Reconstruction of Three-dimensional Images from Electron Micrographs of Structures with Helical Symmetry

D. J. DeRosier† and P. B. Moore‡

Medical Research Council Laboratory of Molecular Biology
Hills Road, Cambridge, England

(Received 14 February 1970)

The application of the method of three-dimensional image reconstruction to electron micrographs of biological structures with helical symmetry is presented in detail.

1. Introduction

The optics of the electron microscope are such that, to a good approximation, a micrograph represents a projection of the density distribution of the specimen onto a plane. DeRosier & Klug (1968) have proposed a general method for reconstructing a three-dimensional image of a specimen from electron micrographs. Their method depends on the fact that the Fourier transform of a two-dimensional projection of a three-dimensional object is identical with the corresponding central section of the three-dimensional transform of the object. The three-dimensional transform can therefore be built up plane by plane using transforms of different projected views of the object. If sufficient data are available, the object can then be reconstructed by Fourier inversion of the resulting three-dimensional transform.

If an object has symmetry one can calculate several central sections from a single projection. Crowther, DeRosier & Klug (1970) have considered the application of the reconstruction method to particles possessing point group symmetries for which, despite their symmetry, more than one view of the structure is required. The purpose of this publication is to describe in detail the application of the method to those structures with helical symmetry for which only one view is required.

2. Materials and Methods

(a) Densitometry

A flying spot densitometer was used in this work. (For a complete description of this instrument, see Arndt, Crowther & Mallett, 1968). The densitometer is able to make several hundred measurements per second. Because many structures are being examined in this laboratory, an automatic densitometer is necessary. However, for research groups whose primary interest is solving one or two structures, it is possible to gather these data by hand, using a recording densitometer such as that produced by Joyce–Loebel. For example, to

† Present address: Clayton Foundation Biochemical Institute, Department of Chemistry, The University of Texas, Austin, Texas 78712, U.S.A.
‡ Present address: Department of Molecular Biophysics, Yale University, New Haven, Conn., U.S.A.
§ This is valid for negatively stained preparations to resolutions of 20 Å provided any defocusing is less than about 10,000 Å (Erickson & Klug, personal communication).
solve the structure of bacteriophage T4 tail (DeRosier & Klug, 1968) only 2800 optical densities were needed per tail image. Out of the several dozen images examined by optical diffraction, four images were processed on the computer to arrive at the final model, making a total of roughly 11,000 densities. This amount of data can be collected easily by hand in a few days.

(b) Computation

The computer used in processing the data was the IBM 360/44 at the Institute of Theoretical Astronomy, Cambridge, England. Programs were written in FORTRAN IV, but use was made of several machine-code subroutines which were made available by the computer facility. The program system was built around the Cooley–Tukey (1965) fast Fourier algorithm which was obtained from the IBM Scientific Subroutine Library. This algorithm greatly reduces the time required to calculate Fourier coefficients compared to programs not using fast Fourier transformations. However, its use does place restrictions on the processing; these will be considered later in the discussion.

The programs used in this work capitalized on the particular machine configuration at the computer center. Groups wishing to use this method will want to design and write programs tailored to their needs and their computational facilities. For this reason, we will confine our discussion to the basic algorithms and decisions which must be made along the way and omit any discussion of the encoding of the algorithm into a computer language.

3. Results

(a) An outline of the theory

The nature of helical symmetry makes it advantageous to use cylindrical polar co-ordinates in describing helical structures and their transforms. Thus the density distribution of the object is \( \rho(r, \phi, z) \) and its transform is \( F(R, \Phi, Z) \), where the axis of helical symmetry is the \( z \) direction. The equations for Fourier transformation in cylindrical polar co-ordinates are contained in an article by Klug, Crick & Wyckoff (1958). The following are those equations from that article which are relevant to the reconstruction process.

The Fourier transform of a structure, \( \rho(r, \phi, z) \), with helical symmetry can be expressed in the form

\[
F(R, \Phi, l/c) = \sum_n G_{n,l}(R) \exp \left( \imath n (\Phi + \pi/2) \right).
\]  

(1)

\( F(R, \Phi, Z) \) is non-zero only for \( Z = l/c \) where \( l \) is the layer line number and \( c \) is the axial repeat of the structure. \( c \) satisfies the condition \( \rho(r, \phi, z) = \rho(r, \phi, z + c) \).

The layer line number, \( l \), is linked to \( n \), the order of \( G_{n,l} \), by the relation

\[
l = in + um
\]

(2)

where \( t \) = the number of turns of the basic helix per axial repeat, \( u \) = the number of subunits per repeat, \( m \) = any integer which satisfies (2).

