam, and hence a part of any of the diffracted beams that it intercepts, can pass round it. The size of these beams at the diffraction plane is inversely proportional to the lateral coherence length of the beam illuminating the object; thus as the degree of coherence is increased there comes a point – when the beam (at the diffraction plane) and the thread diameters become equal – at which the phase plate no longer functions as a device for producing phase contrast, but instead functions as a dark field device. It may seem somewhat paradoxical therefore that although the process of image formation with the phase plate can be described in terms of the Abbe theory for coherent illumination, such ideal conditions do not strictly apply.

3.1. *The 'diffraction effect'*

It is well known from the wave theory of image formation that the image seen on the photographic plate essentially results from a process of twofold diffraction: the incident electron wave is diffracted first on passing through the object, and again by whatever restricting device is placed in the plane of its diffraction pattern. Thus the image is not simply the interference pattern due to the object alone (taking account of aberrations), but is constructed from the convolution of the (complex) amplitudes which would give rise to this pattern with the amplitude pattern, or Fourier transform, of the device in the diffraction plane.

The Fourier transform of the phase plate is essentially that of a circular opening bisected by a narrow, completely opaque strip: the thread in fact scatters most of the impinging electrons, rather than absorbs them; however, it can easily be demonstrated that the contribution made by those electrons scattered by the

† The summed amplitudes must all bear the same sign because the image records only their squares and therefore cannot distinguish between those that are positive and those that are negative. The nature of the 'spurious structures' created when this requirement is not met has been discussed by Hanszen (1969).

thread which remain in the field of view is negligible compared with the number of electrons that make up the image.

Now diffraction by a circular opening gives rise to the well known Airy pattern; this consists of a small, but strong, central maximum (the Airy disk) surrounded by a series of concentric bright and dark rings. Diffraction by an opaque strip, on the other hand, produces a long streak – consisting of a broad central maximum and a series of much weaker peaks – lying in a direction perpendicular to it. The phases in the central regions of the two amplitude functions describing these patterns are opposite and the Fourier transform of the phase plate is therefore their area-weighted difference.

The appearance of the central region of the phase plate transform and the difference between this and the transform of the circular opening alone can be seen from the light interference patterns reproduced in figure 2, plate 23. These patterns are analogous to those which would be formed at the image plane of the electron microscope under our conditions of observation in that they were produced against a uniform coherent background wave and therefore do represent the corresponding amplitude functions and not simply their squares (the intensity functions). Since the Airy disks are practically the same size in both cases (the phase plate actually forms the slightly smaller one) they demonstrate that the presence of the thread does not in any way imply loss in resolution. Note also that the (dark) streak due to the thread leads to a highlighting and not, as one might at first suspect, a smearing out of the (bright) Airy disk.

An example would not normally be found in electron microscope images showing much more than just the very central part of the perturbation due to the thread, as in the light optical analogue of figure 2*a*. The perturbation would normally appear in poor contrast here because the lateral coherence length involved is relatively short: with typical illumination conditions the lateral coherence length referred to the object plane, $\lambda/2\theta_c$ (where λ is the electron wavelength and θ_c is the semi-angle subtended at this plane by the second condenser aperture), is only *ca.* 7 nm, whereas the half-width of the central maximum referred to this plane typically extends over 14 nm. To make it as conspicuous as possible electron optically, therefore, a special phase plate having an exceptionally thick thread (about 1 μm in diameter) was constructed and used to examine the edges of a carbon film which had curled up into the incident beam so as to form lines of very strongly scattering material. The images obtained in this way were able to show not only the full extent of the broad central maximum but some of the weak subsidiary peaks as well. An example is given in figure 3, plate 24. To confirm that the observed peaks do indeed represent amplitudes, rather than intensities, they were densitometered and compared with the amplitude function appropriate to the diameter of thread that was used. The result is given in figure 4, where it can be seen that the actual positions of the peaks and the calculated positions of the amplitude peaks are in good agreement.

A more characteristic and much less disturbing manifestation of the 'diffraction

effect' is the very slight streakiness associated with the fine background structure in the normal phase plate images of the carbon support films (see figure 5, plate 24); other possible causes for the streakiness, such as astigmatism and vibration, must be discounted, as the same effect can easily be reproduced by inserting a wire in the back focal plane of the objective lens of a light microscope, with, for example, fine grained photographic emulsion as a test object.

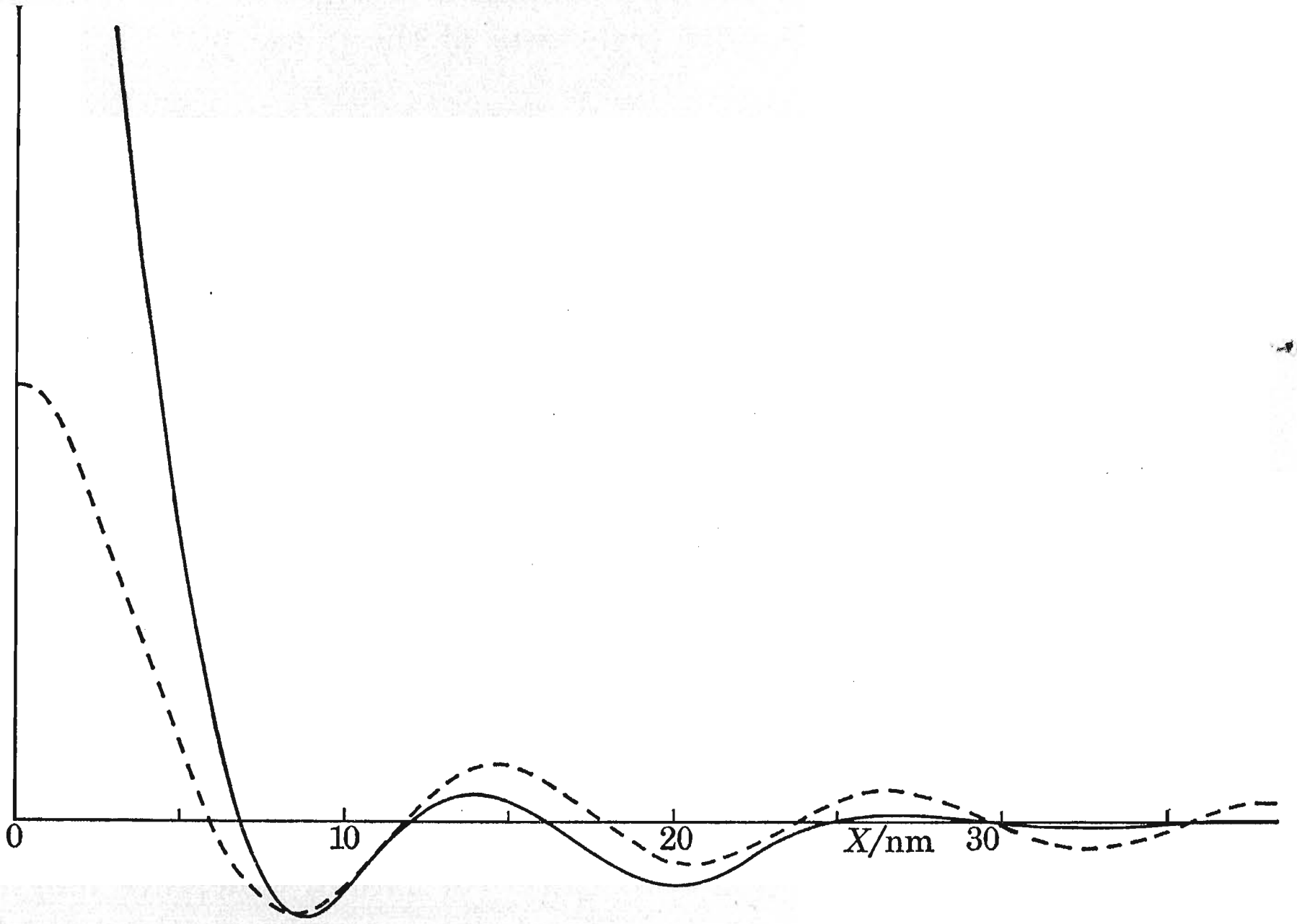
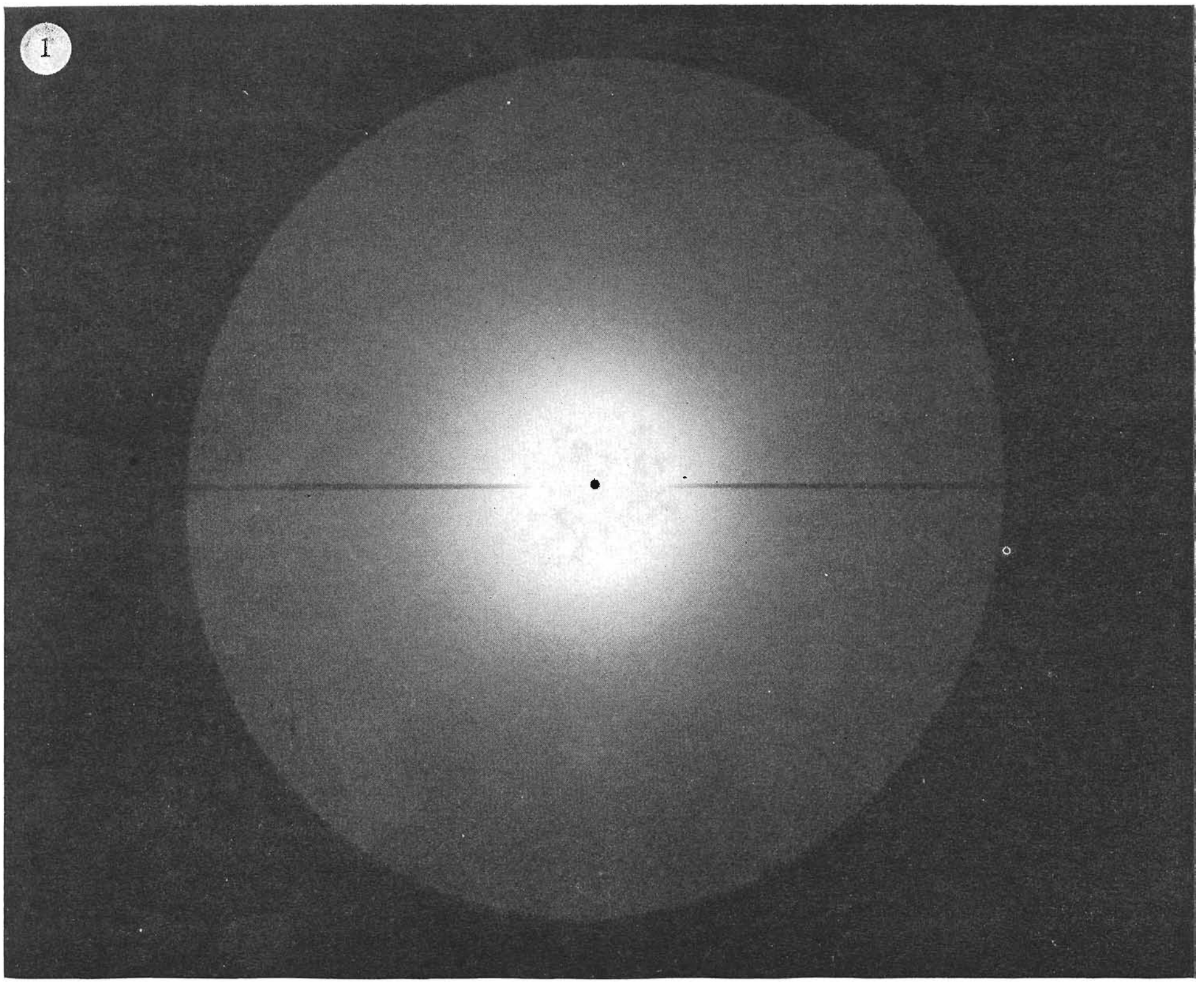


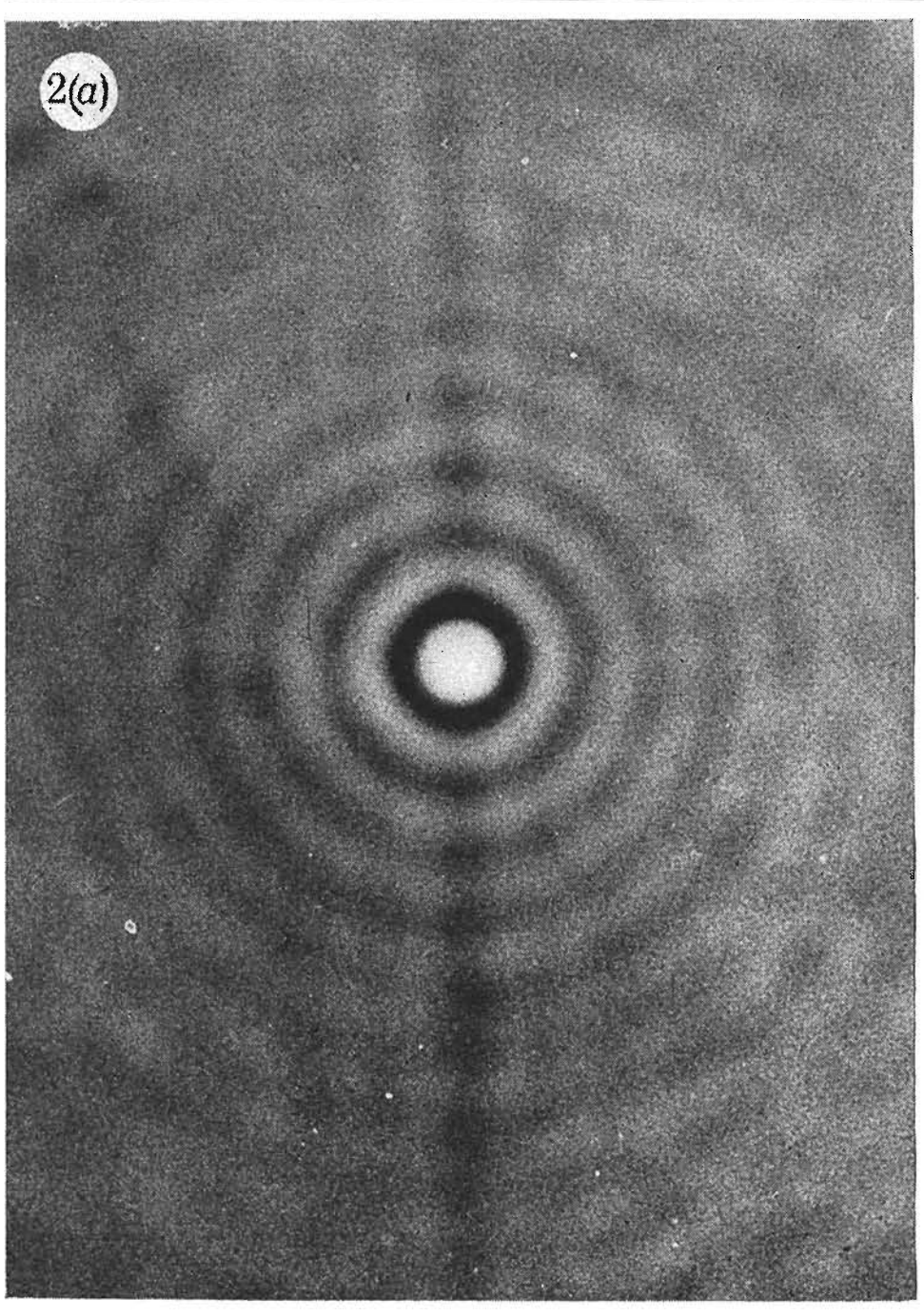
FIGURE 4. Comparison between the spacings of the diffraction fringes shown in figure 3 and the calculated positions of the amplitude peaks; ——, densitometer trace along the line indicated in figure 3; – – –, part of the amplitude function $\sin\gamma/\gamma$ due to a slit having a width, W, of one micron plotted against the X-coordinate in the object plane ($\gamma = \pi WX/f\lambda$).

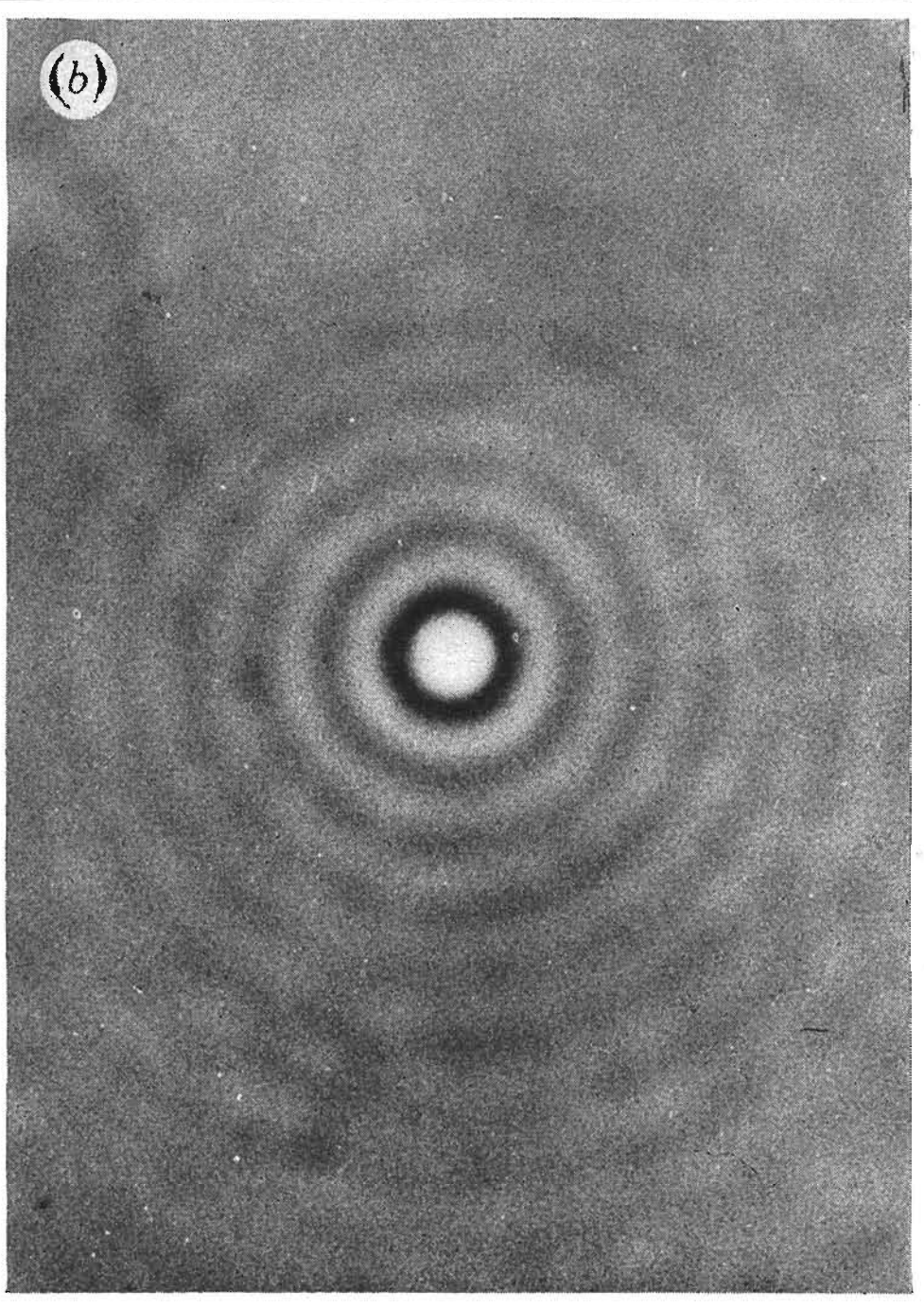
DESCRIPTION OF PLATE 23

FIGURE 1. The electrostatic phase plate in position at the back focal plane of the objective lens, illuminated by the scattering pattern from a carbon film: the circular aperture is a thin foil one and the filament extending diametrically across it is a spider's thread coated with a thin, discontinuous, layer of gold. A black dot has been drawn in the centre to indicate the typical size of the beam of unscattered electrons. (Magn. × 3000.)

