LARSEN, S. E., KILPATRICK, J. E., LIEB, E. H. & JORDAN, H. F. (1965). *Phys. Rev. A*, **140**, 1129-1130.
LEVINE, R. D. (1980). *J. Phys. A: Math. Nucl. Gen.* **13**, 91-108.
MCQUARRIE, D. A. (1973). *Statistical Thermodynamics.* Mill Valley, California: Univ. Science Books.
MAIN, P. (1976). In *Crystallographic Computing Techniques*, edited by F. R. AHMED. Copenhagen: Munksgaard.
MAIN, P. & ROSSMANN, M. G. (1966). *Acta Cryst.* **21**, 67-72.
MAJOR, F., GUATHERET, D., LAPALME, G., JOLICOEUR, L., FILLION, E. & CEDERGREN, R. (1990). In PhD thesis (François Major), Univ de Montréal, Canada.
MANDELBROT, B. B. (1989). *Phys. Today*, **42**, 71-73.
MARGENAU, H. & KESTNER, N. R. (1969). *Theory of Intermolecular Forces.* New York: Pergaman.
MAYER, J. E. & MAYER, M. G. (1940). *Statistical Mechanics.* New York: Wiley.
MISTRIOTIS, A. D., FLYTZANIS, N. & FARANTOS, S. C. (1989). *Phys. Rev. B*, **39**(2), 1212-1218.
MOHANTY, U. (1982). *J. Math. Phys.* **23**, 1489-1492.
MOHLING, F. (1982). *Statistical Mechanics: Methods and Applications.* Publishers Creative Services Inc.
MORITA, T. & HIROIKE, K. (1961). *Prog. Theor. Phys.* **25**, 537-543.
NAVAZA, J. (1985). *Acta Cryst.* A**41**, 232-244.
NAVAZA, J. (1986). *Acta Cryst.* A**33**, 212-223.
PARR, R. G. (1983). *Ann. Rev. Phys. Chem.* **34**, 631-656.
PARR, R. G. & WEITAO, (1989). *Density-Functional Theory of Atoms and Molecules.* Oxford Univ. Press.
PESCHAR, R. & SCHENK, H. (1986). *Acta Cryst.* A**42**, 309-317.
RABINOVICH, D. & SHAKKED, Z. (1984). *Acta Cryst.* A**40**, 195-200.
REE, F. H. & HOOVER, W. G. (1967). *J. Chem. Phys.* **46**, 4181-4196.
RIUS, J. & MIRAVITLLES, C. (1991). *Acta Cryst.* A**47**, 52-55.
RIVERS, R. J. (1987). *Path Integral Methods in Quantum Field Theory.* Cambridge Univ. Press.
ROMAN, S. (1979*a*). *J. Approx. Theory*, **26**, 340-381.
ROMAN, S. (1979*b*). *Adv. Math.* **31**, 309-329.
SEMENOVSKAYA, S. V., KHACHATURYAN, K. A. & KHACHATURYAN, A. G. (1981). *Acta Cryst.* A**37**, 742-754.
SEMENOVSKAYA, S. V., KHACHATURYAN, K. A. & KHACHATURYAN, A. G. (1985). *Acta Cryst.* A**41**, 268-273.
SERFLING, R. J. (1968). *Ann. Math. Stat.* **39**, 1158-1175.
SHELDRICK, G. (1990). *Acta Cryst.* A**46**, 467-473.
SHMUELI, U., RABINOVICH, S. & WEISS, G. H. (1989). *Acta Cryst.* A**45**, 361-367.
STELL, G. (1976). In *Phase Transitions and Critical Phenomena*, Vol. 5B, edited by C. DOMB & M. S. GREEN. London: Academic Press.
STILLINGER, F. H. & BUFF, J. (1962). *J. Chem. Phys.* **37**, 1-12.
STOUT, G. H. & JENSEN, L. H. (1989). *X-ray Structure Determination: a Practical Guide*, 2nd ed. New York: John Wiley.
TERSOFF, J. (1986). *Phys. Rev. Lett.* **55**(6), 632-635.
WILSON, A. J. C. (1981). *Acta Cryst.* A**37**, 808-810.
WILSON, B. G. & ROGERS, F. J. (1986). *Physica (Utrecht)*, A**139**, 359-386.
YVON, J. (1935). *Actual. Sci. Ind.* **203**, 67-72.
ZUBAREV, D. N. (1974). *Nonequilibrium Statistical Thermodynamics.* New York: Consultants Bureau.