The reconstruction process consists of determining the functions \( G_{n,l}(R) \) from \( F(R, \Phi, l/c) \) and then calculating \( \rho(r, \phi, z) \) using the inverse Fourier–Bessel transformation

\[
\rho(r, \phi, z) = \sum_l \sum_n g_{n,l}(r) \exp(\imath n \phi) \exp(-2\pi \imath l z/c)
\]

(3)

where

\[
g_{n,l}(r) = \int G_{n,l}(R) J_n(2\pi R r) 2\pi R dR
\]

A three-dimensional image of a structure can be reconstructed from a single view if, to the resolution to which one is working, there is on any layer plane, \( l \), only one value
of \( n \) for which \( G_{n,t} \) is non-zero. Even though equation (2) generates many values of \( n \) for every \( t \), the properties of \( G_{n,t} \) are such that for large values of \( n \), \( G_{n,t}(R) \) is effectively zero for \( R < (n-2)/2\pi a \) where \( a \) is the largest radial dimension of the structure.

These relations thus fix the effective resolution, that is, the maximum value of \( R \), for which there is only one non-zero \( G \) function. In this region, equation (1) reduces to the following equation:

\[
F(R, \Phi, l/c) = G_{n,t}(R) \exp \text{in}[\Phi + \pi/2]. \tag{4}
\]

As an example, consider the first layer plane corresponding to a particle of radius 100 Å whose selection rule is given by \( l = 3n + 49m \) (these values apply to tobacco mosaic virus). We obtain \( n = -16 \) and \( n = +33 \) as the two lowest values of \( n \) for the first layer plane \( (l = 1) \). The \( G_{33} \) term does not interfere with the \( G_{10} \) term until \( R > (33 - 2)/2\pi 100 = 1/20 \) Å\(^{-1} \). On this layer plane one could work to a resolution of \( 1/R = 20 \) Å. The resolution limit is in general different for different layer planes so that the resolution of the final model in this case may be different from 20 Å.

The distribution of density in a micrograph corresponding to a projection of a structure along the \( x \) direction (normal to the helix axis) can be expressed as

\[
\sigma(y, z) = \int \rho(x, y, z) \, dx.
\]

The Fourier transform of \( \sigma(y, z) \) is \( F(0, Y, Z) \), the \( X = 0 \) central section of the three-dimensional transform. Conversion of \( F(0, Y, Z) \) to polar co-ordinates yields \( R = |Y| \) and \( \Phi = 0 \) for \( Y > 0 \) and \( \Phi = \pi \) for \( Y < 0 \). Thus using equation (4) we can obtain two estimates for each \( G_{n,t} \). These are given by:

\[
G_{n,t}(R) = F(R, 0, Z = l/c) \exp [-\text{in} \pi/2]
\]

\[
G_{n,t}(R) = F(R, \pi, Z = l/c) \exp [+\text{in} \pi/2]
\]

Differences in the two estimates of \( G_{n,t}(R) \) reflect differences between the near and far sides of the particle. As in the case of optical filtering of electron micrographs of helical particles (Klug & DeRosier, 1966; Moody, 1967), it is possible to assign one of the estimates of \( G \) on each layer line (except for \( n = 0 \)) to the near side of the particle and the other to the far side. Thus from a single particle one can carry out two three-dimensional reconstructions, one corresponding to the near side and the other to the far side of the particle.

While the final three-dimensional density map is generated from \( G_{n,t}(R) \) by Fourier-Bessel transformation given in equation (3) we could have generated the full three-dimensional transform from \( F(R, 0, Z) \) using the relation \( F(R, \Phi, Z) = F(R, 0, Z) \exp(\text{in} \Phi) \) and then carried out the inverse transformation in Cartesian co-ordinates. The Fourier-Bessel transformation has the advantage of utilizing the symmetry of the structure thus reducing the number of coefficients needed for the transformation.

The remainder of this paper is devoted to describing how equation (4) is utilized. In order to guide the reader through the wealth of detail, a brief outline of the steps involved in tackling an unknown structure is presented (see Fig. 1).

† It is possible to obtain from a single view, values of \( G_{n,t} \) in regions where two orders of \( G_{n,t} \) overlap provided one order is odd and the other is even. In other cases it is necessary to use more than one view (Crowther et al., 1970).
Fig. 1. The scheme presented shows the flow of data in the process of three-dimensional reconstruction.

(b) An outline of the general procedure

(1) Optical diffraction patterns are obtained of available images. The axial repeat distance of the structure is estimated and an indexing of the diffraction pattern is proposed.

(2) A small group of images is selected on the basis of the quality of their diffraction patterns. Each of these images is converted by densitometry to an array of optical densities which can be processed by a computer.
(3) The Fourier transform is computed from each density array and the Fourier coefficients are used to locate the axis of the particle and confirm the indexing. The transform also provides a measure of the amount and quality of the structural information present.

(4) Fourier coefficients lying along layer lines in each transform are used to compute a three-dimensional density map.

(5) The density maps from each of the images are used to construct a final model.

