$(1|0,\frac{1}{2},\frac{1}{2})Pnam$
$= \{(1|0,\frac{1}{2},\frac{1}{2}), (\bar{1}|0,\frac{1}{2},\frac{1}{2}), (2_{1[100]}|\frac{1}{2},0,0),$
$(m_{(100)}|\frac{1}{2},0,0), (2_{[010]}|\frac{1}{2},0,\frac{1}{2}),$
$(n_{(010)}|\frac{1}{2},0,\frac{1}{2}), (2_{[001]}|0,\frac{1}{2},0), (b_{(001)}|0,\frac{1}{2},0)\}$
$(2_{1[100]}|\frac{1}{2},\frac{1}{2},0)Pnam$
$= \{(2_{1[100]}|\frac{1}{2},\frac{1}{2},0), (b_{(100)}|\frac{1}{2},\frac{1}{2},0), (1|0,0,\frac{1}{2}),$
$(\bar{1}|0,0,\frac{1}{2}), (2_{[001]}|0,0,0),$
$(m_{(001)}|0,0,0), (2_{1[010]}|\frac{1}{2},\frac{1}{2},\frac{1}{2}), (n_{(010)}|\frac{1}{2},\frac{1}{2},\frac{1}{2})\}$
$(2_{1[100]}|\frac{1}{2},0,\frac{1}{2})Pnam$
$= \{(2_{1[100]}|\frac{1}{2},0,\frac{1}{2}), (c_{(100)}|\frac{1}{2},0,\frac{1}{2}),$
$(1|0,\frac{1}{2},0), (\bar{1}|0,\frac{1}{2},0), (2_{1[001]}|0,\frac{1}{2},\frac{1}{2}),$
$(b_{(001)}|0,\frac{1}{2},\frac{1}{2}), (2_{[010]}|\frac{1}{2},0,0), (a_{(010)}|\frac{1}{2},0,0)\}$

Here we have three antiphase boundaries $(1|0,\frac{1}{2},\frac{1}{2})$, $(1|0,0,\frac{1}{2})$ and $(1|0,\frac{1}{2},0)$.

The actual shape and extension of the domains are not determined by symmetry alone and are controlled by local elastic energy (*i.e.* bond energy) minimization. The weak energy of van der Waals bonds strongly suggests that most of the small perturbations in the arrangement of octahedra and molecular groups due to the domain boundaries occurs mainly through them.

It should be noticed that, in addition to these domain structures essentially due to the transition, there may also appear independently twins by pseudomerohedry because of the close values of a and b on the one hand and of b_m and c_m on the other.

References

GUYMONT, M. (1978). *Phys. Rev. B*, **18**, 5385-5393.
GUYMONT, M. (1981). *Phys. Rev. B*, **24**, 2647-2655.
GUYMONT, M., GRATIAS, D., PORTIER, R. & FAYARD, M. (1976). *Phys. Status Solidi A*, **38**, 629-636.
JIANG, W. J., ZHANG, Z., HEI, DZ. K. & KUO, K. H. (1985). *Phys. Status Solidi A*, **92**, K5-8.
MERESSE, A. & DAOUD, A. (1989). *Acta Cryst.* C**45**, 194-196.
ZANGAR, H., MIANE, J. L., COURSEILLE, C., CHANH, N. B., COUZI, M. & MLIK, Y. (1989). *Phys. Status Solidi A*, **115**, 107-115.

Acta Cryst. (1991). A**47**, 562-567

An Estimate of the Number of Structural Parameters Measurable from a Fiber Diffraction Pattern

BY LEE MAKOWSKI

Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, *USA*

(*Received* 6 *June* 1990; *accepted* 16 *April* 1991)

Abstract

An estimate of the number of independent structural parameters that can be determined from a fiber diffraction pattern is derived and its application is demonstrated. At resolutions where independent estimates can be made for the intensity of every layer line, this number is set by sampling limits along each layer line [Makowski (1982). *J. Appl. Cryst.* **15**, 546-557]. At resolutions beyond which separation of intensities due to individual layer lines is possible (the deconvolution limit) there may still be usable structural information in the pattern. Even though intensities on individual layer lines cannot be uniquely determined from these data, the data may still represent useful constraints on structural models of the diffracting particles. Here it is shown that beyond the deconvolution limit the total number of structural parameters obtainable increases linearly as a function of resolution.