FIGURE 2. The central regions of the Fraunhofer patterns formed by two different masks on placing them in the diffraction plane of a light-optical diffractometer having, in its object plane, a small hole in a very fine mesh screen. The screen transmits sufficient light that the patterns record the amplitude, rather than the intensity, variations. (*a*) is the pattern due to a mask having a similar geometry to that of the phase plate in figure 1 (but with a small hole in the centre to let the 'unscattered beam' through), and (*b*) is the pattern due to the same mask, but with the 'thread' removed.

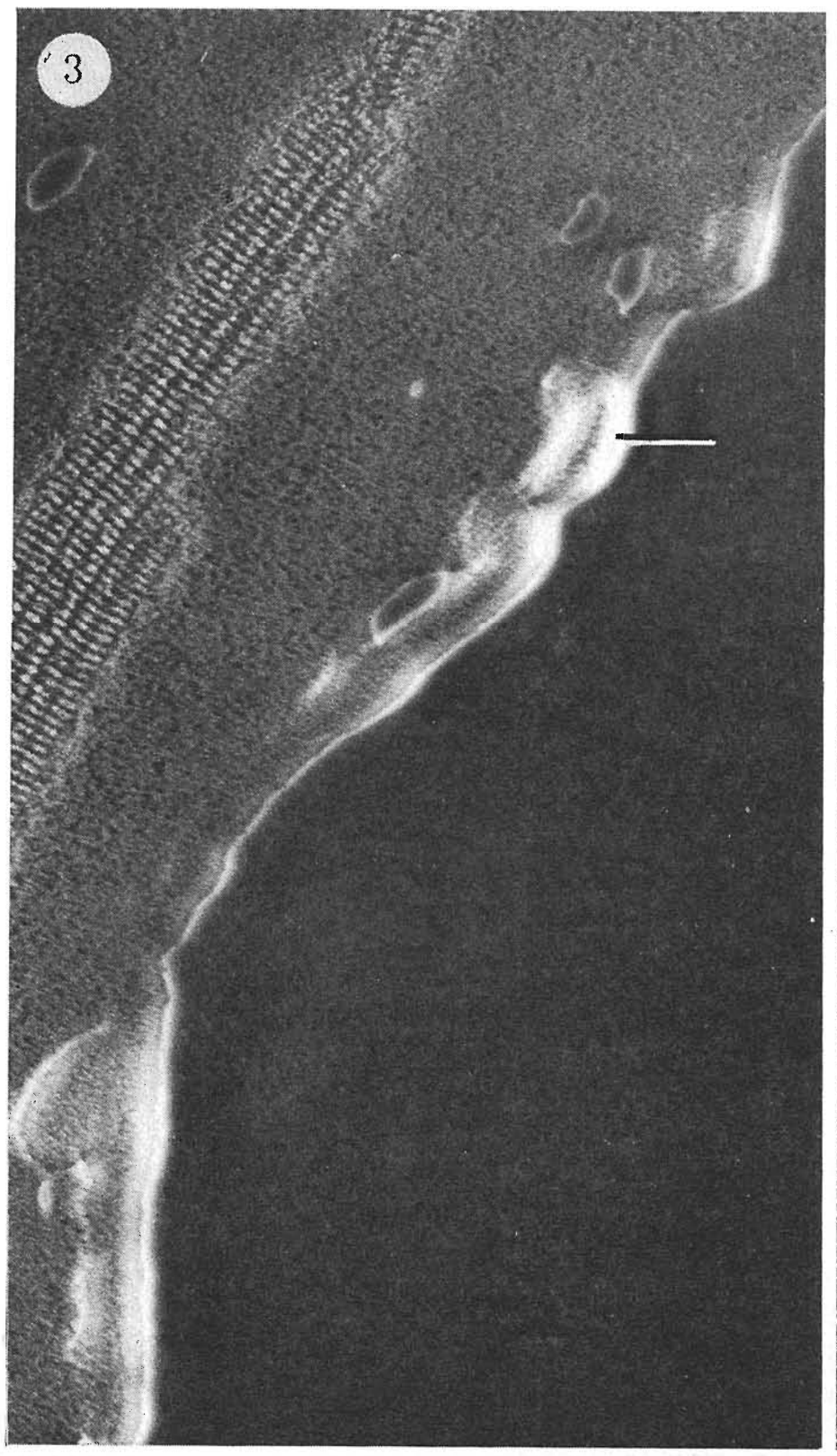


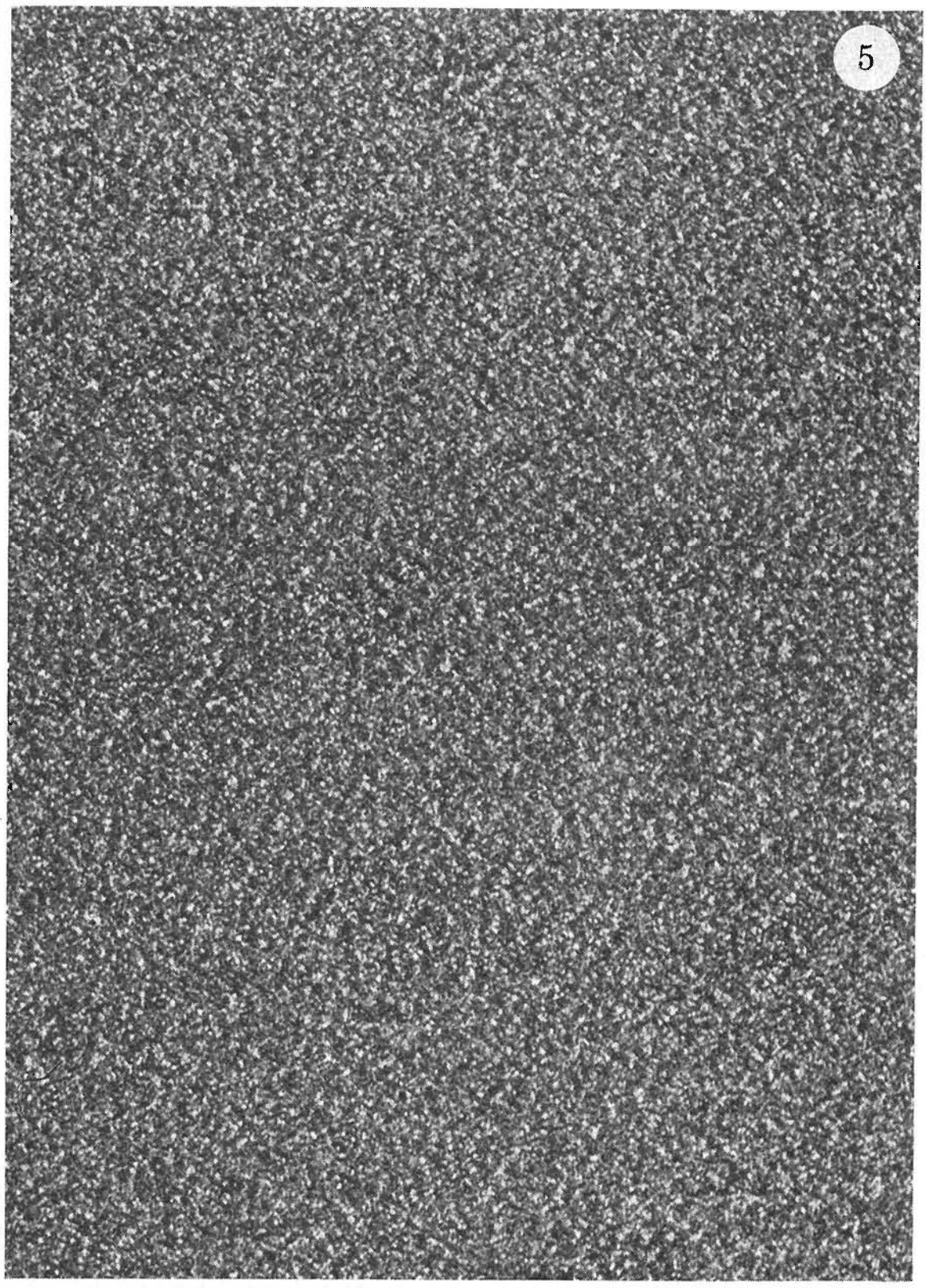


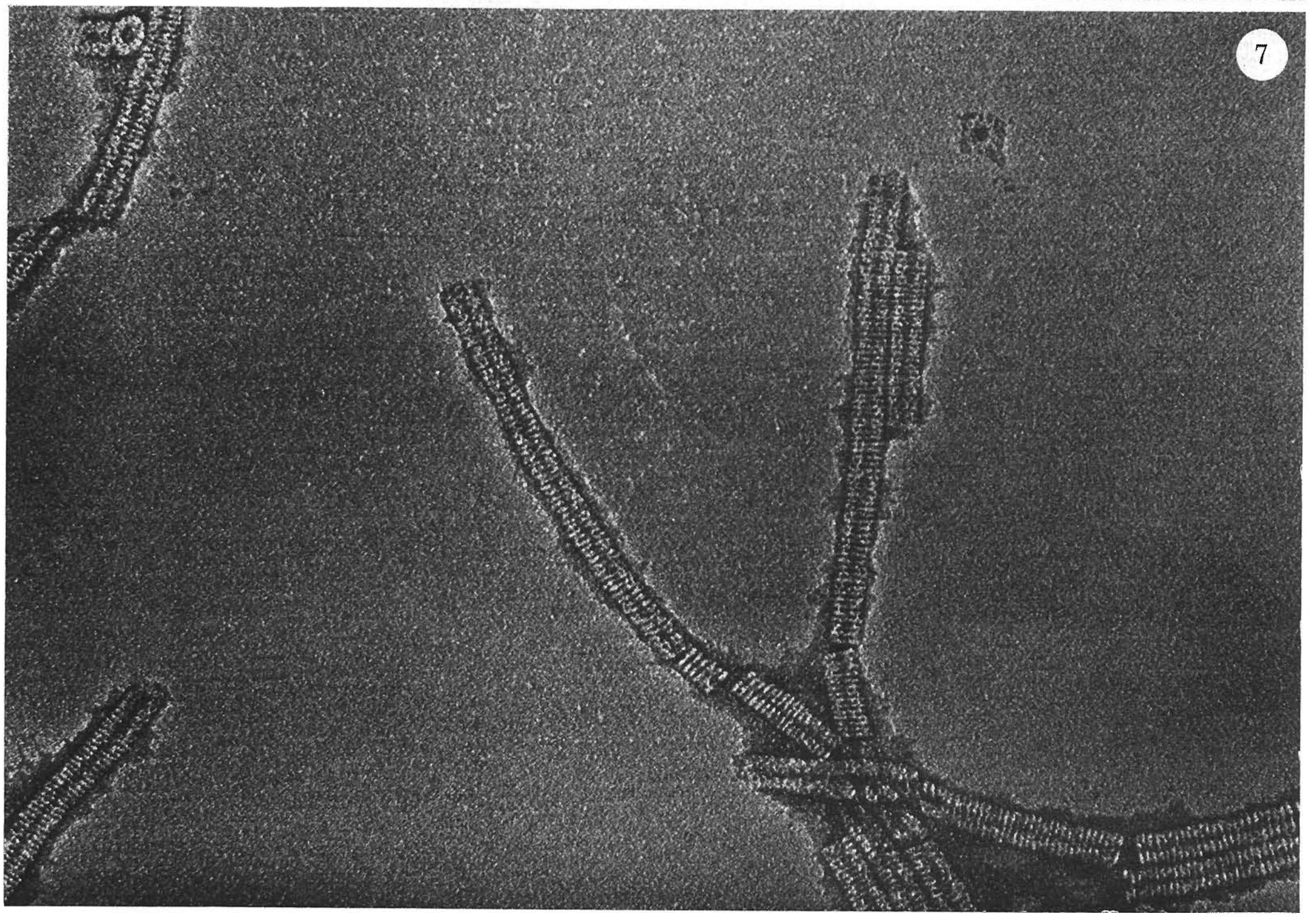


FIGURES 1 and 2*a*, *b*. For legends see facing page.

(*Facing p.* 332)







FIGURES 3, 5 and 7. For legends see facing page.

So far we have only considered the 'diffraction effect' associated with what are essentially points (or lines) in the object plane. Most real objects however have comparatively uniform densities spread over sizeable areas. These objects, on interacting with any particular incident wave, give rise to a varying phase and amplitude distribution in the diffraction plane. In being illuminated from the many separate points in the source, they therefore produce a large number of separate, but equivalent, laterally displaced diffraction patterns.

This partial coherence effect causes each diffraction maximum to be spread uniformly over a disk whose diameter, determined by the extent of the source, is somewhat larger than that of the thread; see figure 6. The thread thus behaves

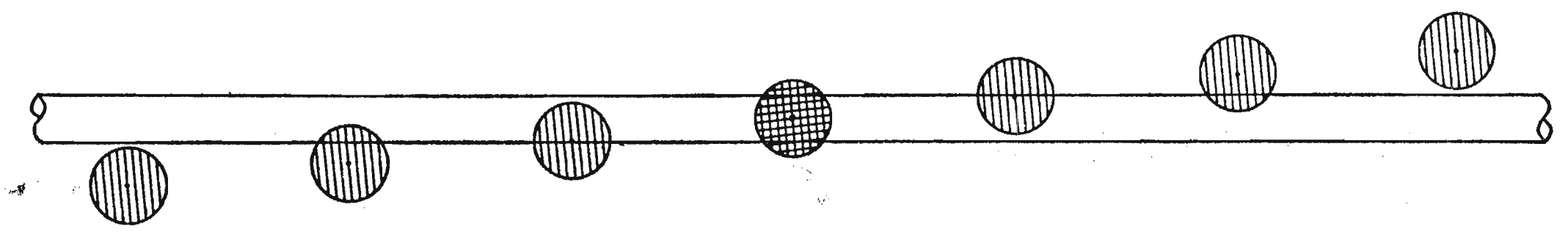


FIGURE 6. Showing how the electron diffraction maxima interact with the thread during normal operation of the phase plate. For the purpose of illustration, a grating is assumed to exist in the object plane which, if the illumination were coherent, would produce a row of bright spots in the diffraction plane making a small angle to the thread.

With the actual conditions of partially coherent illumination, these spots become smeared out into disks, as shown. The disk diameter is greater than that of the thread; moreoever, since critical illumination is used (see §2) this diameter is *ca.* $2f\theta_c$ (where θ_c is the semi-angle subtended at the object plane by the second condenser aperture) and the intensity over each disk is more or less uniform.

Only the part of each beam which overlaps the thread is, in effect, absorbed by it. Thus both the zero and first order maxima due to the grating contribute to the image even though they are centred in the direct line of the thread. The second order maxima in being centred outside, but within a disk radius of the thread's circumference, contribute to the image with less than their full weight.

DESCRIPTION OF PLATE 24

FIGURE 3. Diffraction fringes produced at the edge of a thin carbon film which has been allowed to curl up into the incident electron beam so as to effectively form lines of very strongly scattering material. The fringes were densitometered along the line indicated, and it was found that their spacings closely matched those which would be formed by the Fourier transform of a strip of width equal to the diameter of the thread used for the phase plate (see figure 4). (Magn. × 265 000.)

This, and the following two micrographs, are so oriented with respect to the page that the thread makes a vertical line.

FIGURE 5. Manifestation of the diffraction effect in the image of a carbon film. The micrograph has a slightly streaky appearance, each point in it being associated with a long, but weak, perturbation in the direction perpendicular to the thread; i.e. in the horizontal direction. (Magn. × 565 000.)

FIGURE 7. The diffraction effect associated with an object having comparatively uniform densities spread over sizeable areas. The dark regions of negative stain surrounding the test specimens are highlighted by a bright fringe which is about 8 nm wide (cf. figure 8*a*). The effect is not apparent in the direction of the thread. (Magn. × 250 000.)

as if it absorbs some of each diffraction maximum with which it interacts, the amount depending on the source size and the exact position of the maximum in the diffraction plane. The Fourier transform of an absorbing function of this nature is similar to that of the corresponding strip function, which would apply for the coherent case (where there is no spread in the diffraction maxima), but has a narrower central maximum and weaker subsidiary peaks. Its convolution with the original object would therefore lead to a highlighting effect which is similar but less extensive; a calculated profile for a typical case is shown in figure 8*a*.