(c) Selection of micrographs

In selecting images it is important to examine as many particles as possible by optical diffraction. One usually finds variation in the particles and should pick for reconstruction several particles which are distributed over the range of variation. The particles chosen should not be curved and should be sufficiently far from other particles so that they can be masked off free from contributions of their neighbours. While the mathematical equations deal only with the projections of an isolated particle, it should be added that in certain cases it is possible to separate the contributions of three or four particles which might make up a limited segment of a twodimensional array (see Moore, Huxley & DeRosier, 1970). However, sections of crystalline arrays should be treated with great caution because one may have more than a single layer superimposed in the image.

Layer lines in the diffraction patterns of selected particles should be straight and should be as free of noise as possible. One should note that since the result of the flattening of a particle is to increase the diffracted intensity along layer lines (Moody, 1967), the diffraction pattern of a flattened particle is likely to appear cleaner than that of a comparable unflattened particle. Right–left symmetry in each of the layer lines provides an indication of the uniformity of staining and preservation, but need not be a requisite for selection for reconstruction.

It is not desirable, given our present knowledge concerning negative staining, to attempt to set out more detailed guidelines for accepting or rejecting images. Every structure should be treated individually and a sufficient body of experience built up on it in order to make decisions.

(d) Indexing of helical diffraction patterns

The most fruitful approach to indexing an optical diffraction pattern is to begin by constructing the reciprocal lattice. Several examples showing how this lattice is found are given in papers on optical filtering (Klug & DeRosier, 1966; Kiselev, DeRosier & Klug, 1968; Kiselev & Klug, 1969).

Once the lattice is known, values of \( n \) can be estimated from the position of maximum intensity on each layer line. For this purpose, the maximum can be thought of as resulting from the maximum of the Bessel function \( J_n(2\pi Ra) \) where \( a \) corresponds to an average radius for the \( n \)th helical family. Using a reasonable range for \( a \) one arrives at a range of values for \( n \) on each layer line. Since the values of \( n \) on different layer lines must obey a selection rule (equation (2)) it is usually possible to narrow the choice down to a few alternatives. The computed transform usually provides sufficient information to make the final choice (see section below on refinement of axis direction).
In order to process electron micrographs in a digital computer, the micrograph, a photographic transparency, must be converted into digital form. An efficient way to carry out this process is to measure the optical density of the micrograph at the points on a square grid and represent the picture as an array of optical density values. The instruments used for this purpose necessarily measure the average density of a small region of the micrograph rather than the density at a geometric point. The size and shape of the region measured will influence the relationship of the digital array to the micrograph. Thus, compared to the corresponding optical transform, the Fourier transform computed from the digital array will reflect (1) the sampling frequency at which densities were measured, and (2) the size and shape of the area "looked at" by the instrument each time it makes a measurement.

Analytically the result of Fourier transforming the array of optical densities can be expressed as follows:

$$F_c(R) = \left[ (F_m(R) \cdot F_s(R)) \ast F_1(R) \right] \ast F_0(R)$$

where $\ast$ denotes convolution, $F_c$ is the computed transform, $F_m$ is the transform of the continuous density distribution, $F_s$ is the transform of a function which has value of one everywhere inside the area transformed and zero elsewhere (i.e. the function which masks the particle from its surroundings), $F_0$ is the transform of the instrument sampling area and $F_1$ is the transform of the array of grid points at which the optical density is sampled. $F_m$, the micrograph transform, is the function needed to reconstruct the structure of the particle in question. The following six sections deal with estimating $F_m$ from $F_c$.

(f) Contribution of the densitometer to the transform

The effect of sampling the optical density of an image with a finite aperture is roughly equivalent to first convoluting the image with an aperture function and then sampling at geometrical points.† The transform of the image convoluted with the aperture function is given by $F_m(R) \cdot F_s(R)$. If the aperture is of radius $b$ then $F_s(R) = 2J_1(2\pi b R)/(2\pi b R)$.

(g) Contribution of sampling to the computed transform

The transform of a sampled function can be described by the convolution of the transform of the continuous function with the transform of the sampling function, that is, including the contribution of the instrument

$$\left[ (F_m(R) \cdot F_s(R)) \ast F_1(R) \right]$$

If the sampling function is a regular square grid with points at intervals of $a$ then $F_1$ is a square grid of points of spacings of $1/a$. The convolution can be regarded as a regular array made up by placing the transform $F_m(R) \cdot F_s(R)$ at every lattice point of $F_1$. One should adjust the parameters $a$ and $b$ to minimize the interference caused by the overlapping of neighboring transforms $F_m(R) \cdot F_s(R)$ in the array. For example, if resolution of 1/20 Å⁻¹ is desired, then a safe procedure is set to $a = 5$ Å, and $b = 2.5$ Å. The value of $F_s(1/20)$ will be 0.9. The coefficient $F_m(1/20)$ will be overlapped by $F_m(3/20)F_s(3/20)$. Since the transforms of micrographs are usually weak beyond $R = 1/10$ Å⁻¹ this overlap will make virtually no contribution to $F_c$.