Introduction

Every diffuse arc of intensity in a fiber diffraction pattern has the potential for acting as a constraint on models of the diffracting object. Quantitative measurement of the intensity of a reflection can be made to some limiting resolution dependent on the distribution of reflections in reciprocal space and the degree of disorientation in the specimen. The limiting resolution is set by the degree of overlap that can be corrected using a numerical deconvolution procedure (Makowski, 1978). Reflections falling beyond this deconvolution limit cannot be accurately measured because their overlap with neighbors due to disorientation cannot be corrected by any numerical procedure. Nevertheless, it is clear that in many cases substantial structural information exists in the diffracted intensity beyond this limit. Fig. 1 is a diffraction pattern from a fiber of the Pap adhesion pili from *E. coli*. With angular deconvolution (Makowski, 1978),

0108-7673/91/050562-06$03.00

it is possible to measure the intensities along layer lines in this pattern to about 30 Å spacing. However, there is strong discernible intensity in the diffraction pattern well beyond this spacing and intensity can be seen to extend beyond 3 Å spacing in stronger exposures. Currently, the intensity beyond the deconvolution limit is used only qualitatively to estimate the prevalence of various structural patterns in a fiber (*e.g.* Marvin, 1966; Henderson, 1975). Refinement of structural models directly against the optical densities on the film (rather than the intensities along layer lines) provides a potential method for utilizing the intensities beyond the deconvolution limit (*e.g.* Tibbitts, Caspar, Phillips & Goodenough, 1988). The design and testing of structural models refined against data extending beyond the deconvolution limit must take into account the number of independent intensity measurements that can be made in this region of the pattern. In this paper, an estimate of the number of independent structural parameters measurable from