Figure 7, plate 24, is a good example of the 'diffraction effect' associated with the type of object we have just considered. It has produced a strong highlighting of the edges of the dark film of stain that surrounds the test specimens.

3.2. *Lateral positioning errors*

In the preceding section it was tacitly assumed that the phase plate was positioned exactly centrally and in the Fraunhofer diffraction plane. In practice this ideal is by no means easy to achieve, and it is therefore of interest to examine the effects of small displacements of the phase plate both within this plane and along the optic axis.

The major effect of a small lateral displacement is to disturb the exact balance between the amplitudes of the corresponding (Friedel related) diffraction maxima on either side of the thread. To show how this imbalance affects the image we will consider the image formed by a simple one dimensional phase grating, firstly assuming the phase plate is positioned symmetrically in the diffraction plane. To take account of the *ca.* $\frac{1}{2}\pi$ phase change upon scattering (see for example, Haine & Cosslett 1961) we will represent the Friedel related maxima by complex numbers whose real parts, a_n, are of opposite sign, but whose imaginary parts, b_n, are equal, i.e. by $a_n+\mathrm{i}b_n$, $-a_n+\mathrm{i}b_n$. In the absence of aberrations, the amplitude distribution in the image plane, as a function of the distance, γ (the unit of which is so chosen as to make the period of the grating extend from $-\pi$ to $+\pi$) from a certain fixed origin, then becomes

$$\Psi(\gamma) = \Psi_0 + 2\mathrm{i}\sum_{1}^{\infty}(w_n a_n \sin n\gamma + w_n b_n \cos n\gamma), \tag{1}$$

where Ψ_0 is the amplitude of the unscattered wave and w_n is a weighting term which accounts for the weakening or elimination of certain maxima by the components of the phase plate (and through the small angular variation in the atomic scattering amplitude – see §4.3).

In order to represent the actual situation in the electron microscope, the effect of the $\frac{1}{2}\pi$ phase change introduced into a large part of the scattered wave by the combined effect of the electric field of the phase plate and of spherical aberration must also be included in the expression for $\Psi(\gamma)$. This can be done by multiplying the quantities $a_n+\mathrm{i}b_n$, $-a_n+\mathrm{i}b_n$ of those maxima which are affected, by $-\mathrm{i}$. A similar expression is thereby deduced, except for the imaginary part of $\Psi(\gamma)$

which is associated with these maxima becoming real. On squaring, neglecting the small second order quantities, to get the intensities, only this real part and Ψ_0^2 are retained. It therefore, together with the uniform background, Ψ_0^2, specifies the intensity distribution which will be observed at the image plane.

If now the phase plate is displaced laterally slightly from the symmetrical position, the real parts (a_n) of some of the lowest order Friedel related maxima will no longer cancel. Such a state of imbalance can be accounted for in equation (1) by altering the values of some of the w_n very slightly and introducing an additional real term into the right hand side. This real term will be of the form $(c_n \cos n\gamma - d_n \sin n\gamma)$ summed over the region of imbalance, with coefficients c_n and d_n representing the differences in amplitudes between each related pair of maxima; it is a schlieren term, which, if appearing in the intensity expression, provides a type of image resembling the original object viewed under oblique illumination (Zernike 1942). In practice, because it is impossible for the electric field to produce a useful phase shift amongst the very low resolution Fourier components, where c_n and d_n are important, it normally will appear in the intensity expression.

The result of a simple one dimensional calculation, shown in figure 8, illustrates the importance of this term in relation to the term specifying the 'true' image. It is apparent from this figure that only very small lateral displacements of the phase plate (*ca.* 100 nm) from the symmetrical position should be sufficient to make the schlieren term detectable in the image.

The high sensitivity to errors in lateral positioning is borne out in practice and in some micrographs, e.g. figure 13*d*, plate 26, the schlieren effect becomes fairly conspicuous. Even so, unless the errors involved are unduly large, this image defect is not considered to be a serious one, because it is generated by only low resolution Fourier components and, as can be seen by comparing figure 8*a* and 8*b*, there is not much accompanying alteration to the 'true' image.

3.3. *Incorrect height adjustment*

Some preliminary tests indicated that it was possible to adjust the height of the phase plate to lie within about 0.1 mm of the diffraction plane without too much difficulty. This distance is rather large in comparison with, say, the diameter of the phase plate aperture, and one might therefore have expected to need considerably finer adjustments to produce a good phase contrast image. Practical experience, on the contrary, has shown that such large errors in longitudinal alinement can be tolerated without any detectable deterioration of the image.

This perhaps rather surprising finding can be understood by making use of the Huygens–Fresnel principle and carrying out some geometrical constructions. We consider each point in the diffraction pattern to act as a source for spherical wavelets, and each such original source, S, to be in effect shifted a certain amount by the electric field of the phase plate to a new virtual position, S′. The consequence of S moving to S′ is then deduced, in the knowledge that the image is

just the interference pattern formed by the recombination of all such wavelets reaching it.

In order to locate the virtual position, S′, by simple geometrical constructions it has to be assumed that all electrons emerging from a given point in the diffraction plane pass through approximately the same field and therefore suffer

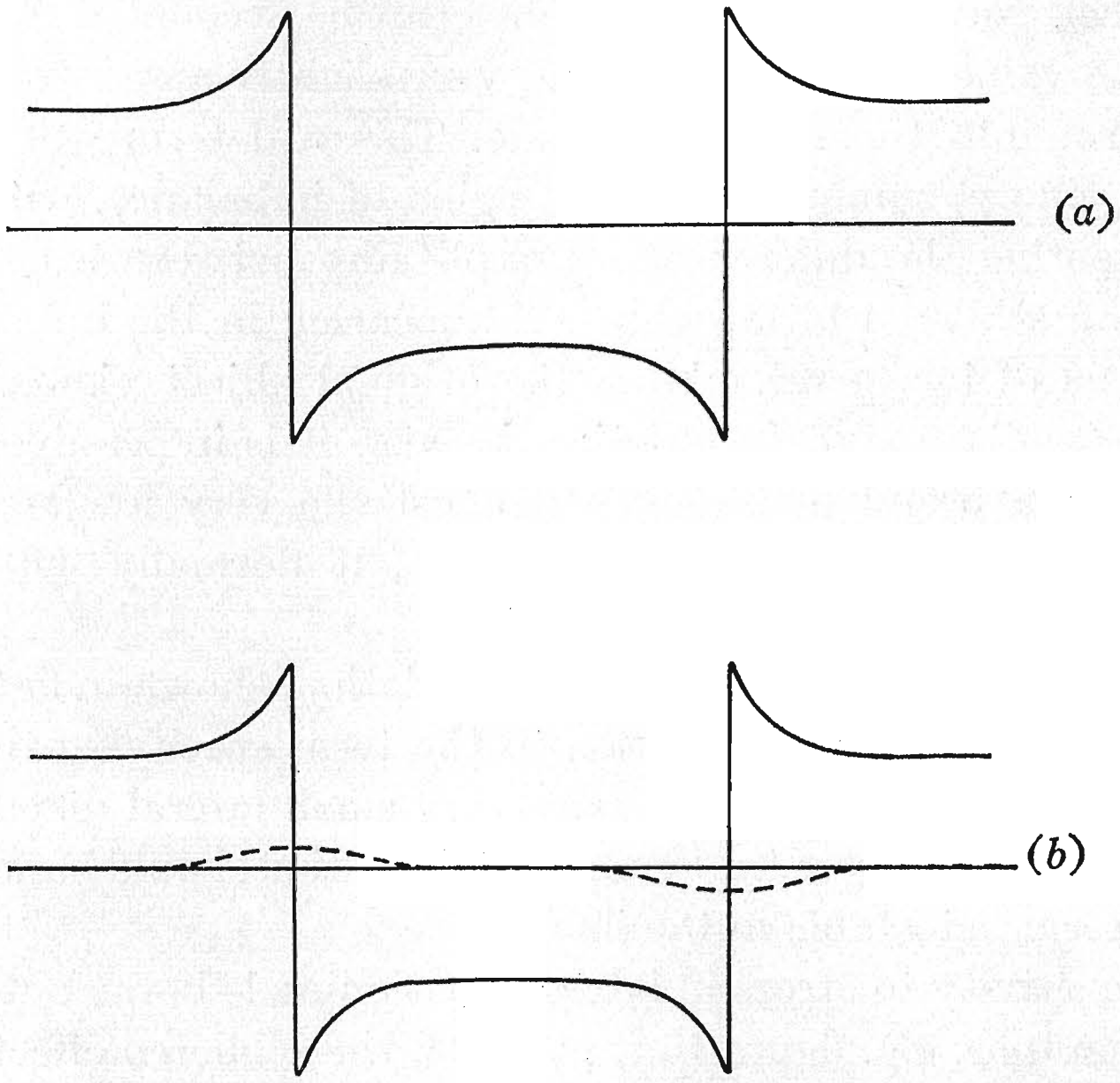


FIGURE 8. Amplitude profiles showing the effect of a small lateral displacement of the thread from the central position in the Fraunhofer diffraction plane. It is assumed that the thread is perpendicular to the plane of the paper, that its diameter is 0.4 μm, and that it alters the amplitudes in the diffraction plane as if each Fourier component were spread uniformly over a circular patch twice its diameter. The image formed by allowing all Fourier components to be delivered at the weight determined by the contrast transfer characteristics of the phase plate (see §4.3) is a 20 nm wide rectangular pulse.

(*a*) The 'true' image formed on accounting for the thread being in the central position;

(*b*) the 'true' and schlieren (broken line) components formed when the thread is displaced from the central position by 100 nm; in practice, the schlieren component adds directly to the 'true' image, generating a bright fringe on one side and a dark fringe on the other.

the same deflexion. But this assumption is justified for the magnitudes of vertical displacement we are considering (*ca.* 0.1 mm), since the angle subtended by the normal high magnification image is very small (*ca.* 10^{-4} rad) and those electrons striking the photographic emulsion which originate from the same point in the diffraction plane are therefore separated by *ca.* 10 nm, at the most, in the plane of the phase plate.

The construction method is given in figure 9*a*. For small deflexions, S′ lies on the line through S that is perpendicular to the line drawn between S and the

centre of the phase plate. Whether it is closer or farther from the optic axis than S depends on whether the phase plate is above or below the diffraction plane. Note also that with deflexions towards the optic axis, S′ always moves towards the image plane; this corresponds to a decrease in the optical path of the scattered wave.

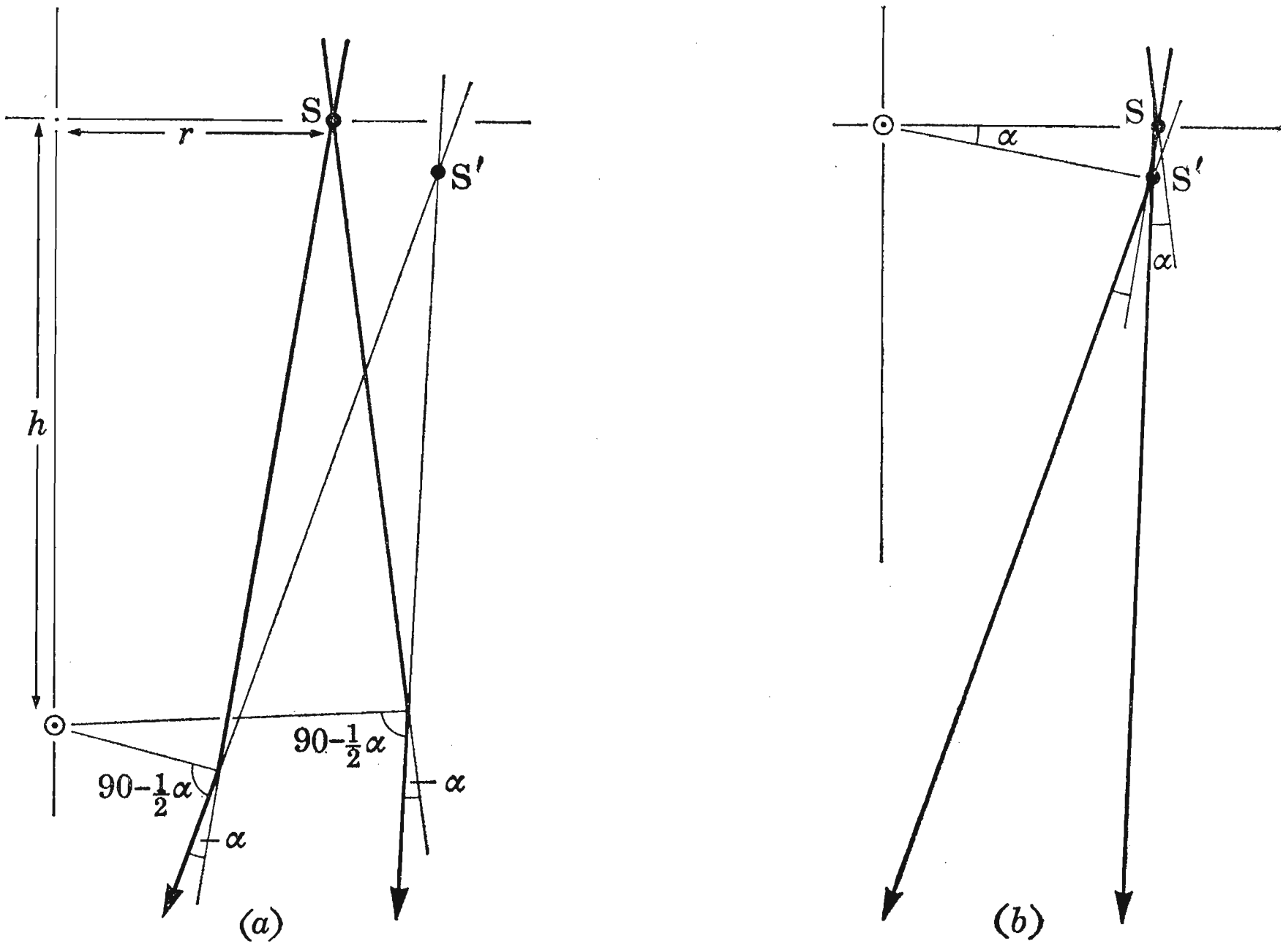


FIGURE 9. A simple construction for showing the effect of small displacements of the phase plate along the optic axis. S is a point in the Fraunhofer diffraction plane at a distance, r, from the optic axis, and it is transposed to a new vertical position, S′, below S by an inward deflection, α, about the thread axis (which is normal to the plane of the paper). With small values of α the vertical displacement is always *ca.* $r\alpha$, producing a phase shift, $2\pi\lambda^{-1}(r\alpha)$, which does not depend on the height, h, of the phase plate above or below the diffraction plane.

(*a*) shows how the position of S′ is determined for the general case: two arbitrary straight lines are drawn through S, bent by an amount, α, about the thread axis, and projected backwards; S′ is their point of intersection; (*b*) shows the movement S → S′ for the same deflexion, when the phase plate lies exactly in the diffraction plane.

It is clear from figure 9 that the height, h, of the phase plate above or below the diffraction plane is unimportant in determining the magnitudes of the phase shifts generated; however, only when it lies exactly in this plane does the phase plate effect a movement S → S′ corresponding to a phase shift only (figure 9*b*). In all other cases there is a displacement component, δr, within the diffraction plane. This displacement is in the same or opposite direction to that of the deflexion, and since $r = f\lambda/a$ (where a is the object spacing involved) it is manifested in the image as a distortion.

With the magnitudes of error in height adjustment likely to be encountered in

practice, $h/r \gg 1$. Under these conditions the maximum fractional distortion in object spacing can be shown to be related to the phase shift, χ, by

$$\frac{\delta a}{a} \simeq \frac{\lambda h}{2\pi r^2}\chi, \tag{2}$$

it therefore varies linearly with h; but more important, because $r \propto 1/a$ it should become especially severe among large object spacings. This dependence on object spacing can be observed in images formed by purposely making h and χ very large; however, for normal high resolution images, in which the maximum values for h and χ are respectively 0.1 mm and π, equation (2) indicates that these distortion effects are unlikely to be significant. For example, object spacings would need to be greater than 40 nm to show as much distortion as 1 %.

3.4. *Charge and focusing errors*

Incorrect adjustment of charge on the thread and inaccurate focusing lead to essentially the same type of artefact. This can be most easily recognized in the Fourier transform of the micrograph since it is associated with a contrast transfer function (see §4.3) which begins to oscillate at lower spatial frequencies than it should. Oscillations of the contrast transfer function mean that different spatial frequencies appear in opposing contrast at the image plane. Micrographs can therefore only be considered to be free from charge and focusing artefacts if the first reversal in contrast appears at a resolution which is better than that of the finest detail which is *characteristic* of the specimen.

4. Contrast properties

The contrast apparent in the electron micrographs can be divided into two types: amplitude contrast (also referred to as scattering or absorption contrast) and phase contrast. The theoretical background to the mechanisms involved and calculations dealing with some specific applications have been given in a number of papers (for a recent review see Hanszen 1971).

In what follows, we are concerned with the relative contributions made by each mechanism, taking account of the phase plate's property of being able to prevent a large proportion of the unscattered electrons from contributing to the image.

4.1. *Amplitude contrast*

Amplitude contrast arises as a result of a difference in the proportion of scattered and unscattered electrons that make up the image to the proportion leaving the object. Many of the inelastically scattered electrons which pass through the phase plate aperture are probably unable to contribute effectively to the high resolution image, but, for present purposes, it is sufficient to consider just those which are scattered outside it to be unavailable. Thus, if the phase plate reduces the intensity of the unscattered wave to a fraction, q_0, of its original intensity, and the intensity

of the scattered wave to a fraction, q_s, of its original intensity, this source of contrast must vanish when $q_0 = q_s$. Furthermore, dark amplitude contrast (i.e. thicker, more strongly scattering parts of the object appear darker) must obtain when $q_0 > q_s$, and bright amplitude contrast when $q_0 < q_s$. In normal bright field microscopy, with q_0 always equal to 1, only the dark type of contrast can be achieved; in dark field microscopy ($q_0 = 0$) only the bright type of contrast can be achieved.

With the phase plate, the value of q_0 can of course be varied according to the diameter of the beam of unscattered electrons illuminating the thread. It therefore enables both types of amplitude contrast to be achieved.