† The photomultiplier averages the transmitted light intensity over the aperture, whereas the convolution requires an averaging of the logarithm of the optical density. However, this formalism will certainly account for the observed results in a qualitative way.
(h) Contribution of the image boundary to the transform

The array densitometered for computation can be considered as the product of an infinite density array multiplied by a function which is one everywhere within the densitometered area and zero everywhere else (i.e. a square wave or box function which masks off the desired portion of the image from its surroundings). When the densitometered array is transformed, the result is to convolute the infinite array transform, which is \( F_m(R) \cdot F_e(R) \), with the transform of the area function, \( F_v(R) \) (see equation (5)). For an array with a rectangular boundary, of dimensions \( c \) by \( d \),

\[
F_e(Y) = \frac{\sin \pi c X}{\pi c X} \cdot \frac{\sin \pi d Y}{\pi d Y}.
\]

The result of the convolution is to produce strong spikes running out from the transform origin which are, in general, the most striking feature of the transform.

The strength of these spikes is easily explained. Most micrographs are optically dense and the fluctuations in density within the image of a particle relatively small. When masking off an area of the image for transformation, points lying outside this area are assigned an optical density of zero and, unavoidably, a boundary between zero density and high density is introduced into the image. No other feature in the image is associated with so large a fluctuation in density; this fact is reflected in high amplitude spikes in the transform. When large areas of the transform are computed, it is often found that these spikes carry from replace one transform by unit cell to another in the convolution. A method of reducing this effect is given below.

(i) Alignment of micrographs in the densitometer

Ideally a micrograph of a helical object is densitometered so that one axis of the replace scan by grid is parallel to the helix axis. A rectangular boundary is constructed around the object delimiting the area to be transformed with the long axis of the rectangle parallel to the helical axis. When the transform of the image array is computed, the boundary spikes then lie on the lines connecting the centres of adjacent transforms as do the equators and meridians of the helical transforms. Since the equator and the meridional regions of most layer lines are often omitted from further steps in the analysis, this arrangement of the scan raster and the box conveniently puts the boundary spikes in parts of the transform where they do little harm. However, this ideal geometry is often not achieved in practice. When the box spikes are not perfectly aligned with the scanning raster, the box spikes of one transform invade the off-equatorial and off-meridional regions of neighbouring transforms leading to values for phases and amplitudes on layer lines considerably different from those of \( F_m \), the values needed. Thus in practical situations, one often wishes to suppress the contributions of the box to the transform.

(j) Minimizing the effect of the boundary

A procedure called "floating" has been devised to reduce the importance of \( F_e \) in the transform. In this procedure, the average value of the optical density around the perimeter of the box is subtracted from all the optical densities in the image array so that the optical density change at the edge of the image is reduced to an average value of zero while the magnitude of the density fluctuations inside the box remains the same. When the altered array is transformed, the box spikes have greatly reduced amplitudes while the required layer line data are unaffected. Our experience is that com-
Plate I (a). The optical transform of the extended tail sheath of the T4 bacteriophage. (b). The corresponding amplitude portion of the calculated transform.

The patterns are, of course, nearly identical. The differences which exist are due to irregularities in the emulsion of the electron micrograph. These irregularities affect the optical but not the computed transform. The use of a line printer display for the computed transform accounts for the 10% horizontal compression of the computed transform relative to the optical transform.
puted transforms correspond more closely to the transform of the continuous distribution as a result of this procedure. We feel that a floating step is a necessity when computing transforms. Once floating has been done, misorientation of the scan raster relative to the object can generally be ignored.

(k) The length of particle required

A projection of the infinite helix can be considered as a convolution of a single repeat with a set of points arranged in a straight line at intervals of c, the linear repeat distance. It is a simple matter using the properties of convoluted functions to show that the values of amplitude and phase measured along the layer lines of the transform of an integral number of repeats is the same as that of the transform of the infinite helix.

Thus by including an integral number of repeats in the image array and by controlling the contribution of the boundary, the computed transform can be made to correspond closely to the transform of a continuous, infinite particle.

(l) Line printer display of densitometered image

The optical density data is displayed on the line printer in the form of an array of symbols which vary in blackness so as to give a pictorial representation of the area scanned. Features of the original micrograph are easily recognized in these displays (see Moore et al., 1970). The repeat of the helix can be determined so that the scan grid dimensions can be related to distances on the micrograph. Furthermore, the boundaries of the object are quickly discerned so that the subset of the density array to be used in further computation can be defined.

The first portion of the Fourier transform routine is devoted to producing the subset of the original scan the operator specifies after inspecting the scan display (see section (m) below). The form this subset takes will depend on the programs which must subsequently use it. If desired, the subset of the scan array is then subjected to a floating step (see above) after which it is ready for Fourier transformation.

While precision scanning of micrographs would have some advantages, e.g. display of the scan and “boxing” (see below) the scan would be unnecessary, we feel that the generous scanning we currently use also has its good points. In the first place, it eliminates the problem of setting the micrograph in the densitometer with precision. Second, one often finds after doing some preliminary processing that one wishes to redefine or otherwise alter the area selected for reconstruction. These alterations can be done without densitometering again if a large area is scanned to begin with.