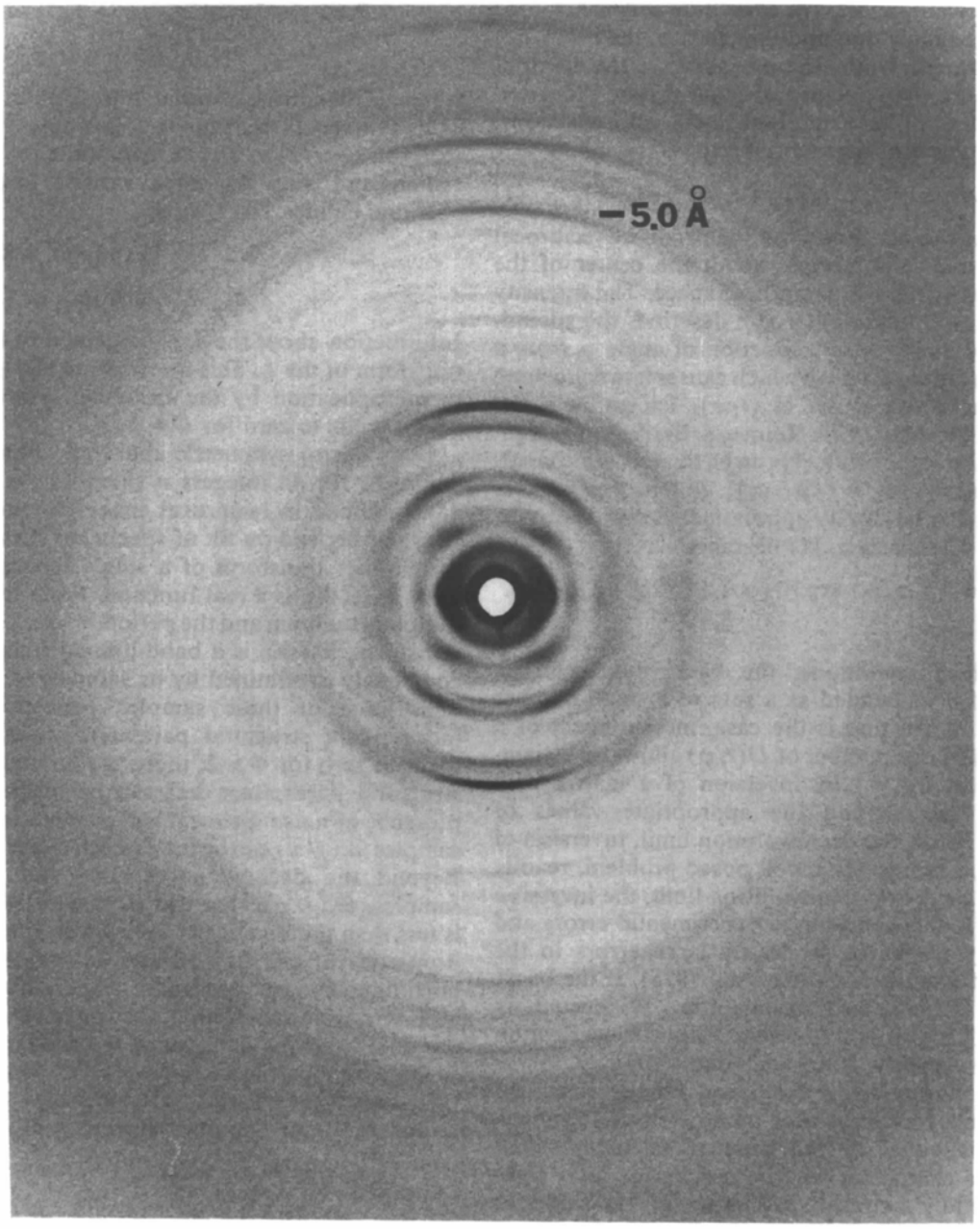

Fig. 1. Diffraction pattern from the Pap adhesion pili from *E. coli* (taken by M. Gong). The axial repeat for this specimen is 240·5 Å. The disorientation in this pattern corresponds to $\sigma = 6{\cdot}0°$.

a fiber diffraction pattern to any resolution is derived and its application demonstrated.

Analysis of intensities at a single radius

For simplicity, the treatment here is limited to diffraction from a partially oriented specimen of helical objects of diameter d and axial repeat c with random orientation about their long axes. Diffraction from a specimen of this type is limited to layer planes separated by a distance of $1/c$ in reciprocal space. These planes are observed as layer lines at their intersection with the Ewald sphere. The layer lines will be assumed to be vanishingly thin and the effect of the beam size will be ignored. Under these conditions, the distribution of intensity in reciprocal space can be expressed as the sum of contributions from all reflections, $I_i(r, \varphi_i)$, plus a background term,

$$D(r, \varphi) = \sum I_i(r, \varphi_i) f(\varphi, \varphi_i) + B(r, \varphi), \quad (1)$$

where r is the distance from the origin of reciprocal space and φ is the angle about the center of the diffraction pattern in reciprocal space. The intensity distribution function, $f(\varphi, \varphi_i)$, describes the spreading of the intensity as a function of angle φ from a reflection falling at (r, φ_i) which causes it to contribute to the measured signal at (r, φ). Except near the meridian (Stubbs, 1974; Holmes & Barrington-Leigh, 1974; Makowski, 1978), (1) can be treated as a convolution [*i.e.* $f(\varphi, \varphi_i) \rightarrow f(\varphi - \varphi_i)$] and in many cases $f(\varphi - \varphi_i)$ can be closely approximated by a Gaussian function. In that case, (1) becomes

$$D(r, \varphi) = \sum I_i(r, \varphi_i) \exp[-(\varphi - \varphi_i)^2/2\sigma^2] + B(r, \varphi). \quad (2)$$

Under most conditions, the background can be measured or expanded as a sum of a small number of terms. When that is the case, measurement of a sufficiently large number of $D(r, \varphi)$ allows an estimation of the $I_i(r, \varphi_i)$ by inversion of a matrix with elements representing the appropriate values of $f(\varphi, \varphi_i)$. Below the deconvolution limit, inversion of this matrix represents a well posed problem, readily solved. Above the deconvolution limit, the inversion becomes highly sensitive to experimental errors and ultimately results in arbitrarily large errors in the estimated intensities (Makowski, 1978). If the background is circularly symmetric or accurately measured, it will not affect the following derivation and, consequently, is not included.