Assuming the normal logarithmic form for the contrast, the maximum amplitude contrast available to the image, in terms of q is

$$C_A = -\ln\{1 + p(q_s/q_0 - 1)\}, \tag{3}$$

where p is the fraction of electrons striking the object that are scattered by it. This expression shows dark contrast to be the least effective of the two types of amplitude contrast; its improvement, requiring a smaller figure for q_s, can only be accomplished by a reduction in the size of objective aperture – i.e. at the expense of resolution. With bright contrast, where q_0 is made smaller and q_s is allowed to remain large, there is no such disadvantage. Moreover, providing q_0 can be made small compared to q_s, bright contrast obviously provides opportunity for the greatest enhancement. In practice, insufficient illumination to facilitate accurate focusing and charge adjustment imposes a lower limit of 0.1–0.2 on this parameter.

The magnitude of q_s, unlike that of q_0, is fixed by the dimensions of the phase plate. For given phase plate dimensions it is determined by the distribution of total (elastic plus inelastic) scattered electrons in the diffraction plane. Unfortunately, although this distribution should be essentially independent of the thickness of the scattering material for the small mass thicknesses we are concerned with, it is strongly dependent, because of interference effects, on the arrangement of the constituent atoms.

This point is illustrated very clearly in figure 10, plate 25, where a direct comparison is made between the electron diffraction pattern of a thin carbon film and of a thin film of stain. The type of pattern exhibited by the carbon film resembles that which would be expected from an amorphous material, but the type of pattern exhibited by the stain film is typical of a material composed mainly, if not entirely, of small crystals.

The sampling effect seen in the stain pattern is due to Bragg scattering from the individual small crystals, and all the sampling peaks lie outside the phase plate aperture because of the small lattice spacings involved. From the observed half-peak breadths and by direct examination of the image of the film itself, it is estimated that the majority of the crystals range in diameter from 1 to 2 nm, though there may be a considerable number of even smaller ones.

Although the Bragg scattering contribution presents difficulties in the case of the stain, it is possible to derive an estimate of the magnitude of q_s for the carbon – and for the biological matter, for which its value must be essentially the same (Burge & Silvester 1960) – by considering this material to be completely amorphous, so that its electron scattering behaviour conforms strictly with the theoretical scattering distributions derived for single atoms. We then have $q_s \approx (S_T - S_t)/S_T$, where S_T and S_t are, respectively, the theoretical mass scattering cross-sections for scattering at all angles and for scattering outside the phase plate aperture (the edge of which corresponds to a scattering angle, θ, of 9×10^{-3} rad). Taking S_T as 24×10^4 cm^2 g^{-1} (Burge & Smith 1962*a*, figure 1) and S_t as 6×10^4 cm^2 g^{-1} (Burge & Smith 1962*b*) we estimate q_s for carbonaceous materials to be 0.75. A value of 4×10^4 cm^2 g^{-1} for S_t, obtained by interpolating some experimental data (Hall 1966, p. 212, table 8.2), suggests that this figure may be a little too low, but at the same time gives some confidence in the theoretical derivations.

The above method obviously cannot be used to obtain an estimate of q_s for the stain. Nevertheless, we know from figure 10 that it is certainly much smaller than the value for carbon, whilst the experiments of §4.3, in which the stain remains dark over a wide range of q_0 values, indicate that it is less than 0.2.

Now, as explained earlier, we cannot readily lower q_0 much below 0.2 without running into illumination difficulties. If then we assume this lower limit for q_0 and the foregoing estimated values for q_s, it is evident that q_s for the stain is too low for the phase plate to be effective in improving its amplitude contrast (lowering q_0 to 0.2 will only make it weaker). On the other hand, by using equation (3), it can easily be shown that q_s for the carbonaceous material is sufficiently high that the bright amplitude contrast available is a substantial improvement over the maximum dark amplitude contrast attainable by bright field methods.

4.2. *Phase contrast*

Phase contrast is created by introducing a path difference between the scattered and unscattered waves. For optimum interference between these waves the corresponding phase shift is $\frac{1}{2}\pi$. The maximum phase contrast attainable also depends on the fraction of the original intensity of the unscattered wave remaining available to the image (q_0); for the *weakly* scattering objects we are considering this improves as $q_0^{-\frac{1}{2}}$.

Description of plate 25

Figure 10. Electron diffraction patterns from a thin film of carbon (at the top) and a thin film of uranyl formate stain (at the bottom), obtained under identical illumination conditions. To make the comparison as realistic as possible a mask was used so that a half of each diffraction pattern was exposed onto the same photographic plate. The central dark line is the region of overlap for the two mask positions; its width was made considerably greater than the diameter of the intense central order beam in order to avoid halation effects. The broken circle drawn over the patterns with a radius corresponding to 1/0.39 nm^{-1} represents the edge of the phase plate aperture.

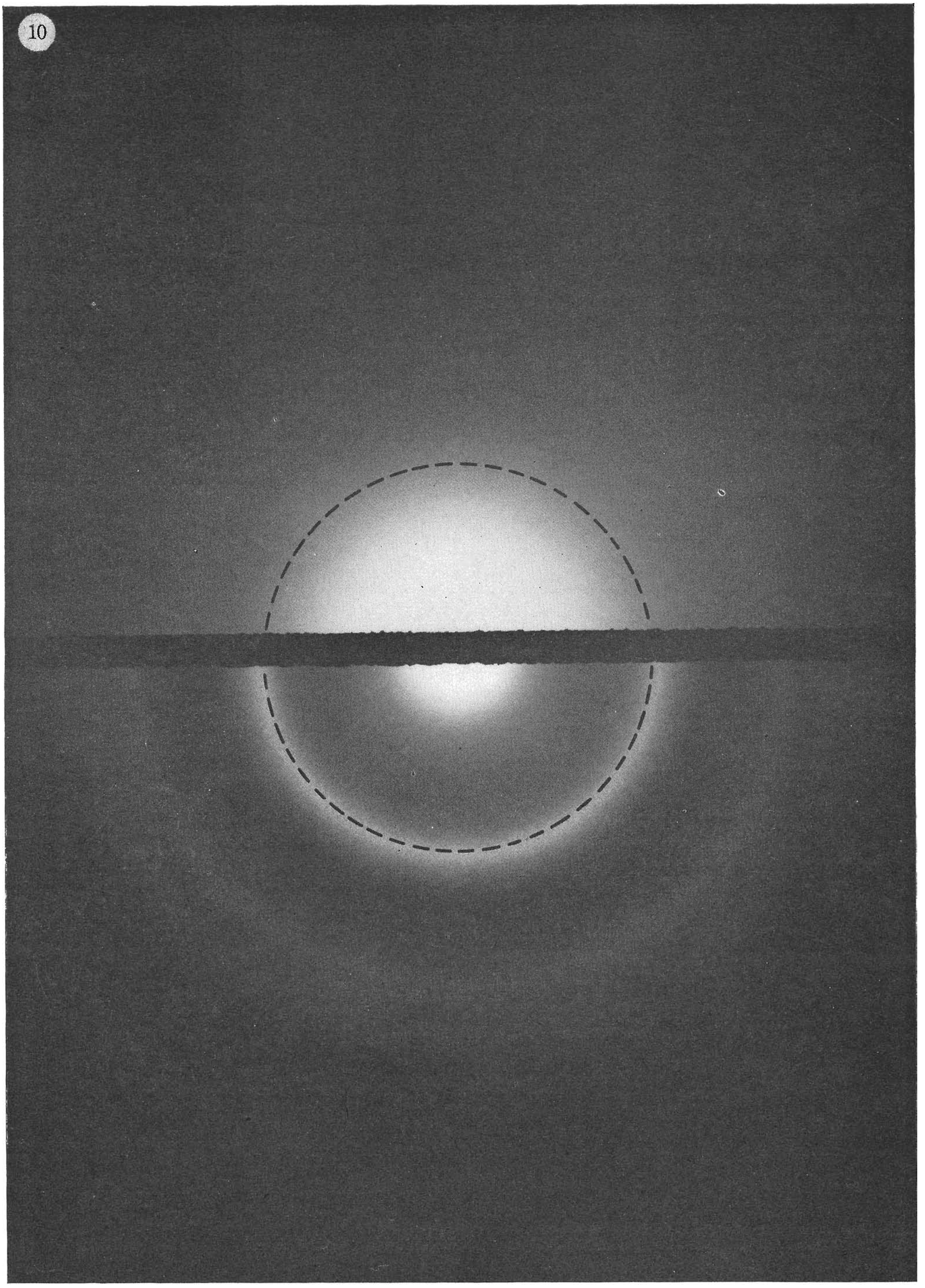


FIGURE 10. For legend see facing page.

(*Facing p.* 340)

The phase shifting properties of our type of phase plate have been discussed in some detail previously (Unwin 1971): the aperture diameter and the distribution and magnitude of charge on the thread are the important parameters. When we take into account the phase shift due to spherical aberration, the total phase shift under optimum conditions should vary with the reciprocal coordinate in the diffraction plane, θ/λ, in the manner shown in figure 11. This curve is approximately the same for all azimuths, except in the region very close to the origin,

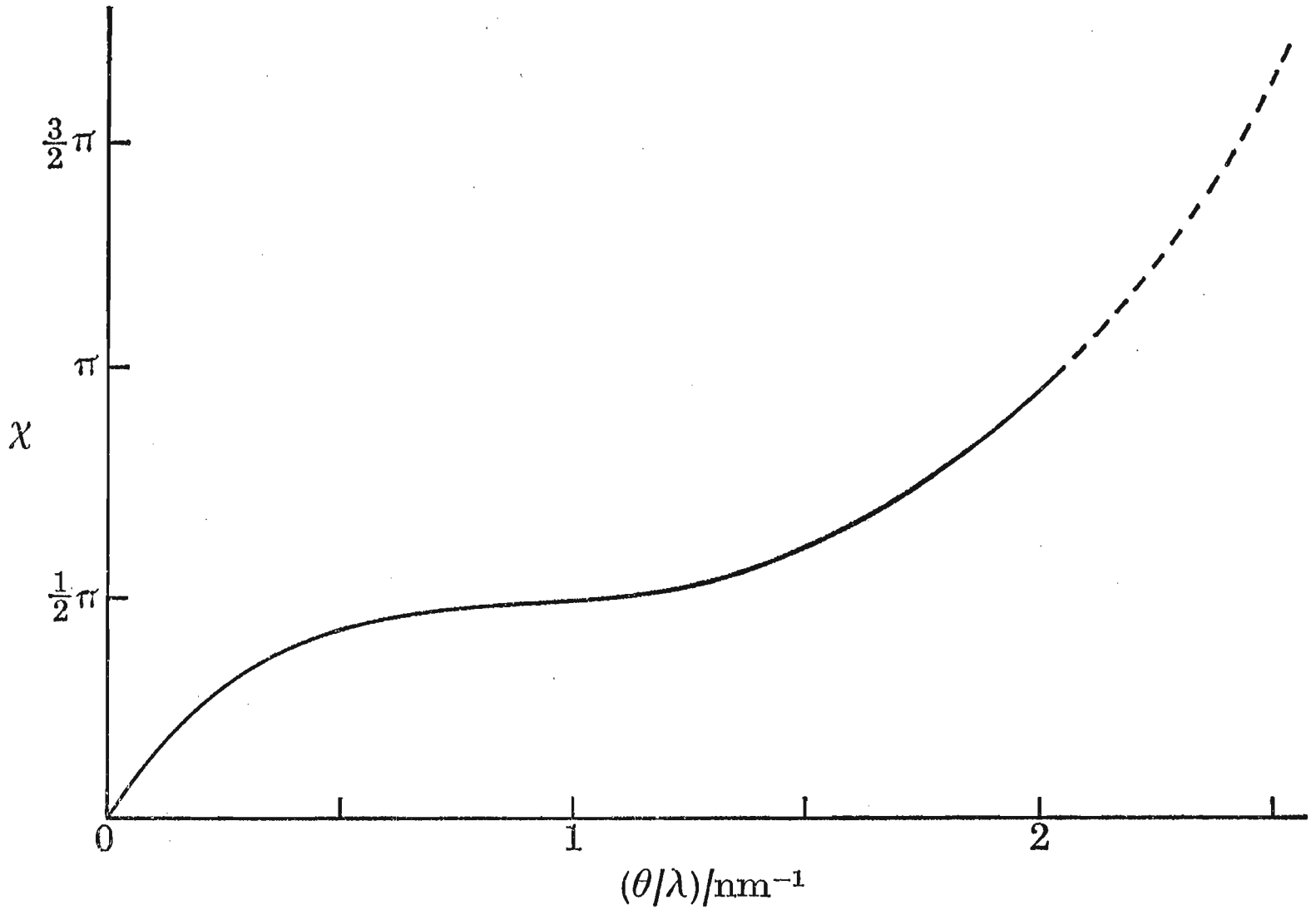


FIGURE 11. The variation of the phase shift, χ, due to the combined effects of the phase plate's electric field and of spherical aberration, with the reciprocal coordinate, θ/λ. The shape of the curve up to $\theta/\lambda \approx 0.15$ has an azimuthal dependence; as drawn, it refers to the phase shift variation in the direction perpendicular to the thread. The positive phase shift indicates a shortening of the optical path.

where the electric field is no longer nearly rotationally symmetric. The point at which the $\frac{1}{2}\pi$ plateau is reached can be varied to some extent, depending on the length of thread illuminated by the beam of unscattered electrons, but the severity of the phase shift due to spherical aberration imposes an upper limit to its extension to higher resolutions.

Now the electric field generated by the phase plate deflects the scattered electrons towards the optic axis; it therefore shortens their optical path (§3.3). The $\frac{1}{2}\pi$ plateau in figure 11 thus corresponds to the condition for bright phase contrast. Bright phase contrast (in which thicker, more strongly scattering parts of the object appear brighter) is the opposite to the type of phase contrast produced in bright field by defocusing, and leads naturally to further enhancement of object detail displayed in bright amplitude contrast. Taking account of the findings in §4.1, this means that our method of observation is more suited to enhancing detail in the carbonaceous material than in the stain, and works best for this

material when $q_0 < 0.75$ (i.e. when more than one-quarter of the unscattered electrons have been stopped off).

Since preventing unscattered electrons from reaching the image also improves the phase contrast, our method of observation clearly provides opportunity for greater enhancement of carbonaceous material than does the bright field method in terms of both contrast mechanisms.

4.3. *The total image*

It is well known from contrast transfer theory (see, for example, Hanszen & Morgenstern 1965; Hanszen 1971) that, to a first approximation, the image of a weakly scattering object is just the sum of the phase contrast image and the amplitude contrast image considered separately. Furthermore, the two images are

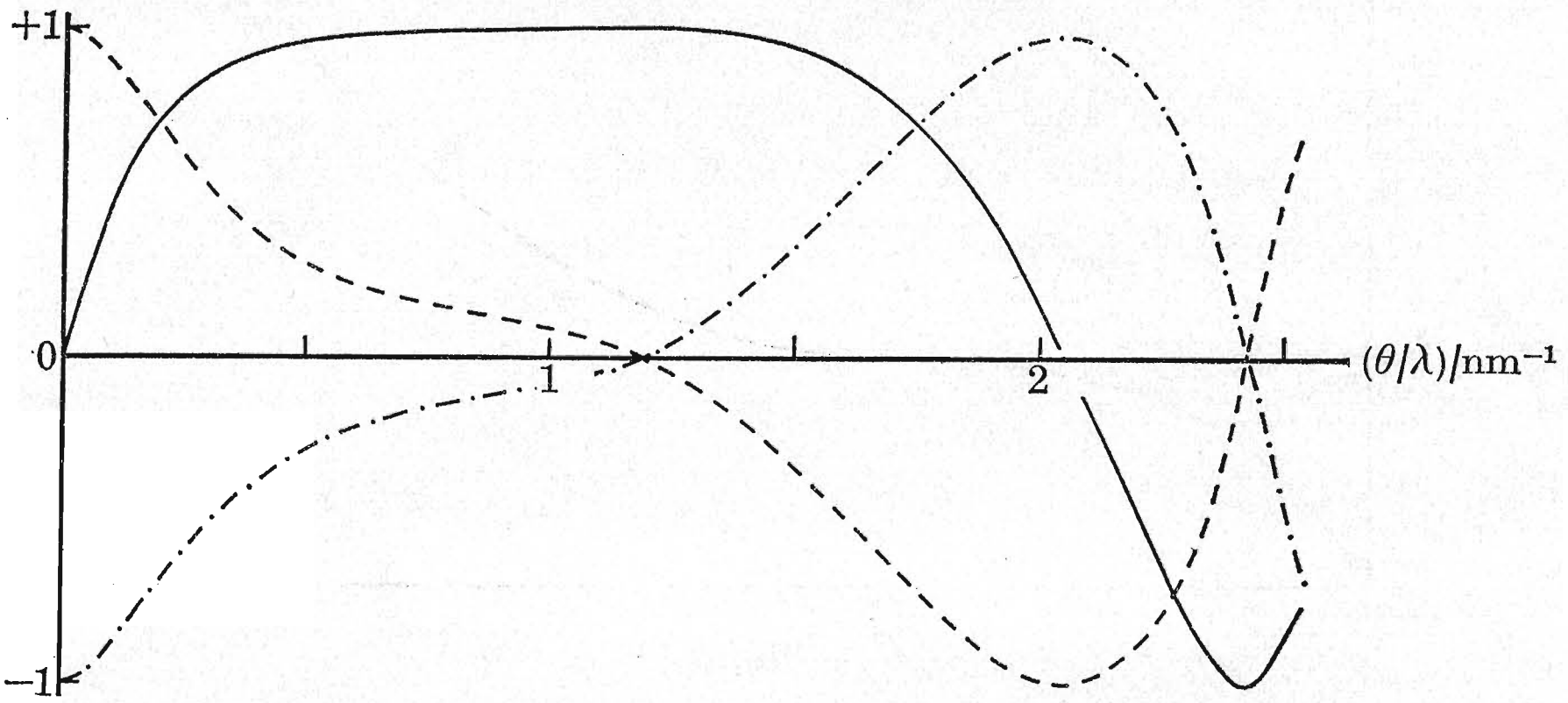


FIGURE 12. Contrast transfer functions for the phase plate: ——, $\sin\chi$, the phase contrast transfer function; – – –, $\cos\chi$, the amplitude contrast transfer function for bright contrast, i.e. for carbonaceous material when $q_0 < 0.75$; ·–·–·–, $-\cos\chi$, the amplitude contrast transfer function for dark contrast, i.e. for the stain for all values of q_0 and for carbonaceous material when $q_0 > 0.75$. As in figure 11, the exact shape of the curves up to $\theta/\lambda \simeq 0.15$, depends on the azimuth.

essentially the Fourier transform of the amplitude distribution due to those elastically scattered electrons† at the diffraction plane which remain available to the image modulated, respectively, by the sine and the cosine of the phase shift, χ. These conditions of linear transfer apply only if the amplitudes of the various diffracted orders remain weak in comparison with the amplitude of the central order, i.e. providing q_0 is not made too small. Curves of $\sin\chi$ and $\cos\chi$ plotted against the reciprocal coordinate, θ/λ (contrast transfer functions), then denote the efficiency with which the various Fourier components are transferred to the image under a given set of optical conditions.