(m) The box algorithm

A routine is required which, by selecting a subset of the density array, effectively masks off all but that portion of the image which is needed for the Fourier transformation. This is achieved by setting to zero all those densities which lie outside a polygonal region. The polygon is specified by a set of m pairs of numbers corresponding to the m vertices of the polygon taken in sequential order. Each two adjacent vertices define an edge of the polygon. The program takes each edge in turn and computes its intersections (if any) with each of the rows of the array and stores them according to row. When all edges have been processed, the program takes the set of intersections for each row and orders them smallest to largest according to their distance from the left-hand
edge. Any density points between the left-hand edge and the first intersection are outside the polygon and are set to zero. Those between the first and the second intersections are inside the polygon. Those points between the second and third are outside and so on. The simplifying feature of this method is that in order to know which densities in a given row are inside the box, one need know only the locations of the intersections of that row with the box.

If the transform is to be calculated using some function of the optical density (e.g. the logarithm), this stage provides a good place for its insertion. Thus the program sets the density to zero if it is outside the box or replaces it by some function of its value if it is inside.

(n) Fourier transformation

The Fourier transformation program makes use of a fast Fourier algorithm (Cooley & Tukey, 1965). While this algorithm imposes restrictions on the computation, these are more than offset by its speed relative to other algorithms not in this class. The Cooley–Tukey algorithm requires that the number of densities of input be a power of two. In addition, the number of coefficients calculated is the same as the number of input densities and represents a regular sampling of a single repeat of the computed Fourier transform.

We find it convenient to use an input array of 256 by 256 points, giving a total of 65,536 densities. The selected area of the array of optical densities is usually much smaller than this and is brought up to the standard 256 by 256 size by the addition of density points of zero magnitude. We double the sampling frequency of the transform along the direction of the helix axis and, at the same time, eliminate the Friedel-related coefficients to maintain the output array at the standard size. The increased sampling ensures that narrow layer lines arising from long lengths of particles in the image array will be adequately sampled in the transform. It also allows bilinear interpolation of coefficients in the event the layer lines do not fall precisely on the rows of the computed transform.

In our program, the input and output array sizes were fixed, although there is no reason why they could not be made adjustable with some additional programming. The array size was chosen as the largest consistent with a reasonably short running time and yet able to handle all structures which were then under examination.

For example, in the densitometry of the extended form of bacteriophage T4 tail sheath, densities were measured every 8 \( \text{Å} \) so that the total input array corresponded to a square area \( 8 \times 256 = 2048 \text{Å} \) on a side. The transform sampling interval was \((5 \times 256)^{-1} = 0.00049 \text{Å}^{-1}\) in the \( y \) direction and \((2 \times 5 \times 256)^{-1} = 0.00024 \text{Å}^{-1}\) in the \( z \) direction (due to the doubled frequency of sampling). The repeat distance in the computed transform was \( 1/8 = 0.125 \text{Å}^{-1}\). Plate I shows the optical transform of extended bacteriophage T4 tail alongside its computed transform.

The running time for the transformation itself is about 2-5 minutes. However, the size of the density array made it necessary to factor the two-dimensional transform into sets of one-dimensional transforms in each dimension. Intermediate results were stored on disc. The additional data handling required added another 3-5 minutes to the program’s running time. The line printer display of the transform takes 4 minutes giving a total time for Fourier transformation and display of 10 minutes.

Several methods exist for altering the sampling frequency of the transform. The easiest way to increase the transform sampling from say 256 points per repeat to 512
is to add an extra 256 zero density points to the starting line of 256 densities and
carry out at 512 point transformation. Such increases can only be in powers of 2.
Another method which allows any integral increase is illustrated in the following
example. It is possible to shift the transform sampling grid along by 1/2 of a sampling
unit. In the case of phage tail this would correspond to sampling at \( Z = 0.00024 \),
0.00072 etc. as well as at \( Z = 0 \), 0.00048, etc. Sampling at these in-between points
is achieved by applying a complex phase-shift factor to the density array so that \( \rho(z) \)
is replaced by \( \rho'(z) = \rho(z) \exp(2\pi i z \Delta) \) where in this case \( \Delta = 0.00024 \). One then transforms \( \rho(z) \) and \( \rho'(z) \) and interleaves the two sets of coefficients.

**Computing the transforms of very large arrays**

There may be cases in which one wishes to transform arrays that are so large that
both execution time and storage space become limiting. It is possible to reduce the
size of the array by replacing groups of sampled densities with a single density which
is a weighted average of the densities in the group. Such a procedure is roughly
equivalent to sampling the image with decreased frequency using a sampling aperture
the size of the group of points. The decrease in the sampling frequency (i.e. the increase
in sampling interval) results in a reduced repeat distance in the computed transform.