At any radius below the deconvolution limit, the total number of structural parameters that can be derived is equal to the number of independent intensities at that radius (the number of layer lines intersected by a circle at that spacing). Intensities at adjacent radii may not be independent of one another and that issue is addressed below. Beyond the deconvolution limit, the total number of structural parameters that can be estimated will be less than the number of layer lines intersecting a circle at that spacing. To demonstrate this, (2) is rewritten for a radius, r, as (ignoring background)

$$D_r(\varphi) = \exp(-\varphi^2/2\sigma^2) * \sum I_i(r, \varphi_i)\delta(\varphi - \varphi_i) \quad (3)$$

where $*$ denotes convolution. If one takes $D_r(\varphi)$ to be defined for all φ, D_r is periodic with period π. If one defines the repeating unit of D_r to be in the interval $(-\pi/2, \pi/2)$, the Fourier transform, $\mathscr{D}_r(\Phi)$ of the repeating unit of D_r is real and equal to

$$\mathscr{D}_r(\Phi) = (2\pi\sigma^2)^{1/2}[\exp(-2\sigma^2\pi^2\Phi^2)] \times \mathscr{F}[\sum I_i(r, \varphi_i)\delta(\varphi - \varphi_i)], \quad (4)$$

where $\mathscr{F}$ denotes a Fourier transform and the sum extends over all reflections (layer lines) within the interval $(-\pi/2, \pi/2)$. Φ has units of reciprocal radians and is the reciprocal variable to φ.

If one defines $\Sigma = 1/2\pi\sigma$,

$$\mathscr{D}_r(\Phi) = (2\pi\Sigma^2)^{-1/2}[\exp(-\Phi^2/2\Sigma^2)] \times \mathscr{F}[\sum I_i(r, \varphi_i)\delta(\varphi - \varphi_i)]. \quad (5)$$

Information about the I_i is contained in the Fourier transform of the I_i. This information is degraded due to multiplication by the Gaussian, $\exp(-\Phi^2/2\Sigma^2)$ which tends to zero for $\Phi \gg \Sigma$.

D_r is mirror symmetric about all points at which $\varphi = n\pi/2$ for all integers n (meridian and equator; D_r is defined in reciprocal space and consequently does not depend on tilt of specimen). Consequently, the Fourier transform of a single repeating unit of $D_r(\varphi)$, $\mathscr{D}_r(\Phi)$, is a real function. From the Shannon sampling theorem and the period, π, of D_r, its Fourier transform, $\mathscr{D}_r(\Phi)$, is a band-limited function and is completely determined by its samples at intervals of $1/\pi$. Each of these samples corresponds to an independent structural parameter. Because $\mathscr{D}_r(\Phi)$ tends to zero for $\Phi \gg \Sigma$, there is a limited number of structural parameters that can be measured in the presence of noise. Below the deconvolution limit, the samples $\mathscr{D}_r(j/\pi)$ can be used to determine all $I_i(r, \varphi_i)$. Beyond the deconvolution limit, the number of samples, $\mathscr{D}_r(j/\pi)$, that can be accurately estimated is less than the number of layer lines. Analysis of the properties of $\mathscr{D}_r(\Phi)$ provides an estimate for the maximum number of structural parameters obtainable at any radius. The actual number depends on the quality of the data, since the fundamental limit is the amplification of noise due to division by the Gaussian, $\exp(-\Phi^2/2\Sigma^2)$. In particular, at any given radius, the layer-line intensities are derivable as

$$[\sum I_i(r, \varphi_i)\delta(\varphi - \varphi_i)] = \mathscr{F}[(2\pi\Sigma^2)^{1/2}\mathscr{D}_r(\Phi)/\exp(-\Phi^2/2\Sigma^2)]. \quad (6)$$

For large values of Φ, the noise amplification due to division by the Gaussian is limiting.

Let the largest factor by which noise can be reasonably amplified in the analysis of a diffraction pattern be ε. In (6), any noise in $\mathscr{D}_r(\Phi)$ is multiplied by

$$\varepsilon = 1/\exp(-\Phi^2/2\Sigma^2). \quad (7)$$

For any given choice of ε, the maximum number of structural parameters, N, obtainable at ANY radius in a fiber diffraction pattern is given by

$$\varepsilon = \exp(N^2/2\Sigma^2\pi^2). \quad (8)$$

Since there are (N/π) independent samples of $\mathscr{D}_r(\Phi)$ out to a limiting value of Φ, it follows that

$$N = [2\pi^2\Sigma^2 \ln(\varepsilon)]^{1/2} = [\ln(\varepsilon)/2]^{1/2}/\sigma. \quad (9)$$