† It is assumed that the image formed by the inelastically scattered electrons is sufficiently blurred as a result of delocalization of the scattering events and of chromatic aberration (Crick & Misell 1971; Misell & Crick 1971) that it makes a negligible contribution to the high resolution contrast. The inelastically scattered electrons will however contribute to the general background level on which the high resolution detail is formed.

Figure 12 shows the curves appropriate to our observation conditions. These were calculated directly from figure 11 and therefore assume that the object is exactly in focus. It was not considered necessary to include the effects of partial coherence in these calculations, since with the small angles of illumination employed the transfer function for the coherent case and the (normalized) transfer function, corrected for the effects of partial coherence, are practically identical over the region the curves have been drawn (Hanszen & Trepte 1971).

The sin χ curve indicates that very nearly the true phase contrast image should be obtained with all object spacings lying between about 0.6 and 5 nm. The $\pm\cos\chi$ curves show, on the other hand, that there can be important additional contributions from the amplitude contrast among parts of the image deriving from the very low and the very high spatial frequencies. Actually, because the amplitudes are certain to be much stronger at low angles, it is among the parts of the image deriving from the very low spatial frequencies that the effect of the amplitude contrast will be most noticeable. Indeed the amplitude contrast must determine the relative mean intensities of all gross features in the image.

For a complete description of the total contrast, not only is it necessary to know the efficiencies of transfer of the two types of contrast, but also the ratio of the maximum amounts available to the image. As before, by assuming it to be completely amorphous (and so neglecting the modifications to the scattering behaviour of single atoms brought about by the type of binding involved and by interference effects) it is possible to get a useful estimate of this ratio for the carbonaceous material. In the case where there is no reduction in the intensity of the unscattered wave (i.e. $q_0 = 1$), the ratio of maximum amplitude to maximum phase contrast is given approximately by (Erickson 1972):

$$Q = S_t A/2N_0\lambda f_e, \tag{4}$$

where N_0 is the Avagardro constant and A and f_e are, respectively, the atomic mass and the atomic scattering amplitude for electrons of the material; f_e varies slowly with scattering angle over the angular range of interest (to $\theta = 9\times10^{-3}$ rad) and taking the mean figure for this range – using the values calculated by Haase from Thomas–Fermi–Dirac potentials for free atoms (Haase 1970) – of 0.28 nm and the previous theoretical value for S_t, we estimate Q for carbon, when $q_0 = 1$, to be 0.06.

Because of the approximations involved, this figure for Q cannot be taken too literally; it does show, however, that the contrast available from the amplitude of the *elastic* wave scattered by the carbon inside the phase plate aperture is large in comparison with the contrast available from the intensity of the *total* wave scattered outside it.

It remains to calculate how Q, and also the total contrast, varies with q_0. This can be done with the aid of equation (3) by putting $p = 1-\exp(-S_T\rho t)$ (ρt is the mass thickness = density × thickness) and using the above estimated values for q_s and Q. Table 1 shows the results of some calculations based on a mass thickness

of 5×10^{-7} g/cm^2; this corresponds to about 2.5 nm thickness for the relatively dense carbon support film, but probably more like 5 nm thickness for the biological material.

Note in this table the reversal in the type of amplitude contrast either side of the zero point in Q (at $q_0 = 0.75$) and the greater magnitude of amplitude contrast, relative to the bright field figure ($q_0 = 1$), when q_0 is made lower than 0.6. The type of phase contrast is of course always the same (bright), improving steadily as q_0 is lowered towards 0.2. It does not improve as rapidly as the amplitude contrast, however, so that Q is greatest when very large proportions of unscattered electrons are stopped off.

TABLE 1. THE CALCULATED VARIATION IN THE MAXIMUM AMPLITUDE CONTRAST CONTRIBUTION, C_A, THE MAXIMUM PHASE CONTRAST CONTRIBUTION, C_p, AND THEIR RATIO, Q, WITH THE ABSORPTION PARAMETER, q_0, FOR A PIECE OF AMORPHOUS CARBONACEOUS MATERIAL HAVING A MASS THICKNESS OF 5×10^{-7} g/cm^2

q_0	C_A†	C_p	Q
1.0	−0.03	0.50	0.06
0.8	−0.01	0.56	0.02
0.6	+0.03	0.65	0.05
0.4	+0.10	0.79	0.13
0.2	+0.27	1.12	0.24

† The figures in this column have been multiplied by −1 so that their signs conform with the usual convention for electron microscopy, in which *dark* contrast is associated with the *negative* parts of the contrast transfer function. This convention is the opposite of that introduced for light microscopy by Zernike (1942), where dark and positive are considered to be equivalent.

The set of micrographs, reproduced in figure 13, plate 26, shows that the observed changes in contrast as q_0 is decreased are consistent, at least qualitatively, with the calculations in table 1. Each micrograph shows a hole and a relatively thick carbon film as well as the thin carbon film on which the specimens are supported.

The contrast between the background intensity within the holes and the carbon support film is weak in all cases because this film is very thin. To get a good indication of the nature and magnitude of the amplitude contrast contribution in each micrograph we therefore examine the relative mean intensities of the two

DESCRIPTION OF PLATE 26

FIGURE 13. A set of micrographs showing the changes in contrast brought about by a variation in the absorption parameter, q_0. Each micrograph shows some of the negatively stained test specimens, the thin carbon support film (T), a small hole (H) and another, much thicker, carbon film (S) on the right. The approximate figures for q_0 are (*a*) 0.85, (*b*) 0.7, (*c*) 0.4, (*d*) 0.2. The rather conspicuous bright band on the lower side of the 'raft' in (*d*) is a schlieren effect brought about by a small error in the lateral positioning of the phase plate (see §3.2). (Magn. ×325000.)

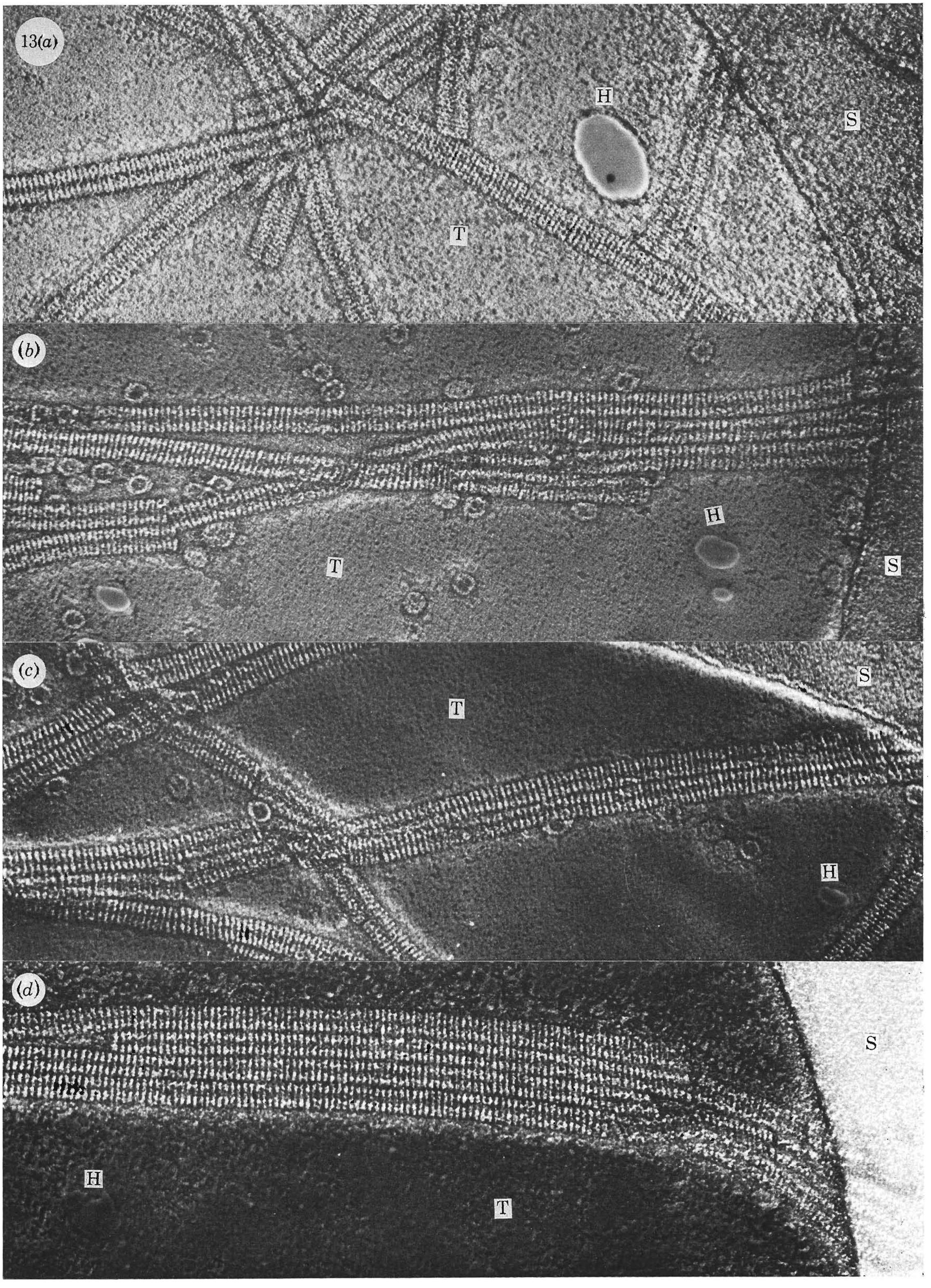


FIGURE 13. For legend see facing page.

(*Facing p.* 344)

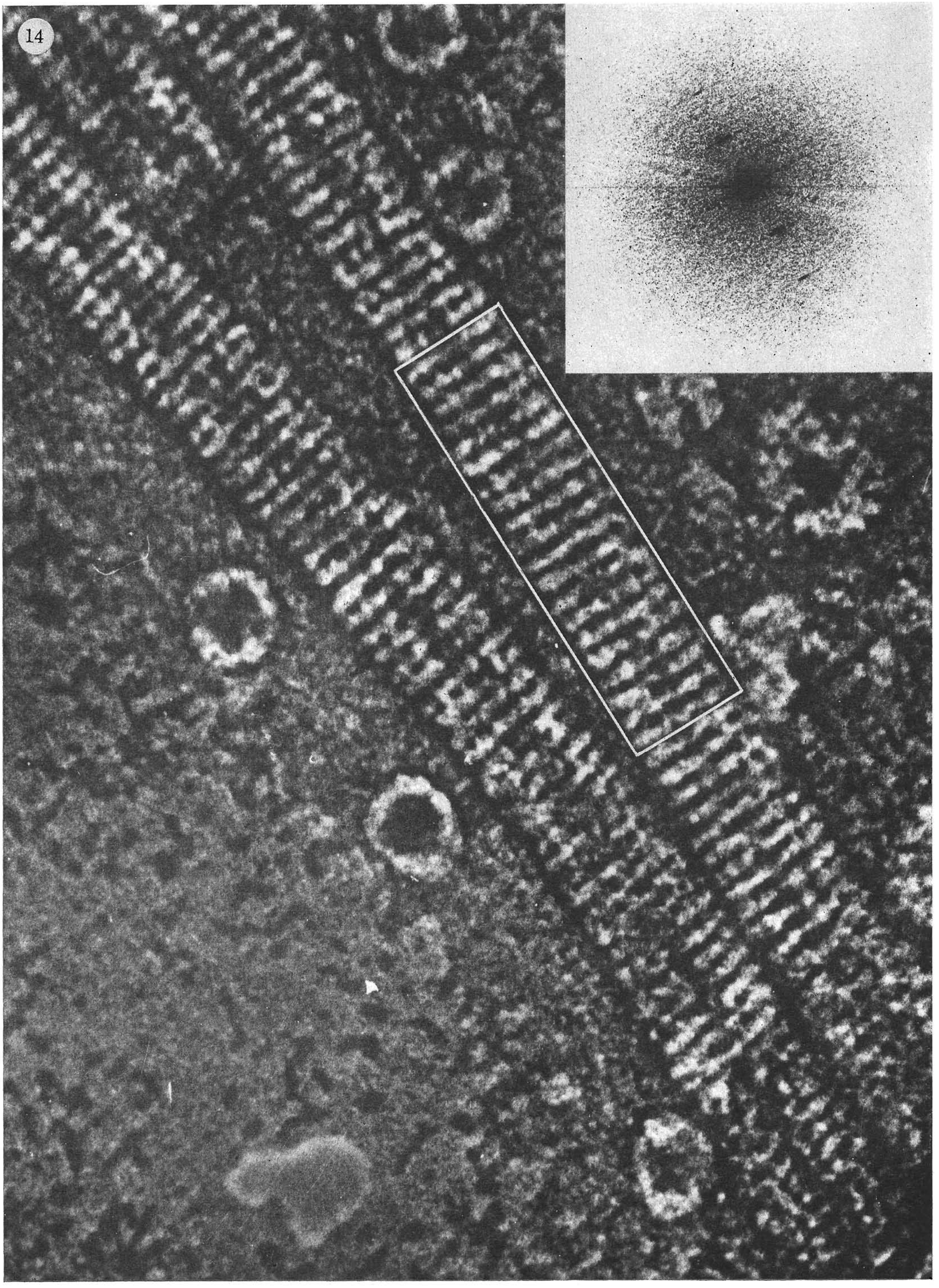


FIGURE 14. For legend see facing page.

carbon films. It is seen to change in the predicted fashion; thus in (*a*) where $q_0 > q_s$ it is dark but weak, in (*b*) where q_0 and q_s are almost the same it is almost non-existent, in (*c*) where $q_0 < q_s$ it is definitely bright, and in (*d*) where $q_0 \ll q_s$ it is bright and strong.

To get an indication of the nature and magnitude of the phase contrast contribution we can look to the specimens themselves. They provide an appropriate guide because, as will be seen later, nearly all their diffraction maxima fall into the spatial frequency region where $\sin \chi \simeq 1$. Again, qualitatively, there is agreement with table 1: the repeating biological units in all cases appear brighter than the background level, even in (*a*) where the amplitude contrast is dark, and this difference increases as q_0 gets smaller.

The fact that the stain appears dark throughout the sequence of micrographs indicates that its Q value is very large, i.e. it scatters only a very small number of electrons, *elastically*, inside the aperture in comparison with the total number of electrons it scatters outside. This proportion is not sufficient to prevent the dark amplitude contrast, which is strong because q_s is also low, from having the dominating effect. These properties of the stain are expected on the basis of its crystallinity (see later discussion, §6).

5. Analysis of electron micrographs

The rapid computing techniques developed by DeRosier & Klug (1968) and further extended by DeRosier & Moore (1970) for calculating Fourier transforms of selected areas in electron micrographs and reconstructing images from them, provided an opportunity of obtaining a strict, quantitative, evaluation of the phase plate image.

In brief, our method of analysis was as follows. First, we examined the Fourier transforms of the phase plate images of chosen specimens, checking that the positions and phases of their diffraction maxima were in accord with what one would expect from their known structure. Secondly, we reconstructed 'noise' filtered images, comparing these with others in which appropriate corrections had been applied to eliminate artefacts arising from the presence of the thread. And thirdly, we compared the amplitudes of corresponding diffraction maxima arising from both phase plate and bright field images of the same specimens (so checking the predictions of §4.3 that the phase plate image would be formed with each

---

Description of plate 27

Figure 14. Micrograph of the test specimens (and also some ferritin molecules) taken with a phase plate having a thread diameter of 0.4 μm and a beam of unscattered electrons having a diameter of 1.0 μm (giving $q_0 = 0.5$). A rectangle has been drawn over the micrograph to indicate the section of specimen selected for analysis. (Magn. ×1380000.)

The inset shows the correctly oriented optical diffraction pattern of a large area containing both the test specimen and the carbon support film. (1 mm = 0.05 nm^{-1}.)

Fourier component over a wide range of spatial frequencies contributing at the correct weight). Representative examples illustrating the findings of the analysis are discussed below.

5.1. *The Fourier transform of the phase plate image*

This section and the following one will be concerned with the micrograph shown in figure 14, plate 27. This micrograph was taken with a fairly low value for the 'absorption' parameter ($q_0 = 0.5$) and it therefore shows up the biological material in fairly strong bright contrast (but not so strong that the linear relation between optical density and electron dose would no longer apply).

None of the avoidable artefacts discussed in §3 are immediately apparent in the micrograph. Furthermore, optical diffraction patterns from it which show some of the diffraction maxima from the specimen as well as the 'noise' pattern from the support film (inset to figure), indicate that all the rays diffracted from the specimen fall within the radius at which the first reversal in contrast takes place. Charge and focusing errors are therefore also too small to be of any consequence (see §3.4).

The 'diffraction effect' manifests itself in the optical diffraction pattern as a bisecting band of weakened intensity having a width corresponding to the extent to which the various (electron) diffraction maxima have interacted with the thread. As is made clear in figure 6, it is not only the maxima centred in the direct line of the thread, but also those centred in a narrow zone on either side, which have been weakened. The 'noise' intensities in the central region of this band are not as reduced as might at first be expected: a similar, relatively intense, central line does not appear, for instance, in the area of the optical diffraction pattern corresponding to the supports for the rings of a zone plate (Möllenstedt *et al.* 1968, figure 4). Its presence in our case can be attributed to the diffraction streaks associated with the fine detail in the carbon support film being only weakly recorded in the image in comparison with other, real, structural features. This situation arises because such extremely thin threads as are used for the phase plates produce interference effects which extend for distances that are considerably larger than the distance over which the illumination remains almost coherent (see §3.1).