The weighted averaging formula can be represented as a convolution of the density
array with the weighting function. The effect will be to multiply the transform of the
continuous image with the transform of the weighting function, thereby reducing the
amplitudes of the computed high resolution coefficients and also reducing the error
which would arise by the overlapping of neighboring transforms in the convolution.
The following example shows how this step is carried out.

Suppose the bacteriophage T4 tail sheath were densitometered every 4 Å instead of
every 8 Å. Every group of four adjacent non-overlapping points would be replaced by
a single point whose density was the average of the four. The new sampling interval
would be 8 Å and each repeat in the computed transform would be multiplied by the
normalized transform, \( W(Y, Z) \), of the four points where: \( W(Y, Z) = (1/4) \cos
(4\pi Y) \cos(4\pi Z) \). Note that this function falls to zero at the repeat interval, 0-125 Å, of
the computed transform. It is possible to modify the form of \( W \) by changing the grouping of points and by putting different weights on each point. For example, in a
linear case, one could use \( \rho_1/4 + \rho_2/2 + \rho_3/4 \) in which case \( W \) would be of the form
\( 1 + \sin 8\pi Z \).

It is clear that while the central region of such a computed transform closely
approximates the transform of the continuous object, the outer regions of the computed
transform are weighted down and will not, in general, approximate the transform of
the continuous object due to overlap. Such outer regions can be examined, however,
if one shifts them into the place previously occupied by the central region, i.e. move
the transform without moving \( W \). To do this one shifts the origin of the transform by
replacing \( \rho(y, z) \) by \( \rho(y, z) \exp(2\pi i (y\Delta Y + z\Delta Z)) \) (see above) where \( \Delta Y \) and \( \Delta Z \)
are the co-ordinates of the centre of the desired region. The complex densities are then
averaged and transformed as just described. Averaging the densities before applying
the shift factor would have the effect of shifting \( W \) and the transform as a unit.

**Displaying the transform**

Since the computed Fourier transform has over 65,000 coefficients, it is necessary to
display it in the most compact form possible. As we chose the lineprinter as the
output device for the display of the transform, each coefficient was represented by two alphameric characters, one corresponding to the amplitude and the other to the phase. The logarithm of the amplitudes were represented by integers, 0 to 9, and the phases were divided into 36 ten-degree intervals which were coded by the alphamericis A to Z, 0 to 9. The resulting amplitude array and phase array were printed separately with identical format so they could be superimposed.

(c) Interpolation

If the layer lines do not coincide with the rows of coefficients in the computed transform, it is necessary to interpolate to obtain the layer line data. The simplest form of interpolation is bilinear interpolation. In order to do bilinear interpolation in reciprocal space, the change in coefficients between sampling points must be small relative to their average magnitude. This situation will occur if two conditions are met: (1) the sampling in reciprocal space is fine relative to the inverse size of the object transformed; and (2) the phase is calculated relative to the exact centre of the object. For the purposes of bilinear interpolation, the layer lines were located in reciprocal space by inspection of the amplitude map put out at the end of the Fourier transformation step.

If a finely sampled transform is not required, a substantial saving in execution time and storage requirements could be achieved by reducing the size of the input array. In this case, however, if layer lines do not fall on the rows of the computed coefficients, a more sophisticated interpolation would be required. If the dimensions of the image array were c by d, the relationship between the computed set of $F(Y, Z)$'s and the desired values, $F(Y', Z')$ is:

$$F(Y', Z') = \sum_y \sum_z \frac{\sin \pi c (Y - Y')}{\pi c (Y - Y')} \frac{\sin \pi d (Z - Z')}{\pi d (Z - Z')} F(Y, Z)$$

(6)

(see Crowther et al., 1970). The question of what frequency of sampling is sufficient to ensure reasonable data using bilinear interpolation and the relative merits of bilinear interpolation and the interpolation scheme implied by equation (6) are still under investigation.

The layer line data used as input for reconstruction programs must be sampled more frequently than the reciprocal of the particle's diameter. If sampling is done less frequently, extra density will appear inside the computed density maps of the object due to the real space convolution which results from having sampled the layer lines in reciprocal space. It is our usual practice to sample the layer lines at intervals of the reciprocal of twice the diameter of the object so that its boundaries can clearly be seen in reconstructions.