Acta Cryst. (1992). A**48**, 209-215

R Factors in X-ray Fiber Diffraction. IV. Analytic Expressions for Largest Likely *R* Factors

BY R. P. MILLANE

Whistler Center for Carbohydrate Research, Smith Hall, Purdue University, West Lafayette, Indiana 47907, *USA*

(*Received* 14 *November* 1990; *accepted* 19 *August* 1991)

Abstract

The largest likely R factor is useful for evaluating the significance of R factors obtained in structure determinations. Numerical expressions have been derived previously for calculating largest likely R factors in fiber diffraction analyses. Analytical approximations to largest likely R factors ($\mathcal{R}$) in fiber diffraction are derived here that show the dependence on resolution ($\hat{\rho}$), helix symmetry (u_v) and molecular radius ($\hat{r}$). The simplest approximation is $\mathcal{R} \propto (u/\hat{r}\hat{\rho})^{1/2}$ which represents the overall behavior of R factors reasonably well. More accurate approximations are also derived. These are applied to various structures and the dependence on different structural parameters is examined. These results provide insight into the behavior of R factors in fiber diffraction and may be useful in further analysis.

1. Introduction

The significance of an R factor obtained in a structure determination can be assessed by comparison with the largest likely R factor; that for a structure uncorrelated with the correct structure (Wilson, 1950). R factors in fiber diffraction are generally smaller than in single-crystal analyses because the diffraction pattern is cylindrically averaged, and the largest likely R factor in fiber diffraction has been studied by Stubbs (1989) and Millane (1989*a*, *b*, 1990). The largest likely R factor depends on the number of overlapping complex Fourier-Bessel structure factors at different positions in reciprocal space, and therefore on the diameter and symmetry of the diffracting particle and the maximum resolution of the diffraction data. The largest likely R factor in fiber diffraction, while easily calculated, is a rather complicated

0108-7673/92/020209-07$03.00

expression which makes its dependence on symmetry and resolution obscure (Millane, 1989*b*). Simple approximate expressions for largest likely R factors in fiber diffraction as a function of parameters such as molecular symmetry and diffraction data resolution are derived here. The utility and accuracy of these expressions are illustrated by applications to typical structures and the general behavior of R factors examined.

Previous results for largest likely R factors in fiber diffraction are recalled in § 2, and explicit approximations to these expressions are derived in § 3. In § 4, the accuracy of these expressions is examined by applying them to particular structures, and their general behavior discussed. Concluding remarks are made in the final section.

2. Preamble

The largest likely R factor (denoted by $\mathcal{R}$ to distinguish it from the cylindrical radius R in reciprocal space) for a fiber diffraction analysis is given by (Millane, 1989*b*)

$$R = \sum_{m=1}^{M} N_m \mathcal{R}_m S_m \Big/ \sum_{m=1}^{M} N_m S_m \qquad (1)$$

where the sums are over the different numbers of overlapping complex Fourier–Bessel structure factors (both real and imaginary parts) G_{nl} (Klug, Crick & Wyckoff, 1958) that contribute to the different intensity measurements. N_m of the intensity measurements have m overlapping terms and M is the maximum value of m on the diffraction pattern. This can be applied to noncrystalline or polycrystalline specimens (Millane, 1989*b*) although in this paper only noncrystalline specimens are addressed. The $\mathcal{R}_m$ are the largest likely R factors for m overlapping terms (Stubbs, 1989) and are given by (Millane, 1989*a*)

$$\mathcal{R}_m = 2 - 2^{-m+2} m \binom{2m-1}{m} B_{1/2}[(m/2)+1/2, m/2] \qquad (2)$$

where $\binom{m}{n}$ is the binomial coefficient and $B_x(m, n)$ is the incomplete beta function. Note that there is a minus sign missing in equation (6) of Millane (1989*b*). The S_m are proportional to the mean values of the amplitudes that contain m overlapping terms and are given by (Millane, 1989*b*)

$$S_m = \Gamma[(m/2)+1/2]/\Gamma(m/2) \qquad (3)$$

where $\Gamma(x)$ is the gamma function. For a particular diffraction pattern, the N_m can be easily determined and the largest likely R factor calculated using tabulated values of $\mathcal{R}_m$ and S_m (Millane, 1989*b*).

Asymptotic analysis of the functions $\mathcal{R}_m$ and S_m as a function of m (for $m \to \infty$) allows the following approximations (that are quite accurate) to be derived (Millane, 1990):

$$\mathcal{R}_m \simeq (2/\pi)^{1/2} m^{-1/2} \qquad (4)$$

$$S_m \simeq 2^{-1/2} m^{1/2} \qquad (5)$$

$$\mathcal{R}_m S_m \simeq \pi^{-1/2}. \qquad (6)$$

These expressions in (1) give largest likely R factors that are accurate to 0.01 for typical fiber structures (Millane, 1990).