Optical diffraction patterns can only supply information about the amplitudes and directions of the diffracted rays. Unfortunately, this information alone is not sufficient to establish that the specimen structure is correctly represented in the micrograph: it is also necessary to confirm that the phase plate has not disturbed the phase relationships between any of these rays.

One way of obtaining this confirmation is to calculate the Fourier transform of the specimen image, using optical diffraction to select a section in which its symmetry is well preserved. The phase relations thereby regenerated can then be compared directly with those expected on the basis of the specimen's known symmetry.

This procedure was applied to the apparently well preserved region of specimen boxed off in figure 14, the Fourier transform of the digitized image being computed according to the methods described by DeRosier & Moore (1970). After adjustment of the phase origin so as to lie exactly on the axis of the specimen, it was found possible to accept or reject a given maximum, whose position was consistent with the specimen structure, depending on whether or not the relative

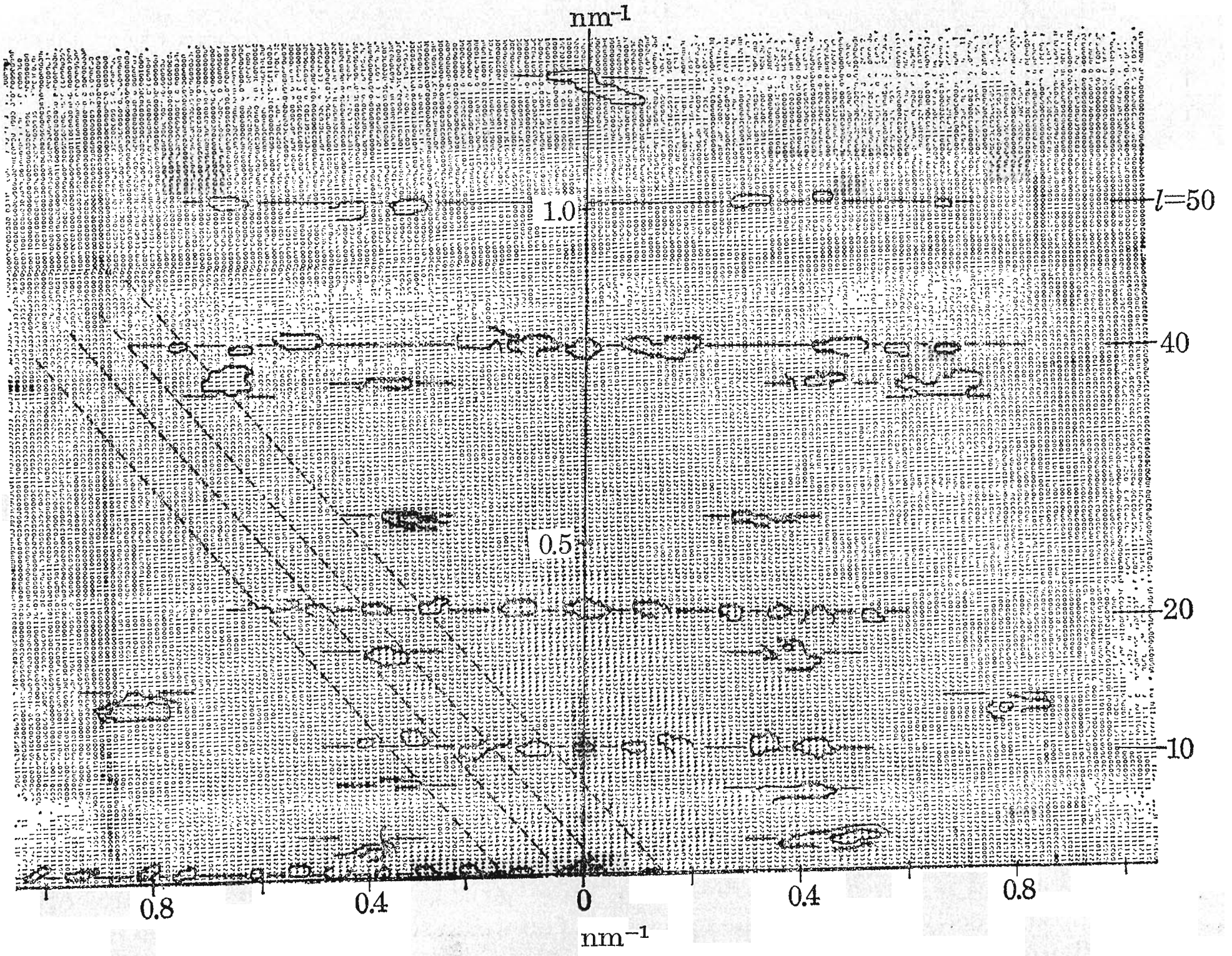


FIGURE 15. Map of those amplitude peaks in the Fourier transform of the region boxed off in figure 14 whose phases and positions indicate that they belong to the specimen structure; four obvious meridional maxima are also included. The tracing of these peaks directly over the line printer display of the computer has meant about 10 % compression of the scale in the horizontal direction. The relative position of the thread (inner pair of broken lines) and the extent to which the diffraction maxima have interacted with it (outer pair of broken lines) is also indicated.

phase of the equivalent maximum on the other side of the meridian was within *ca.* 30° of its expected value (see Finch & Klug 1971). Those maxima in this way considered to be characteristic of the specimen structure were traced over the line printer amplitude output of the computer, together with four distinct meridional maxima associated with the 5 nm disk repeat, giving the display shown in figure 15. This display can be identified with the complete half of the Fourier transform of the specimen structure to a resolution of *ca.* 0.85 nm: it agrees with

the Fourier transforms obtained from electron micrographs of conventional images (Finch & Klug 1971) and correlates closely with the X-ray diffraction patterns obtained from oriented gels of stacked disk rods (J. T. Finch, unpublished). There is therefore no evidence at all for the phase relationships between any of the diffracted rays having been disturbed.

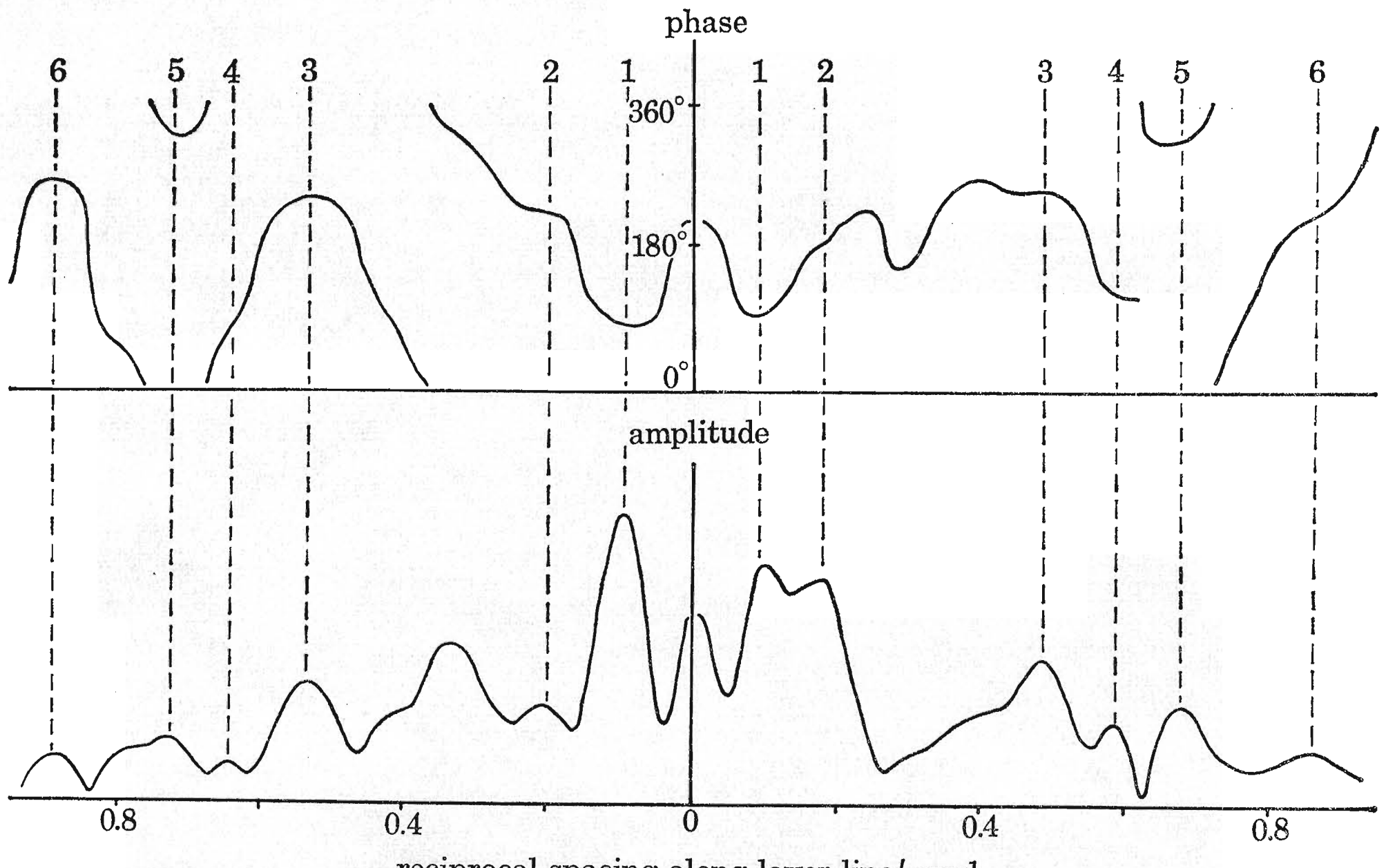


FIGURE 16. Graph of phases and amplitudes along the 40th layer line. The numbers refer to the six peaks for which the correlation is considered to be satisfactory. A fairly prominent peak between those marked 2 and 3 on the left hand side of the meridian shows no correlation and is therefore attributed to 'noise' arising from structural features other than those directly relating to the specimen. The weakest peaks, marked 6, do not show up in figure 15.

A graph of amplitudes and phases along the 40th layer line is reproduced in figure 16 to illustrate the limits of the sort of correlation considered satisfactory. The structure factor along this layer line, as with other layer lines whose numbers are multiples of 10, involves a zero order Bessel function, J_0. This means that the amplitudes and phases of corresponding maxima on either side of the meridian should be equal. As can be seen, there is reasonable correlation both in amplitude and phase down to reciprocal spacings, along the layer line, of 0.85 nm^{-1}. Better correlation is of course achieved with the more intense lower resolution maxima, and there is no evidence for the correlation being any worse in the phase plate image than it is in a corresponding bright field image.

The amplitudes do not correlate as well as the phases in figure 16, but this is also the situation with the bright field image. Irregularities in the staining are

thought to be responsible for the poorer correlation in the bright field case (Klug & Berger 1964), but loss of biological material due to irradiation effects (Thach & Thach 1971) may be more important in our case. In an exact analysis account should also be taken of the fact that the outermost peaks to the left of the meridian in figure 16 fall into the region where the amplitudes have been directly modified by the thread.

The significance of these outermost peaks showing the correct phase agreement should be stressed. This is true for similarly affected peaks on other layer lines, and there is therefore no evidence for any anomalous phase shifts associated with electrons passing immediately adjacent to the thread's surface.

The good phase correlation achieved over considerable regions between some maxima on layer lines other than those involving a zero order Bessel function deserves some comment. The broad high resolution peaks on the 36th layer line – which are also prominent in the X-ray diffraction patterns from oriented gels of stacked disk rods – were particularly striking in this respect. Maxima on these layer lines involve high order Bessel functions – J_{17} and J_{34} – and their good phase correlation must, in the first place, reflect the degree of order within the specimen, but, with the weaker maxima, may in part be a consequence of their relatively greater prominence in the phase plate image (see §5.3).

Finally, it is pointed out that a number of Fourier transforms with essentially the same features as the one in figure 15 have been obtained from phase plate images. The most special feature of this one is that the thread, in having directly modified the amplitudes of 19% of the diffraction maxima, was so oriented as to maximize the difference between an actual image and one which would be obtained after correcting for the defects arising from its physical presence.

5.2. *Reconstructed one-sided images*

By including just those Fourier components associated with the specimen structure in a second Fourier summation, an image of it could be reconstructed in which most of the irrelevant background detail has been eliminated. This image would nevertheless still be complicated by the fact that it is of the two sides of the specimen superimposed. To obtain further simplification one should therefore extract the image corresponding to one side only. This can easily be done by excluding in the summation all Fourier components associated with the other side (see the article by Klug & DeRosier (1966), which also shows how one can determine, from quite elementary considerations, which diffraction maxima correspond to which side). A comparison between a phase plate image thus simplified and a similar one, in which the appropriate compensations have been made, should then disclose, in the most distinct way possible, the nature of the modifications brought about by the physical presence of the thread.

We have made such a comparison, using the information provided by the Fourier transform displayed in figure 15, and the computer methods of DeRosier & Moore (1970) to perform the required manipulations. Maxima associated with the

J_0 give information only on the longitudinal and radial variation in the specimen density and hence derive equally from both sides. All these were accordingly included in the summations and with their amplitudes halved. The remaining maxima arise from the families of 17 parallel helical grooves running up the specimen through the subunits of each disk, and derive from separate sides. Of these, only the ones which derived from the only side having a J_{17} or J_{34}-type maximum (that on the 17th layer line) interacting strongly with the thread were included.

The correct amplitudes for maxima lying within the band of interaction with the thread were found by considering each relevant point in the digitized amplitude output separately. The points were 'placed' in their equivalent positions in the diffraction plane of the electron microscope and expanded over areas equal to that covered by the unscattered electron beam (or any diffracted beam) in this plane. The proportion of the expanded area covered by the thread determines the *intensity* reduction associated with each point, and from this the factor needed to raise the *amplitude* of each point to the value it would have if no thread were present was easily calculated. (Compensating in this way is equivalent to deconvoluting the Fourier transform of the absorbing function, discussed in §3.1, from the object density function).

The uncorrected and compensated one-sided views thus reconstructed for a single disk at a particular orientation are compared, together with an enlarged image of a whole disk at the same orientation, in figure 17, plate 28. Of course, the projection used in these reconstructions, in being the same as for the image, is not the one we would use if we wished to demonstrate the regular nature of the subunit packing within the disk, and even in these one-sided views the projected densities resulting from the subunits in various degrees of overlap have produced a rather complex pattern.

The most striking aspect of this comparison is that the thread, despite its interaction with a high proportion of the diffraction maxima, has only very slightly modified the appearance of the image: the pairing between the protein rings (see §2), and the density fluctuations along them, are just a little less pronounced than they should be. (These features could be expected from the strong interaction of the thread with, respectively, the 10th and 17th layer lines.)

The reason for this modification being a comparatively minor one clearly stems from the thread's behaviour as an absorbing function rather than a straightforward strip or blocking function. As we expected from §3.1 and have now demonstrated, the thread neither eliminates nor changes the phase of any of the diffracted beams with which it interacts, but merely weakens their amplitudes. This property is a characteristic of the method of observation, and we can therefore be confident that the presence of the thread will not *per se* lead to much inaccuracy in the image's representation of the high resolution detail.