(p) Analysis of computed layer line data

As shown in equation (4), the phases of points placed symmetrically on either side of the meridian on a layer line will differ by $\pi$ radians if $n$ on that layer line is odd and differ by zero if $n$ is even. Fourier transforms of helical particles are often computed which have strong amplitudes along their layer lines, but fail to fulfill this expectation. The usual reasons are that the phase origin of the density array was not chosen to lie on the helical axis and that the axis of the particle has been tilted out of the plane of projection.
As shown below it is possible using the calculated transform to determine both the amount by which the phase origin is displaced and also the amount by which the helical axis is tilted. For each calculated phase one can apply a correction $\Delta \theta_n$, the error in the phase angle due to displacement of the phase origin, and $\Delta \theta_t$, the error in the phase angle due to the tilt. The effect of displacing the phase origin from the axis of the particle is to shift the phases of the symmetrically placed coefficients in opposite directions. Let $F(Y', Z') = |F(Y', Z')| \exp i\theta(Y', Z')$ where $Y'$ and $Z'$ form the co-ordinate system in the transform. According to equation (4) the quantity $\Delta \theta$ defined by $\Delta \theta(Y', Z') = \theta(-Y', Z') - \theta(Y', Z') + n\pi$ should be zero. A shift, $\Delta y$, of the origin results in a phase change of $-2\pi \Delta y Y'$ at the point $Y', Z'$. The value of $\Delta \theta_{\phi}$ is given by $\Delta \theta_{\phi}(Y', Z') = +4\pi \Delta y Y'$. Note that $\Delta \theta_{\phi}$ is independent of $Z'$ and the value of $n$.

The effect of tilting the particle axis out of the plane of projection is more involved. In this case, the central section corresponding to the projection intersects a given annulus, $R = \text{constant}$, along a chord rather than a diameter (see Fig. 2). The values of symmetrically placed $F$'s, $F(Y', Z')$ and $F(-Y', Z')$, in the transform are thus $F(R, \Phi, Z)$ and $F(R, \pi - \Phi, Z)$, in polar co-ordinates, where $R = |Y'|/\cos \Phi$, $Z = Z' \cos \omega$, and $\omega$ is the angle of tilt of the particle axis.

![Fig. 2. Shown is the spatial relationship between the co-ordinate systems $R$, $\Phi$, $Z$ (=l/\rho c) of the Fourier-Bessel transformation and the Cartesian co-ordinate system $X'$, $Y'$, $Z'$ for the transform of the micrograph.](image)

As can be seen from Figure 2,

$$\Phi_t = \arctan \frac{d}{Y'}$$

and $d = Z' \sin \omega$. Thus

$$\Phi_t = \arctan \left( \frac{Z' \sin \omega}{Y'} \right)$$

Therefore, the difference, $\Delta \theta_t$, between the two values of $\theta$ at $Y'$ and $-Y'$ is

$$\Delta \theta_t = -2n \Phi_t = -2n \arctan \left( \frac{Z' \sin \omega}{Y'} \right)$$
Note that $\Delta \theta_c$ depends on $Z'$, $Y'$ and $n$ which distinguishes it from $\Delta \theta_s$. The total error $\Delta \theta$ for any pair of points in the transform is given by

$$\Delta \theta = \Delta \theta_s + \Delta \theta_c$$

One can then determine $\Delta y$ and $\omega$ using the values of $\Delta \theta(Y', Z')$ determined from the calculated transform. However, since $\Delta \theta$ is not a simple function of $\Delta y$, $n$ and $\omega$, it was found best to search for a solution numerically. Each point of the search procedure is derived from a single value for $\omega$, $\Delta y$ and an appropriate set of values for $n$ corresponding to a selection rule. For each pair of coefficients on a layer line, the quantity $\theta_c$ is computed; $\theta_c = + 2n \Delta y Y' - n \arctan (Z' \sin \omega |R'|)$ and is then applied to each member of the pair of coefficients. Thus:

$$F(Y', Z') = F'(Y', Z') \exp (i \theta_c)$$

$$F(-Y', Z') = F'(-Y', Z') \exp (-i \theta_c)$$

where $F'$ is the measured value of $F$.

One then evaluates the different trial values of $\omega$ and $\Delta y$ using the following equation:

$$Q = \sum Q(Y', Z') \text{ where } Q(Y', Z') = |F(-Y', Z') \exp (i \pi) - F(Y', Z')|.$$  

A somewhat simpler expression can be used to approximate $Q(Y', Z')$.

$$Q(Y', Z') = F(Y', Z') \Delta \theta \text{ where } F(Y', Z') = \frac{|F(Y', Z')| + |F(-Y', Z')|}{2}.$$  

Note that $\Delta \theta$ must be kept in the range $0 \leq \Delta \theta \leq \pi$ in order to use this approximation.

The physical significance of $Q$ is that it corresponds to the maximum possible error in the reconstructed density due to failure to obey helical symmetry. A summation of

![Fig. 3](image1)

![Fig. 4](image2)

**Fig. 3.** Each of curves a and b is the result of a search procedure for the angle by which the helix axis of a wavy pentamer tube (Kiselev & Klug, 1969) is tilted out of the plane normal to the direction of view. The images corresponding to curves a and b were part of a tilt series on a single tube. The angle $\omega(Q = \text{minimum})$ corresponds to the choice of $\omega$ which is most consistent with the helical symmetry. If the selection rule is changed, the curve of $Q \text{ versus } \omega$ will change. The correct choice of selection rule should give the lowest value of $Q = \text{minimum}$. The sign of $\omega(Q = \text{minimum})$ depends on the hand of the particle. The increase in the value of $Q = \text{minimum}$ from curve a to b is a result of the deterioration in image quality due to the build-up of contamination on the specimen between exposures of the tilt series.