3. Theory

(*a*) *Formulation*

Consider a molecule with u_v helix symmetry (u repeat units in v turns), maximum radius $\hat{r}$ and c repeat c and consider diffraction data measured between minimum and maximum resolution limits (reciprocal-space radii) $\check{\rho}$ and $\hat{\rho}$, respectively. The number of Fourier–Bessel structure factors that effectively contribute to the diffracted intensity at a position in reciprocal space depends on the behavior of the Bessel functions $J_n(2\pi Rr)$ where R and r are the cylindrical radii in reciprocal space and real space, respectively (Makowski, 1982). The Bessel function $J_n(2\pi Rr)$ is considered to make a significant contribution only for

$$|n| < 2\pi R\hat{r} + p \qquad (7)$$

where p is a parameter (this is of course an approximation and is discussed further in § 5). The value of p is usually taken, somewhat arbitrarily, to be 2 (*e.g.* Makowski, 1982), but is considered variable here to examine its effect on the results.

The largest likely R factor depends in general on the parameters u, v, $\hat{r}$, c, $\check{\rho}$, $\hat{\rho}$, p, although the approximations derived here are independent of some of these parameters. It is convenient to consider separately the case where there is effectively no lower limit on the resolution of the diffraction data ($\check{\rho} = 0$) and this is denoted by $\mathcal{R}(\check{\rho} = 0)$.

The summations in (1) over m actually represent a summation over the diffraction pattern (Millane, 1989*b*) and the analysis is conveniently carried out using the latter description. Equation (1) can therefore be written as

$$\mathcal{R} = \frac{\sum_{l=0}^{L} \int_{\check{R}(l)}^{\hat{R}(l)} \mathcal{R}(m(R,l)) S(m(R,l))\, \mathrm{d}R}{\sum_{l=0}^{L} \int_{\check{R}(l)}^{\hat{R}(l)} S(m(R,l))\, \mathrm{d}R} \qquad (8)$$

where $\mathcal{R}_m$ and S_m are replaced by $\mathcal{R}(m)$ and $S(m)$ respectively, $m(R, l)$ denotes the value of m on layer line l at reciprocal-space radius R, $\check{R}(l)$ and $\hat{R}(l)$ are the minimum and maximum values, respectively, of R on layer line l between the resolution limits, and L is the maximum value of l. In order to make the

analysis tractable, the summations in (8) are approximated by integrals and the approximations (5) and (6) are used, giving

$$\mathscr{R} = \left(\frac{2}{\pi}\right)^{1/2} \frac{\int_0^{\hat{\rho}} \int_{\check{R}(Z)}^{(\hat{\rho}^2-Z^2)^{1/2}} U(m(R,Z))\,\mathrm{d}R\,\mathrm{d}Z}{\int_0^{\hat{\rho}} \int_{\check{R}(Z)}^{(\hat{\rho}^2-Z^2)^{1/2}} m^{1/2}(R,Z)\,\mathrm{d}R\,\mathrm{d}Z}$$

$$= (2/\pi)^{1/2} A/B \qquad (9)$$

where $U(m)=0$ for $m=0$ and $U(m)=1$ for $m>0$ and the meaning of the other symbols is obvious, with l replaced by Z. A and B denote the numerator and denominator respectively in (9).

Equation (9) provides a convenient formulation for deriving an analytic expression for the largest likely R factor. This requires an analytic expression for $m(R, Z)$ which is derived in the next subsection, and expressions for $\mathscr{R}$ are derived in the two following subsections.

(b) *Expressions for* $m(R, Z)$

The value of m is equal to twice the number of Fourier-Bessel structure factors that contribute on layer line l at reciprocal-space radius R. It depends therefore on the relevant behavior of the Bessel functions given by (7) and the selection rule

$$l = um + vn \qquad (10)$$

that determines which Fourier-Bessel terms $G_{nl}(R)$ are non-zero on layer line l, where m and n are integers. On a particular layer line, solutions to (10) are given by (Makowski, 1982)

$$n = n_1 \pm mu \qquad (11)$$

where n_1 is any solution to (10). It is convenient to take for n_1 the smallest value of $|n|$, and then inspection of (11) shows that n_1 satisfies the inequality

$$0 \le n_1 \le u/2. \qquad (12)$$

Equation (10) shows that n_1 takes on all integral values between 0 and $u/2$ for any set of layer lines $\{l, l+1, \ldots, l+u/2\}$. Therefore, the average value of n_1 over a set of consecutive layer lines is $\langle n_1 \rangle = u/4$. Inspection of (11) shows that the average difference Δn between consecutive values of $|n|$ on a particular layer line is $u/2$.