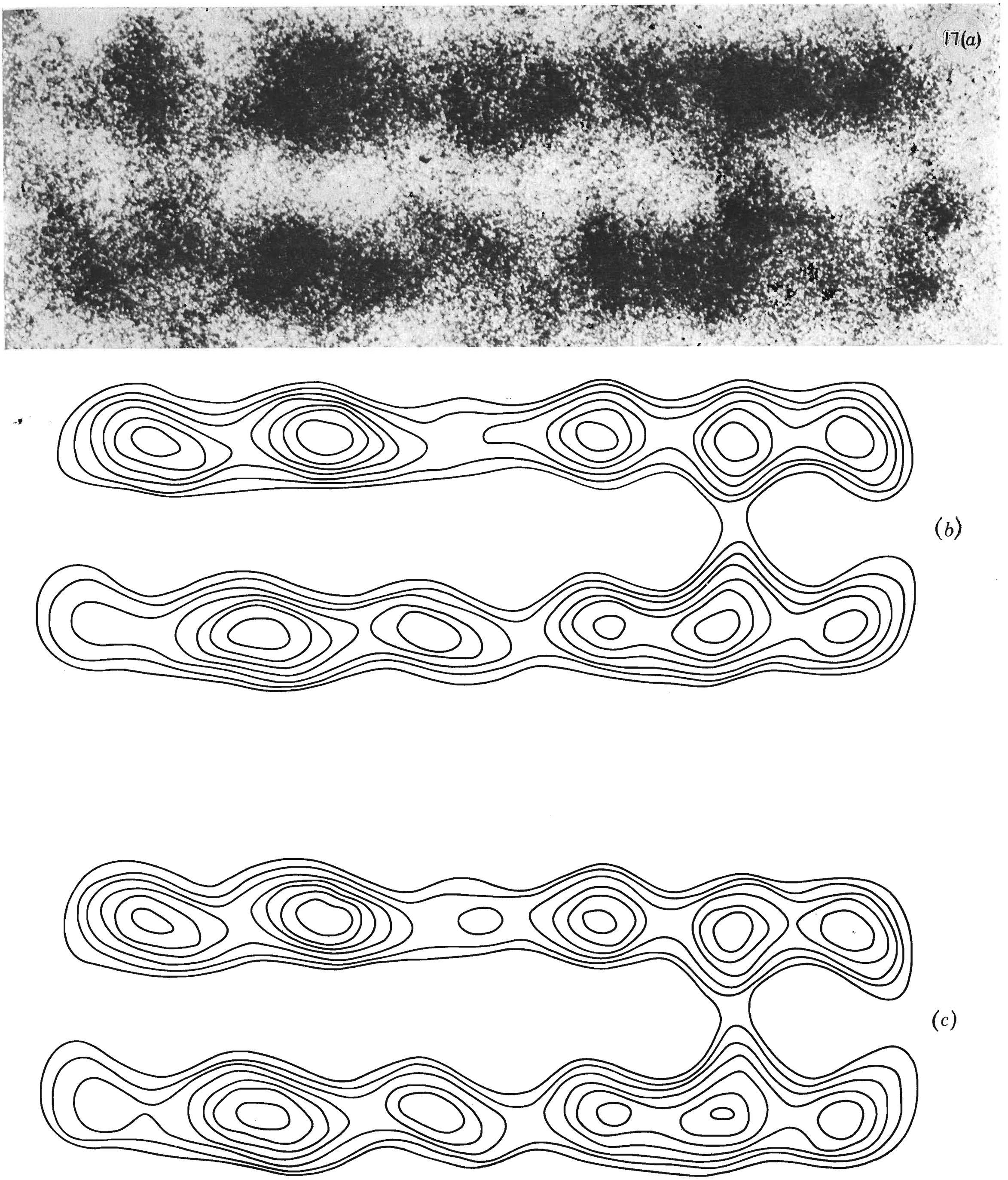
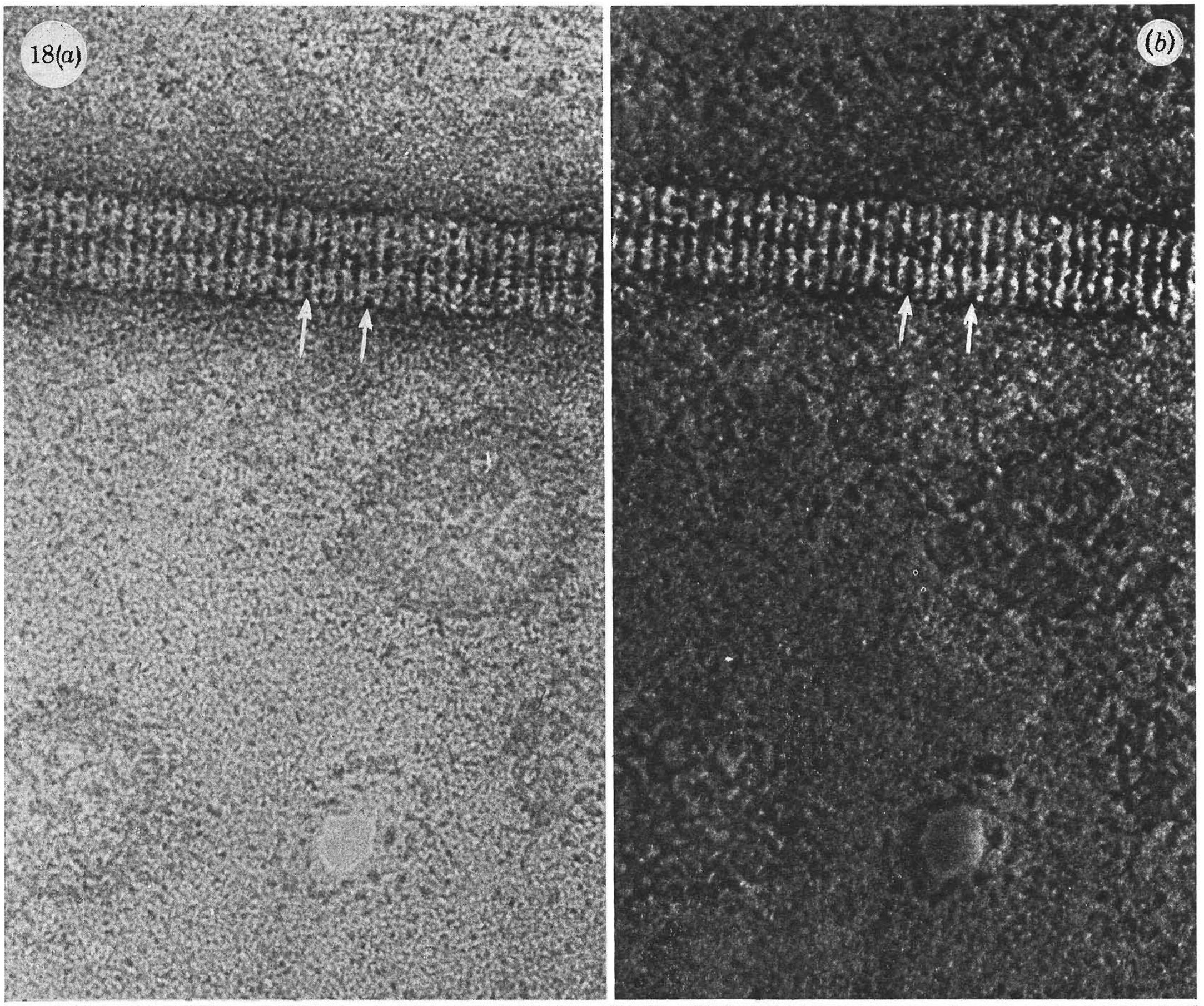
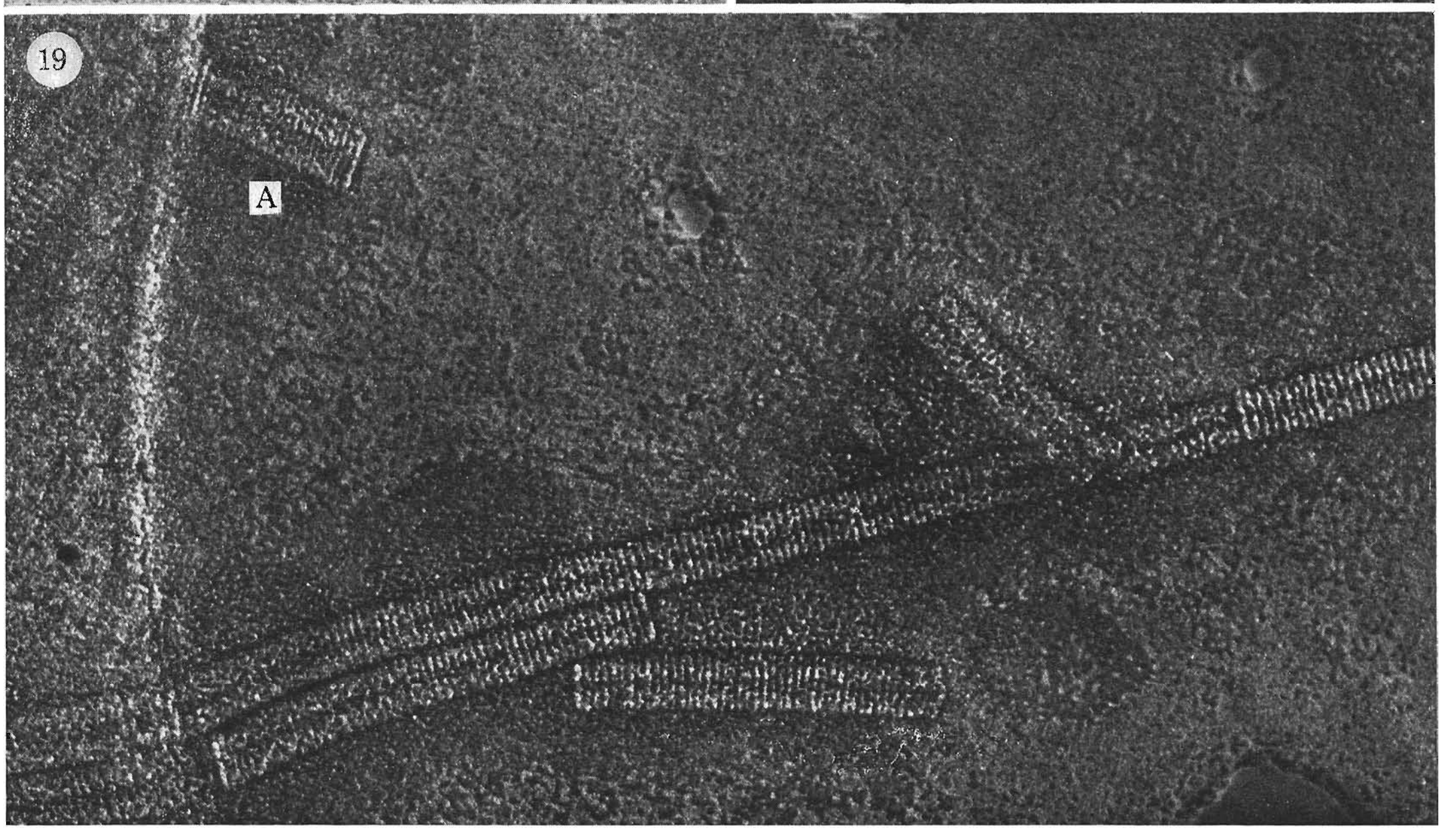


FIGURE 17. A comparison between (*a*) one of the disks in the area boxed off in figure 14, (*b*) its reconstructed one-sided image, and (*c*) its reconstructed one-sided image after compensating for the reduction in the amplitudes of the various diffraction maxima brought about by the physical presence of the thread. The original image has been printed in reversed contrast and, to facilitate the comparison, the reconstructed images have been depicted as equally divided density contours. (Magn. $\times 10\,000\,000$.)

(*Facing p.* 350)





FIGURES 18*a*, *b* and 19. For legends see facing page.

5.3. *Comparison of phase plate and bright field images*

In a number of careful comparisons between the Fourier transforms of the same region of specimen imaged both by our method and the bright field method, no essential difference was found between corresponding maxima in terms of their position and their phase correlation. The relative amplitudes of these maxima, apart from one *consistent* anomaly, could also be accounted for in terms of differences in contrast transfer properties for the two cases.

The behaviour of the amplitudes is important in terms of the predictions of §4.3 and the anomaly is a particularly interesting one. We will therefore discuss these findings in some detail, comparing by way of illustration, the diffracted amplitudes from identical regions of specimen in the pair of images reproduced in figure 18, plate 29.

The bright field image (figure 18*a*) has been defocused sufficiently to maximize the dark contrast among the higher spatial frequency components. An actual defocus, δf, of $+160 \pm 20$ nm was estimated for this micrograph by correlating the positions of the rings of maximum and minimum noise intensity in the optical diffraction pattern of an area containing just the carbon support film with the positions of the peaks and zeros of a set of contrast transfer functions calculated for different values of δf and assuming Q to be 0.06 (§4.3). The phase plate image (figure 18*b*) is an especially simple one in this case, since the 'absorption' parameter ($q_0 \approx 0.7$) had been so adjusted as to give virtually no amplitude contrast from the carbonaceous material (see figure 19, plate 29).

The amplitude ratio, bright field image to phase plate image, for each diffraction maximum was determined by carefully comparing equal equivalent areas in the numerical displays for the two Fourier transforms. The maximum ratio, with the same amplitude for the central order term in either case, was 1.2. It occurs – as we shall see – at a spacing where the efficiency of contrast transfer for both images is a maximum and indicates, therefore, that the maximum contrast in both specimen images, when $q_0 \approx 0.7$ for the phase plate image, is similar. The point most relevant to the present results, however, and one which is readily made

---

Description of plate 29

Figure 18. The same specimen imaged (*a*) in bright field with a 30 μm objective aperture in position and (*b*) with the phase plate ($q_0 \approx 0.7$). The weak, but quite definite, 'splits' in the four protein rings between the arrows, apparent only in (*a*), is probably an imaging, rather than a structural, effect (cf. figure 21*a*). (Magn. × 740 000.)

Figure 19. The general area from which figure 18 was selected. The small difference in intensity levels between the thin carbon support film and the thick carbon film on the left indicates that there is very little amplitude contrast present in the carbonaceous material. The terminating protein rings, particularly those of the specimen lying most nearly perpendicular to the thread (the one marked A), appear slightly brighter than the rest; this effect is probably due to a substantial weakening of the very low resolution Fourier components (cf. figure 21*b*). (Magn. × 355 000.)

apparent by comparing figures 18*a* and *b*, is that this contrast has different origins in the two images; in the bright field image it is contributed mainly by the stain, but in the phase plate image it is contributed mainly by the biological material.

The amplitude ratios were scaled so as to make the maximum ratio unity, and plotted as a function of reciprocal spacing. The results applying to those ratios where, in the case of the phase plate image, interaction with the thread had not led to some modification, are shown in figure 20, each point being labelled to designate the order of the Bessel function involved.

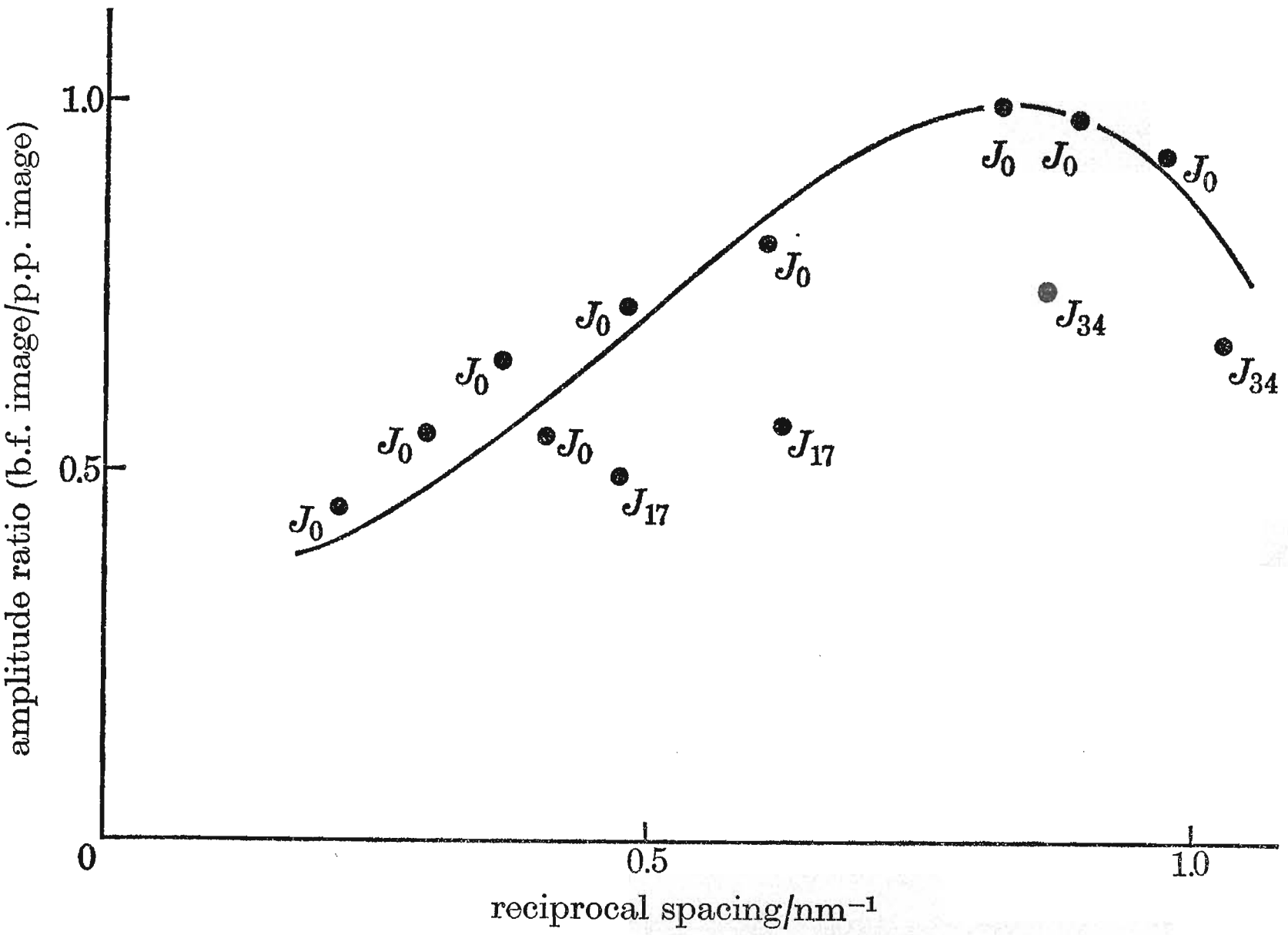


FIGURE 20. Graph of the amplitude ratios (bright field image/phase plate image) of corresponding diffraction maxima for the two specimen images shown in figure 18, plotted as a function of their reciprocal spacing. Each point is labelled to indicate the order of the Bessel function involved; a curve has also been drawn which represents the contrast transfer function appropriate to the bright field image (figure 18*a*).

Also drawn in this figure is part of the contrast transfer function† for the bright field imaging conditions, on the assumptions $\delta f = +160$ nm and $Q = 0.35$. The magnitude of Q in this case is determined by the electron scattering distribution due to a thin film of specimen material and a thin film of stain, summed in proportion to the relative amounts present in the space occupied by the specimen. The value chosen is only an estimate, based on a figure obtained experimentally from a similarly negatively stained specimen examined under similar conditions (Erickson & Klug 1971); but, unless it is substantial, a discrepancy between this

† It is actually the modulus of the contrast transfer function, $\sin\chi - Q\cos\chi$, scaled so that its maximum value is unity; this representation being appropriate to a comparison involving only magnitudes and scaled ratios.

estimate for Q and its actual value will not markedly affect the shape of the contrast transfer function over the region it is drawn.

The curve on which the experimental points should lie is that defined by the ratio of the contrast transfer functions – bright field image to phase plate image. All the points in figure 20 (excluding those designated by J_{17} and J_{34}, which will be considered separately) are, however, seen to lie very close to this bright field contrast transfer curve. This result could have been expected since the contrast transfer function for the phase plate image (figure 12) is near unity over the entire extent the curve is drawn. We thus have experimental confirmation, albeit indirect, of the phase plate's contrast transfer properties outlined in §4.3.

It is, further, reassuring to find no evidence in this sort of comparison that incorrect amplitudes have arisen as a result of second order effects due to the unscattered beam having been weakened too far. The discrepancies associated with the maxima involving the J_{17} and J_{34} Bessel functions cannot be accounted for on this basis since they do not bear a simple geometrical relation to the strongest maxima in the diffraction pattern. That the corresponding points all fall significantly below the curve drawn through the remaining points can therefore only mean that the phase plate images emphasize the structural features giving rise to these maxima (the helical grooves lying between the subunits) relatively more strongly than do the bright field images.