If $f(\varphi)$ is not a Gaussian, then ε is related to $\mathscr{F}[f(\varphi)]$ in an obvious way. For a diffraction pattern with a standard deviation of particle orientation equal to σ $(=1/2\pi\Sigma)$, N is the maximum number of layer lines that can be separated using angular deconvolution. At values of r below the deconvolution limit, the number of structural parameters is equal to the number of layer lines intersecting a circle at radius r. Beyond the deconvolution limit, the number of structural parameters that can be estimated is a constant equal to N. This limit compares with the empirical limit estimated earlier (Makowski, 1978) of

$$N = 1/[\sin(1{\cdot}5\sigma)], \quad (10)$$

which, for small values of σ, is

$$N = 1/(1{\cdot}5\sigma) = 2\pi\Sigma/1{\cdot}5. \quad (11)$$

Equating the two estimates of N, we find that the earlier empirical estimate of N resulted in amplification of a small proportion of the noise in the data by a factor of $\varepsilon = 2{\cdot}25$. Fig. 2 is a plot of N as a function of σ. Fig. 3 shows an example of the analysis of data along a single radius (0·05 Å $^{-1}$) of a diffraction pattern from a specimen with axial repeat of 240·5 Å (as for the diffraction pattern in Fig. 1) for the cases of $\sigma = 0{\cdot}75$, 1·5, 3·0 and 6·0°. The density on this circle is assumed to be a set of Gaussians, each centered at the position of a layer line, as discussed above. The Fourier transforms of these distributions are also shown as an indication of the amount of information in each angular distribution. The extent of the Fourier transform is strongly dependent on the distribution of orientations in the specimen. Given that an independent structural parameter can be obtained from samples every $1/\pi$, (11) indicates that N is 51, 25, 13 and 6 for the four cases. Since the maximum number of layer lines that can contribute to this radius is 13, a complete estimate of all layer-line intensities should be possible for $\sigma < 3{\cdot}0°$. For disorientation with $\sigma > 3{\cdot}0°$, the number of independent parameters that can be estimated is less than the number of layer lines at this radius.

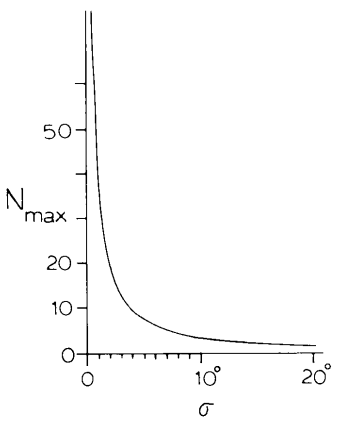

Fig. 2. The maximum number of structural parameters measurable at a single radius, $N(\sigma)$, as a function of the standard deviation of particle orientation in the diffracting specimen, calculated from equation (9) for $\varepsilon = 2{\cdot}25$. Since the deconvolution limit is proportional to N, r_{limit} is a very sensitive function of orientation for $\sigma < 5°$.

Number of structural parameters as a function of radius

Below the deconvolution limit, the total number of independent measurements in a fiber diffraction pattern can be roughly estimated by assuming that independent intensities occur at intervals of $1/2d$ along each layer line, where d is the diameter of the diffracting particle. This is an overestimate because there is no information in the near-meridional data for the distance from the meridian, $R < (n+2)/\pi d$ (Makowski, 1982), where n is the lowest order of Bessel function contributing to the layer line. To a radius of r less than the radius of the deconvolution limit (r_{limit}), the number of independent intensities estimated in this manner is

$$N(r < r_{limit}) = \pi c\, dr^2/2, \quad (12)$$

where c is the axial repeat of the diffracting particle. Equation (12) states the obvious result that, for a two-dimensional data set, the number of independent observations increases with the square of the resolution.

Beyond the resolution limit, the number of structural parameters observable at a given radius is a constant, so that $N(r > r_{limit})$ must increase linearly with resolution. Consequently, once the deconvolution limit is reached, the derivative of $N(r)$ is a constant and $N(r > r_{limit})$ may be written as

$$N(r > r_{limit}) = \pi c\, dr_{limit}^2/2 + \pi c\, dr_{limit}(r - r_{limit}). \quad (13)$$

Fig. 4 contains plots of $N(r)$ for diffraction patterns from tobacco mosaic virus and filamentous bacteriophage Pfl.