From (7), m is a step function of R (Fig. 1). For the purposes of evaluating (9), it is convenient to approximate m as a linear function of R, and the above paragraph and Fig. 1 show that the appropriate approximation is

$$m(R, Z) = m_0 + (8\pi\hat{r}/u)R \qquad (13)$$

where $m_0 = 1 - 4(n_1 - p)/u$. Since m_0 depends on n_1 and $\langle n_1 \rangle = u/4$, the average value of m_0 is $4p/u$, so that (13) can be written as

$$m(R, Z) = (4p/u) + (8\pi\hat{r}/u)R \qquad (14)$$

which is an approximation for m that can be used in (9).

On the equator, the $G_{nl}(R)$ satisfy the relationship $G_{-n0}(R) = (-1)^n G^*_{n0}(R)$, so that $G_{n0}(R)$ and $G_{-n0}(R)$ do not contain independent information. Hence, on the equator (where both n and $-n$ satisfy the selection rule), m is equal to half that given by (14).

(c) *First approximation to the largest likely R factor*

For the case $\check{\rho} = 0$, the numerator in (9) is given by

$$A = \pi\hat{\rho}^2/4. \qquad (15)$$

Evaluation of the denominator in (9) is difficult unless the constant term in (14) is neglected, in which case

$$B = (8\pi\hat{r}/u)^{1/2} \int_0^{\hat{\rho}} \int_0^{(\hat{\rho}^2-Z^2)^{1/2}} R^{1/2}\,\mathrm{d}R\,\mathrm{d}Z \qquad (16)$$

and evaluating the R integral gives

$$B = (2/3)(8\pi\hat{r}/u)^{1/2} \int_0^{\hat{\rho}} (\hat{\rho}^2 - Z^2)^{3/4}\,\mathrm{d}Z. \qquad (17)$$

This can be evaluated in terms of the beta function (Abramowitz & Stegun, 1972, equation 6.2.1), and thence in terms of the gamma function (Abramowitz & Stegun, 1972, equation 6.2.2) giving

$$B = (8/5)[\Gamma(3/4)]^2 (\hat{r}/u)^{1/2} \hat{\rho}^{5/2}. \qquad (18)$$

Substituting (15) and (18) into (9) gives an approximation to the largest likely R factor as

$$\mathscr{R}(\check{\rho} = 0) = k(u/\hat{r}\hat{\rho})^{1/2} \qquad (19)$$

where

$$k = (5\pi^{1/2})/\{16 \times 2^{1/2}[\Gamma(3/4)]^2\} = 0.2608. \qquad (20)$$

A similar analysis with $\check{\rho} \neq 0$ is straightforward and gives

$$\mathscr{R} = k(u/\hat{r}\hat{\rho})^{1/2}(1-x^2)(1-x^{5/2})^{-1} \qquad (21)$$

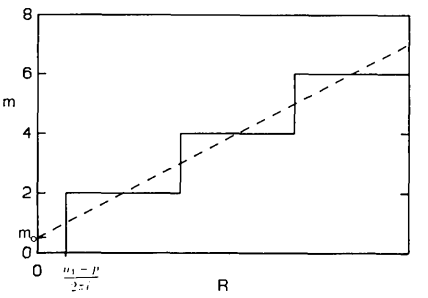

Fig. 1. The number of Fourier-Bessel terms (real and imaginary parts) m contributing to the layer-line intensity as a function of cylindrical radius R (——) and the linear approximation $m(R, Z)$ (- - -) derived here.

where

$$x = \check{\rho}/\hat{\rho}. \quad (22)$$

The approximations (19) and (21) are compared in Fig. 2 (dashed lines) with R factors calculated exactly (solid lines) using (1) (Millane, 1989*b*) for a hypothetical structure with $c = 20$ Å, $\hat{r} = 10$ Å, with 10_1 and 5_1 helix symmetry, as a function of maximum resolution $\hat{\rho}$, and for two minimum resolutions $\check{\rho}$. In all the examples $p = 2$ unless stated otherwise. The approximations mimic the general behavior of the R factors quite well and are reasonably accurate at high resolution but are not particularly accurate at low resolution.