It seems most unlikely that such selective enhancement of certain diffraction maxima over a wide range of spatial frequencies could have arisen as a result of differences between the two images due to increased radiation damage, poorer lens stabilities etc. in the case of the bright field image. Effects such as these are bound to cause a weakening or disappearance of the highest resolution maxima (Kobayashi & Sakaoku 1965; Glaeser 1971) and there is no evidence for any significant differences between the two images in this respect.

On the other hand, the discrepancies could have arisen if the contours of the stain did not quite match up with those of the specimen. This is a state of affairs which must almost inevitably develop to some extent during the setting of the stain and during the subsequent irradiation. Imperfect matching resulting from irradiation induced events does indeed seem to be inferred from electron diffraction experiments (P. N. T. Unwin, unpublished), in which a very noticeable simultaneous weakening of the diffraction peaks from the specimens and sharpening of the crystalline ring pattern from the stain is found to occur during the period *before* the electron dose has become of the order of that required to focus and form a satisfactory high resolution image. Rapid crystallization, involving considerable atomic movement and reorganization, is clearly taking place within the stain during this initial period and presumably, because of surface energy effects and the tighter packing involved when the atoms become integrated into the crystal structures, this leads to some characteristic morphological changes, in particular some rounding off of any particularly sharp edges.

With this picture of events occurring in the stain, the explanation for the

greater emphasis of the grooves between the subunits in phase plate images seems straightforward enough. It is simply that these images (which refer more especially to the specimen) show the real grooves, which may be quite deep, whereas bright field images (which refer more especially to the stain) show only the rounded off projections of the stain into these grooves. Similar differences in emphasis were not detected among other regions of the specimens, but this is not unreasonable since no other regions demand such abrupt changes in the profile of the stain and are therefore likely to be as susceptible to rounding off processes and to coming apart.

It appears then that in being unable to accommodate the quite drastic changes occurring in the surrounding film of stain during irradiation, but at the same time somehow being prevented from collapsing by it, the specimen is the more adept of the two materials at preserving its original shape. A similar explanation for the above difference between the two images could, however, be derived if it was assumed that the stain was unable to penetrate deeply into the grooves in the first place. Until the relative importance of these two sources of mismatch is known, therefore, the primary cause of the discrepancies cannot be established unequivocally.

6. Discussion

The analysis of §5 has served two main purposes: first, it has confirmed that the presence of the thread, providing the phase plate has been reasonably well adjusted, only leads to very slight modifications of the high resolution detail in the image; and secondly, it has demonstrated that the phase plate is capable of providing a realistic image of biological structures to resolutions at least as high as 0.85 nm – a limit that is set by the degree of specimen preservation, not by the device. It is not possible to produce such realistic images, to comparable resolutions, by normal defocusing methods because of restrictions entailed in the range of spatial frequencies over which the bright field contrast transfer function can be made uniform.

The latter difference between the two types of image is not trivial, and to emphasize this we have sketched out in figure 21 the results of two one dimensional Fourier syntheses using the diffraction pattern of a grating-like object and modulating it by the same contrast transfer functions as apply to the pair of specimen images in figure 18. Note we have a positive grating for the phase plate image since the contrast here is most sensitive to small variations in the thickness of the *positive* biological material, and its negative counterpart for the bright field image, where the contrast is most sensitive to small variations in the thickness of the *negative* stain.

Our calculations show the bright field image to look rather different to the original grating: over-emphasis of the higher order Fourier components has caused quite a pronounced dip to form at the centre of each repeating unit. Unfortunately if one attempts to eliminate this defect by increasing the defocus the high resolution detail will be lost, or worse still, appear in reversed contrast. Decreasing the

defocus may lead to a more realistic image, but only at the expense of considerable loss in contrast.

The lines of apparently increased density within some of the protein rings in figure 18*a*, may well be examples of this bright field defect (compare these rings with the same ones in figure 18*b*).

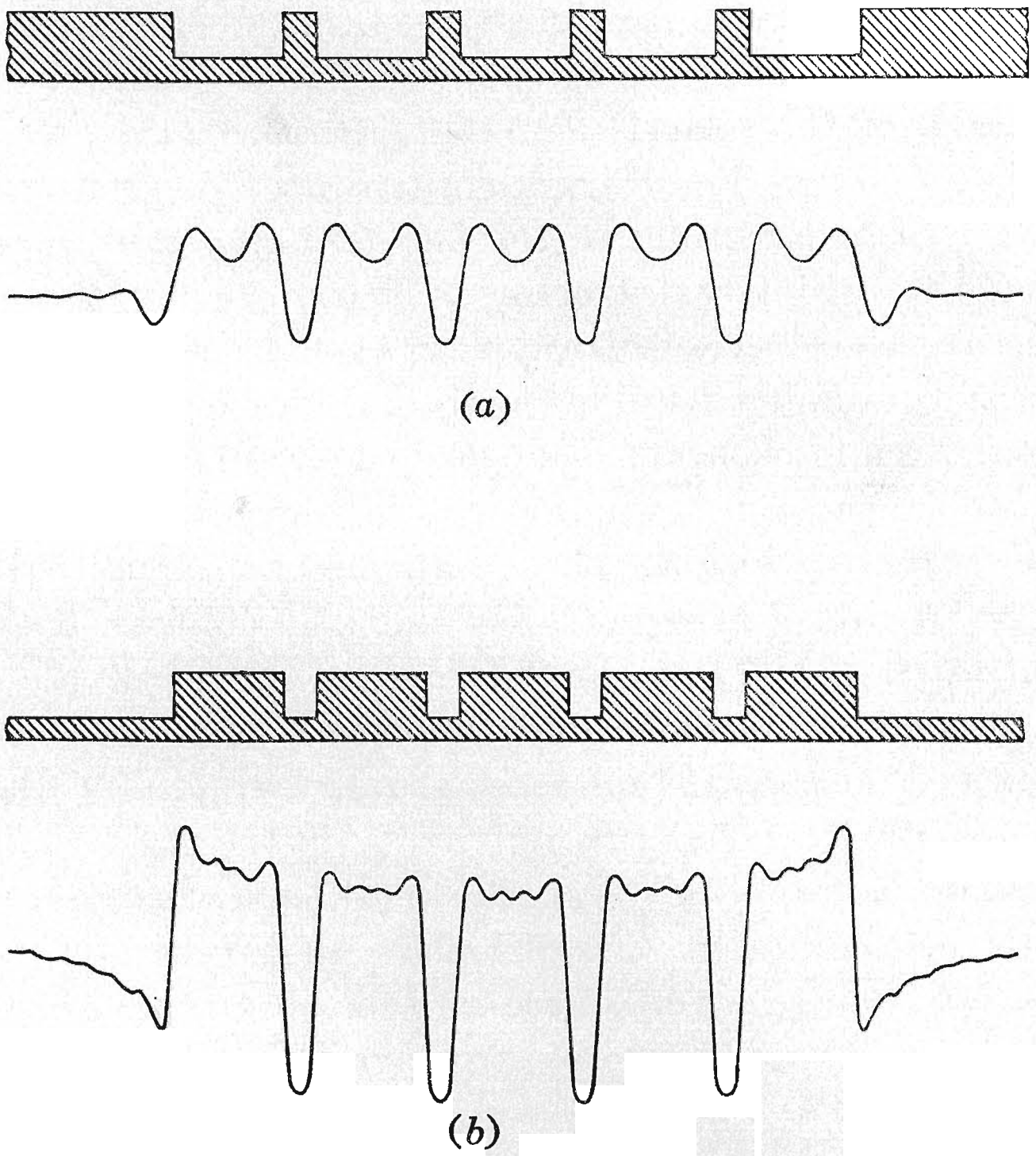


FIGURE 21. Fourier syntheses of the diffraction pattern formed by a one dimensional grating having a restricted number of units with a 2.5 nm repeat. In (*a*) the diffracted amplitudes have been modulated by the same contrast transfer function as was deduced for the image in figure 18*a* (i.e. as for a bright field image at a defocus of +160 nm); in (*b*) these have been modulated by the phase contrast transfer function shown in figure 12 and, additionally, by the same absorption function as used in figure 8*a* (i.e. as for a phase plate image). Only those parts of the contrast transfer functions for which $\chi < \pi$ have been included in the calculations.

Because of the more or less equal weighting for all but the lowest resolution Fourier components, the image calculated for the phase plate bears an altogether closer resemblance to the grating pattern. The two terminal units appear brighter than the rest, but this defect is clearly not very serious. Attention should also be drawn to the improved contrast in this image (which arises partly through relatively greater emphasis of the lower order Fourier components and partly through a weakening of the central order beam) and to the regions between the grating units being considerably darker than the background level. These effects

have been exaggerated somewhat by using a lower value for q_0 (0.4) in the calculations than applies to figure 18*b*.

The darker than background regions are conspicuous between the protein rings in all phase plate images (they should not be interpreted simply in terms of greater stain thickness) and many of these images also show the terminal protein rings of the rods to be slightly brighter than the rest, particularly when the specimen axis lies nearly perpendicular to the thread; see figure 19 or figure 7, bottom right.

Now, while it is clear both from the above similarity between calculated and actual images and from the analysis of §5 that, apart from the possibilities of mismatch (§5.3), the approximation of simple positive–negative gratings is quite a good one, we have yet to take account of the fact that the stain, as a result of its crystallization (§4.1), has internal structure of its own. The effect of this additional departure from the scheme envisaged in figure 21 could be of the utmost importance. This is because the unwanted detail in the stain could be obscuring the structural detail of interest, and it may be possible to make it appear relatively less prominent in the phase plate image than it would in the corresponding conventional image.

There are two ways by which the phase plate might so enhance the information of interest. First, it can, if q_0 is made low enough, simultaneously weaken the amplitude contrast from the stain and strengthen the amplitude contrast from the specimen. Secondly, and much more important since phase contrast is the dominating contrast mechanism for most spatial frequencies, the phase plate allows *opposing* contrast mechanisms to operate on the stain while maintaining *additive* contrast mechanisms for the specimen (at least when $q_0 < 0.75$).

For the latter mechanism to work the phase shift generated would actually need to be just less than $\frac{1}{2}\pi$, as it is necessary for both types of contrast to contribute, yet for the specimen still to be strongly enhanced. But such a condition is achieved normally to $\theta/\lambda \simeq 1.2\ \mathrm{nm}^{-1}$ (see figure 12) and could easily be extended to higher spatial frequencies by having the thread slightly undercharged.

With this small deviation from the $\frac{1}{2}\pi$ phase shift, the efficiency of amplitude contrast transfer is poor. Thus effective cancellation in the stain will only occur if its maximum (bright) phase contrast contribution is much smaller than its maximum (dark) amplitude contrast contribution. Unfortunately, the actual ratio of these two is difficult to ascertain, being dictated largely by a function involving the convolution of the shape transforms of the constituent crystals with the cross-grating patterns formed by their lattices. For large crystals, where the shape transform only extends over very small angles, the available phase contrast contribution from spatial frequencies other than those in the region of the central order beam or the first Bragg ring (which lies just outside the aperture edge) must be negligible. For very small crystals, on the other hand, the shape transform will be sufficiently broad that a significant phase contrast contribution can be expected among all spatial frequencies.

It is of interest to know the crystal diameter at which the phase contrast is just able to contribute among all spatial frequencies. Estimating this by making

the requirement that their central diffraction maxima, when convoluted with their lattice patterns, contribute some amplitude over the entire space within the first Bragg ring ($\theta \approx 10^{-2}$ rad), we obtain a figure of 0.9 nm.

The stains that were used contain large numbers of crystals of about this diameter (§ 4.1). Evidently, therefore, they do have the composition required to produce the relatively small, but *significant*, amount of phase contrast needed for the opposing contrast mechanism to be effective.

Improvements in phase plate images over corresponding bright field images which could be a consequence of this mechanism have been observed (Unwin 1971). The mechanism may also account for the appearance of unusually high resolution maxima as, for example, are displayed in figure 15. The improvements seem to be significant only among the smallest object spacings (*ca.* 1.2 nm and less). But this is not unreasonable because it may be only among these spacings where the efficiency of the dark amplitude contrast becomes poor enough (see figure 12) for the actual contributions from the opposing types of contrast to become comparable.

That improvements of the above nature appear at all is, in itself, a noteworthy result, lending further support to the evidence of § 5.3 that the morphology of the specimen is reasonably well preserved when it is encased in a thin film of stain. This is in marked contrast to the situation with unstained material, where almost instantaneous disintegration and mass loss is known to occur as soon as it is placed in the electron beam (Williams & Fisher 1970; Thach & Thach 1971). The negative stain therefore not only serves as a medium for providing contrast in bright field observation, but also as a very effective medium for keeping the morphology of the specimen intact. Although the mechanism by which it is able to do this is not yet clear (and may involve electrical and heat conduction as well as purely mechanical effects) our results do underline its importance: they show that this preserving role of the stain – now that a high resolution imaging method which strongly enhances detail in carbonaceous material is available – may enable more information to be obtained from biological specimens than has hitherto been possible.

7. Conclusions

The method of observation, in which an 'absorbing' electrostatic phase plate is used to create bright contrast in negatively stained biological specimens, has the following advantages over the conventional bright field method:

(*a*) To a resolution of at least 0.85 nm, it produces more faithful images. The resolution limit in the present investigation was set by the degree of specimen preservation; a real limit of *ca.* 0.5 nm is expected.

(*b*) Partly as a result of improved contrast transfer, and partly because it allows the intensity of the unscattered beam to be weakened, it enables the contrast to be increased several times over that attainable by the bright field method.

(*c*) It strongly emphasizes the contrast within the carbonaceous material, but

reduces the contrast within the stain. The stain is only essential to the method because it keeps the morphology of the specimen more or less intact during irradiation.

(*d*) There is good evidence that, owing to the simultaneous enhancement of the specimen detail and the weakening of detail within the stain, it can provide information about the specimen that it would not be possible to obtain with the bright field method.

Against these advantages are the more stringent focusing requirements and the necessary additional adjustments concerning the positioning of the phase plate and maintaining the correct electric field. These adjustments are not too critical, however, and can be made during observation using the appearance of the image as a guide.

I am very grateful to Dr H. E. Huxley, F.R.S. and Dr A. Klug, F.R.S., for many helpful suggestions and for their critical reading of an earlier version of the manuscript, and to Mrs L. Amos for her help with the computations.

References

Boersch, H. 1947 *Z. Naturf.* **2a**, 615–633.
Burge, R. E. & Silvester, N. R. 1960 *J. Biophys. Biochem. Cytol.* **8**, 1–11.
Burge, R. E. & Smith, G. H. 1962*a* *Nature, Lond.* **195**, 140–142.
Burge, R. E. & Smith, G. H. 1962*b* *Proc. Phys. Soc. Lond.* **79**, 673–690.
Crick, R. A. & Misell, D. L. 1971 *J. Phys. D: Appl. Phys.* **4**, 1–20.
DeRosier, D. J. & Klug, A. 1968 *Nature, Lond.* **217**, 130–134.
DeRosier, D. J. & Moore, P. B. 1970 *J. Mol. Biol.* **52**, 355–369.
Erickson, H. P. 1972 (to be published).
Erickson, H. P. & Klug, A. 1971 *Phil. Trans. R. Soc. Lond.* B **261**, 105–118.
Faget, J., Fagot, M. & Fert, C. 1960 *Proc. Reg. European Conf. Electron Microscopy*, Delft **1**, 18–19.
Finch, J. T. & Klug, A. 1971 *Phil. Trans. R. Soc. Lond.* B **261**, 211–219.
Finch, J. T., Leberman, R., Chang, Y. & Klug, A. 1966 *Nature, Lond.* **212**, 349–350.
Glaeser, R. M. 1971 *J. Ultrast. Res.* **36**, 466–482.
Haase, J. 1970 *Z. Naturforsch.* **25a**, 1219–1235.
Haine, M. E. & Cosslet, V. E. 1961 *The electron microscope*, p. 77. London: E. and F. N. Spon.
Hall, C. E. 1966 *Introduction to electron microscopy*, ch. 8. New York: McGraw-Hill.
Hanszen, K. J. 1969 *Z. Ang. Phys.* **27**, 125–131.
Hanszen, K. J. 1971 *Adv. Opt & Electron Microsc.* **4**, 1–84.
Hanszen, K. J. & Morgenstern, B. Z. 1965 *Z. angew Phys.* **19**, 215–227.
Hanszen, K. J. & Trepte, L. 1971 *Optik* **33**, 182–198.
Kanaya, K. & Kawakatsu, J. 1958 *Proc. 4th Int. Congr. Electron Microscopy*, **1**, 308–309.
Klug, A. & Berger, J. E. 1964 *J. Mol. Biol.* **10**, 565–569.
Klug, A. & DeRosier, D. J. 1966 *Nature, Lond.* **212**, 29–32.
Kobayashi, K. & Sakaoku, K. 1965 *Quantitative electron microscopy*, (eds. G. Bahr and E. Zeitler), pp. 359–376. Baltimore: Williams and Wilkins.
Misell, D. L. & Crick, R. A. 1971 *J. Phys. D: Appl. Phys.* **4**, 1668–1674.
Möllenstedt, G. 1950 *Optik* **6**, 251–255.
Möllenstedt, G., Speidel, R., Hoppe, W., Langer, R., Katerbau, K.-H. & Thon, F. 1968 *Proc. 4th European Regional Conf. Electron Microscopy*, Rome **1**, 125–126.

Thach, R. E. & Thach, S. S. 1971 *Biophys. J.* **11**, 214–210.
Thon, F. & Willasch, D. 1970 *Proc. 7th Int. Congr. Electron Microscopy*, Grenoble **1**, 3–4.
Unwin, P. N. T. 1970 *Ber. Bunsenges, phys. Chem.* **74**, 1137–1141.
Unwin, P. N. T. 1971 *Phil. Trans. R. Soc. Lond.* B **261**, 95–104.
Williams, R. C. & Fisher, H. W. 1970 *Biophys. J.* **10**, 53a.
Zernike, F. 1935 *Z. Tech. Phys.* **16**, 454–457.
Zernike, F. 1942 *Physica* **9**, 686–693.