Discussion

In this paper, the first quantitative measure of the amount of information measurable beyond the deconvolution limit of a fiber diffraction pattern has been presented. This measure will be important in testing structural models of scattering units based on diffracted data extending to the highest possible resolution. It is important to point out here that, just because the number of structural parameters theoretically expected can be estimated, it does not necessarily mean that useful data will exist beyond the deconvolution limit. The intrinsic disorder in a molecular system can result in the fading of diffraction maxima as a function of resolution and the resolution beyond which there is no observable intensity is not directly related to the deconvolution limit. For instance, in the diffraction pattern from Pap pili in Fig. 1, the observable intensities extend well beyond the deconvolution limit. However, in most diffraction patterns from filamentous bacteriophage Pf1, the intensity of reflections falls off sharply beyond about 3·3 Å spacing. In the best-oriented specimens, the deconvolution limit is at about 3 Å spacing. Nevertheless, there is no means for extracting useful structural information beyond the point at which the intensities fade.

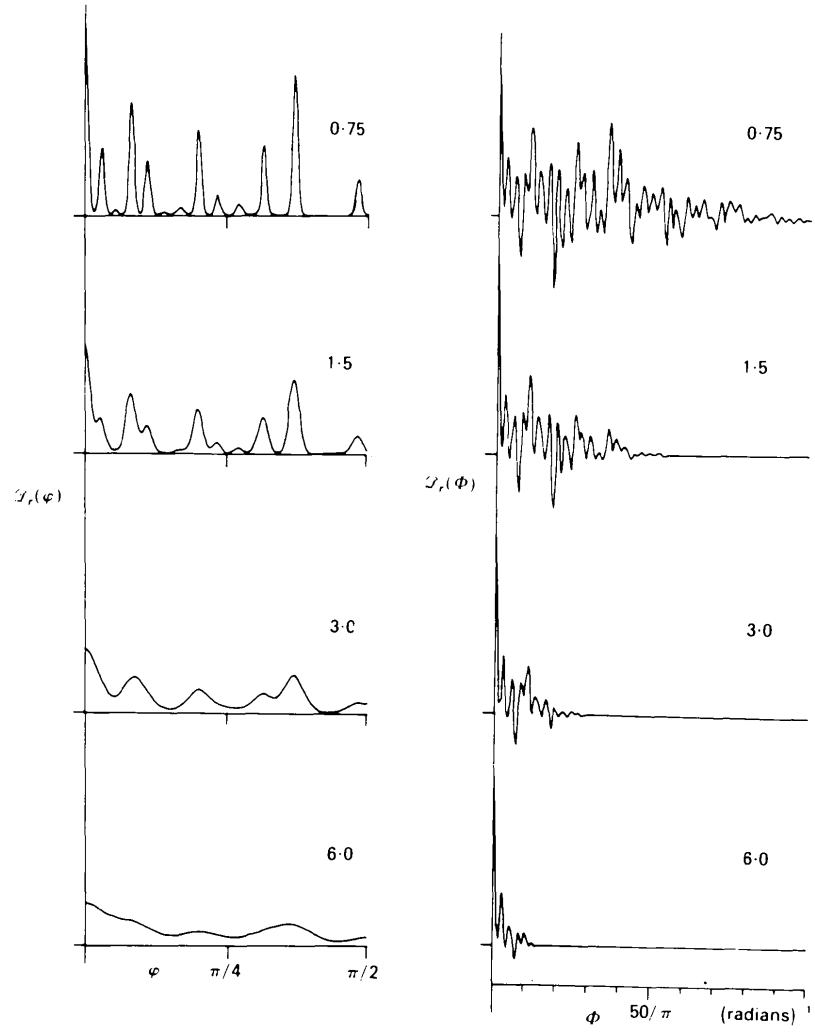

Fig. 3. Model density distributions for a single radius as a function of angle about the center of the diffraction pattern and their Fourier transforms. The calculation was done for a particle with the helical symmetry of Pap pili as determined from the diffraction pattern in Fig. 1 for a variety of disorientations. The plots on the left are the density distributions, as calculated with the standard deviation of orientation indicated. The plots on the right are the Fourier transforms of the density distributions. The theory presented in the text indicates that an independent structural parameter can be determined from each sample of these Fourier transforms, with the samples separated by $1/\pi$. The limited extent of the Fourier transforms of density distributions with $\sigma = 3{\cdot}0$ and $\sigma = 6{\cdot}0°$ is graphic demonstration of the effect of greater degree of disorientation on the information in the density distribution.