**Fig. 4.** Shown is the result of a search procedure to locate the axis of helical symmetry. The data for this curve were first corrected for tilt and correspond to the data used for curve a in Fig. 3.
$\Delta \theta$ alone would not make sense in any case since one wishes to put more weight on those phases corresponding to large amplitudes rather than on those corresponding to small amplitudes, which are more likely to have been affected by noise. Figures 3 and 4 show the result of carrying out such a procedure on a tilt series of a helical array of the protein subunits from human wart virus. From the known direction of the tilt series and the results of the search, it was possible to deduce the sense of this helix.

It should be recognized that if one has a tilt series of a single particle, the above method can be used to determine the hand of the particle, to establish its indexing and to detect departures from perfect helical symmetry.

(q) Two-dimensional filtering by computer

Once the layer line data is available, either before or after phase correction, its inverse Fourier transformation can be computed. This operation yields a two-dimensional density map which is the digital equivalent of an optically filtered image. It is possible to separate the layer line data into its nearside and farside halves and look at the two images corresponding to either side of the object separately, just as in the optical case. The advantage of digital filtering over optical filtering is that the operator has much better control of the data included in the reconstruction when filtering is done by computer (e.g. see Moore et al., 1970).

(r) Three-dimensional reconstruction

The layer line data and indexing information can be combined to produce three-dimensional density maps. For this we used a set of programs which were modifications of those written by Dr K. C. Holmes for application to the X-ray analysis of the structure of tobacco mosaic virus. The essential steps performed by these programs were: (1) combination of the $F(R, \Phi, l/c)$s and $n$'s to yield $G_{n,l}(R)$'s; (2) Fourier-Bessel transformation of the $G_{n,l}(R)$'s to give $g_{n,l}(r)$'s; and (3) expansion of the $g_{n,l}(r)$'s to produce the values of $\rho(r, \phi, z)$. These programs simply realize in digital form the mathematics outlined in the theory section (see above).

There are several possible methods of displaying the data. The first is to produce cylindrical sections (i.e. $r = \text{constant}$). The second is to produce sections of $\phi = \text{constant}$. The third is to generate sections in which $z = \text{constant}$ and the densities in the section are given on a Cartesian grid. In this last case, the Cartesian array of densities is obtained from the polar array by bilinear interpolation.

Each method of displaying the map gives a different perspective because one can only see three-dimensional detail in a direction normal to the sections. The $z = \text{constant}$ sections have been found to be the most generally useful.

(s) Interpretation of the three-dimensional map

The final step of the three-dimensional reconstruction is the interpretation of the map, that is, deciding which features of the map are part of the particle and which are not.

It might be thought that maps of negatively stained particles would have uniformly dense regions corresponding to regions in the structure accessible to stain and empty regions corresponding to protein, which is electron transparent. The map, in fact, represents the distribution of stain averaged over the symmetry-related regions of the structure. If the probability of stain penetration into a given portion of the asymmetric unit is not zero or one, the density found in that corresponding region of the map
will have a value intermediate between that of fully stained and unstained. Thus a range of densities is encountered in the three-dimensional map. Such fluctuations in stain distribution as well as the size of the stain particles themselves limit the resolution of the maps so that the maps are characterized by broad peaks of density of varying maximum value and width.

An additional problem in interpreting the map arises because the distribution of stain outside the particle does not possess cylindrical symmetry. This feature of the stain distribution affects the equator but not appreciably the higher layer lines. The inclusion of the equatorial data in the final Fourier--Bessel transformation is therefore not strictly legitimate since it does not possess the requisite symmetry. The effect is to introduce a distorted radial density distribution. The result of omitting the equatorial data is equivalent to subtracting the radially averaged density distribution. Whichever is done, it is clear that it is not a good practice to set a single cut-off value in density which is used to separate particle from non-particle. Clearly there is room for further experimentation in the interpretation of boundaries.

In some maps, small pieces of structure which float free of the main body of the structure are encountered. To judge whether these pieces belong to the structure or not, it is necessary to examine reconstructions from other particles to see if they exist in other maps and if they sometimes appear to form a connection to the main structure. Obviously, if they are found to connect in some cases, there is reason to accept them as being "real" parts of the structure.

The criteria for choosing the boundary thus depend to a large extent on comparison of density maps of several particles. In general, these comparisons show that while the precise shape of the boundary may not be correct, there is no question that the distribution of mass is correct. Perhaps when more sophisticated methods become available for handling the equatorial data, it may be possible to make more definite rules for determining boundaries.

Both authors are in receipt of a National Science Foundation postdoctoral fellowship. One of us (P. B. M.) held a U.S. Air Force Office of Scientific Research fellowship during a portion of this work.

REFERENCES