The number of structural parameters measurable from a fiber diffraction pattern ultimately depends on the choice of ε, the amount of noise amplification that can be tolerated in estimation of layer-line intensities from raw data. That number depends directly on the signal-to-noise ratio of the original data. In a previous analysis (Makowski, 1978), the intensities along layer lines were plotted as a function of distance from the meridian. Near the deconvolution limit, the noise in these estimated layer lines was greatly amplified. However, experience has indicated that this is not always the case and that systematic errors may be amplified and that this amplification does not vary greatly with radius - it does not result in an apparent increase in noise levels along a layer line. In the analysis of most fiber diffraction patterns using angular deconvolution, we have found that using a value of $\varepsilon = 3$ results in acceptable results, but this may not always be the case. A comparable value is likely to be appropriate for estimates of structural parameters beyond the deconvolution limit.

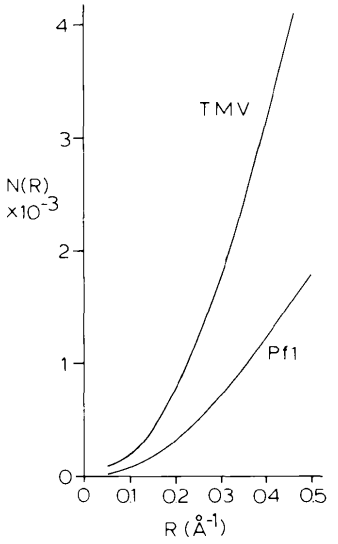

Fig. 4. The total number of structural parameters derivable from fiber diffraction patterns from tobacco mosaic virus (TMV) and filamentous bacteriophage Pf1 as a function of resolution. For this calculation, σ was set to 1·4° for both specimens. The larger number of parameters measurable for TMV is largely a result of its larger diameter (180 Å compared to 65 Å for Pf1).

The analysis presented here does not directly address the question of how to use the measurable structural parameters beyond the deconvolution limit. Equation (6) represents an indeterminant set of equations for the layer-line intensities beyond the deconvolution limit. The measurable structural parameters represent a partial set of data, from which intensities cannot be completely derived. However, they also represent a set of constraints on the structure of a helical assembly that with proper statistical tests can be used to constrain or refine (Tibbitts *et al.*, 1988) the structure of a diffracting unit. The rapidly increasing availability of computing power will provide the means for using these data. This paper provides a quantitative estimate of the amount of information contained in the observable data beyond the deconvolution limit.

The author thanks M. Gong for the use of the diffraction pattern in Fig. 1. This work is supported by a grant from the National Science Foundation.

References

HENDERSON, R. (1975). *J. Mol. Biol.* **93**, 123-138.
HOLMES, K. & BARRINGTON-LEIGH, J. (1974). *Acta Cryst.* A**30**, 635-638.
MAKOWSKI, L. (1978). *J. Appl. Cryst.* **11**, 273-281.
MAKOWSKI, L. (1982). *J. Appl. Cryst.* **15**, 546-557.
MARVIN, D. A. (1966). *J. Mol. Biol.* **15**, 8-17.
STUBBS, G. (1974). *Acta Cryst.* A**30**, 639-644.
TIBBITTS, T. T., CASPAR, D. L. D., PHILLIPS, W. C. & GOODENOUGH, D. A. (1988). *Biophys. J.* **53**, 632a.

Acta Cryst. (1991). A**47**, 567-571

An Atomicity-Constrained Least-Squares Phase-Refinement Method

BY JORDI RIUS AND CARLES MIRAVITLLES

Institut de Ciència de Materials (CSIC), Campus Universitari de Bellaterra, 08193 *Cerdanyola, Barcelona, Spain*

(*Received* 9 *April* 1989; *accepted* 17 *April* 1991)

Abstract

A phase-refinement method based on Sayre's squaring equation and on a suggestion made by Hoppe [*Z. Kristallogr.* (1963), **118**, 121-126] is presented. It takes advantage of a lot of information initially known about a crystal structure such as all the measured structure-factor magnitudes and the atomicity constraint implicit in Sayre's equation. The method assumes that any squared structure factor $|\mathbf{F}(\mathbf{H})|^2$ can

0108-7673/91/050567-05$03.00