(*d*) *Second approximation to the largest likely R factor*

The approximations to the R factors derived above can be improved by using a more accurate evaluation of B in (9). This is done firstly by considering the constant term $4p/u$ in (14) that was neglected in the above analysis. In order to do this but keep the integral for B tractable, $m(R, Z)$ is approximated by (see Appendix)

$$m(R, Z) = (8\pi\hat{r}/u)[1 + 6p/(\pi^2\hat{r}\hat{\rho})]R. \quad (23)$$

Secondly, as described above, the value of m on the equator is half that given by (14). The proportion of the total measurements contributed by the equator (for $\check{\rho} = 0$) is $4/\pi c\hat{\rho}$ so that, referring to (9), this effect can be approximately corrected for by multiplying B by the factor f where

$$f = [1 - 2/(\pi c\hat{\rho})]^{1/2}. \quad (24)$$

Incorporating (23) and (24) into the evaluation of B as above gives the improved approximation

$$\mathcal{R}(\check{\rho} = 0) = k(u/\hat{r}\hat{\rho})^{1/2}[1 + 6p/(\pi^2\hat{r}\hat{\rho})]^{-1/2} \times [1 - 2/(\pi c\hat{\rho})]^{-1/2}. \quad (25)$$

A similar analysis is performed for $\check{\rho} \neq 0$ and, with reference to the Appendix, this gives

$$m(R, Z) = (8\pi\hat{r}/u)[1 + 6p(1 + x^2)/(\pi^2\hat{r}\hat{\rho})]R \quad (26)$$

which allows the largest likely R factor to be approximated by

$$\mathcal{R} = k(u/\hat{r}\hat{\rho})^{1/2}[1 + 6p(1 + x^2)/(\pi^2\hat{r}\hat{\rho})]^{-1/2} \times \{1 - 2/[\pi c\hat{\rho}(1 + x)]\}^{-1/2}(1 - x^2)(1 - x^{5/2})^{-1}. \quad (27)$$

The approximations (25) and (27) are compared in Fig. 2 (dotted lines) with the first approximations derived above, and are seen to be significant improvements over the former. They are accurate to 0.01 for $\check{\rho} = 0$ in these examples, which is sufficient accuracy in practice. For larger values of $\check{\rho}$, they are not quite as accurate at low resolution.

4. Examples and discussion

The general behavior of the largest likely R factor with resolution is illustrated in Fig. 2. The dependence on other parameters is investigated here and the accuracies of the approximations are assessed for some actual structures.

The behavior of the largest likely R factor with different helix symmetries u_1 is shown by the solid line in Fig. 3, and compared with calculations made using the two approximations derived above. Both approximations derived here mimic the general behavior, the second approximation being quite accurate. The first approximation is poorer for larger u since the approximation for $m(R, Z)$ used is less accurate where the first Fourier–Bessel term n_1 begins to contribute at larger distances from the meridian.

The variation of the R factor with v, for u_v helix symmetry, is expected to be small because the effect of different values of v averages out over a sufficient number of layer lines. This is shown to indeed be the case in Fig. 4, where the R factor can be compared

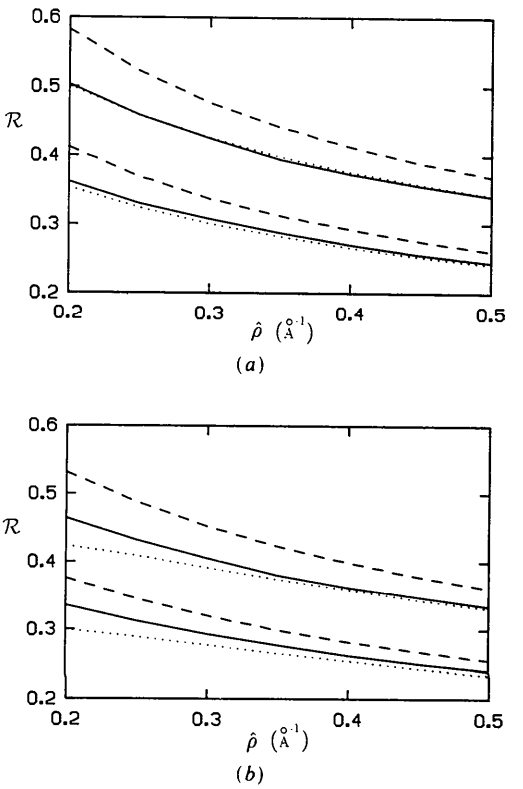

Fig. 2. (*a*) Variation of the largest likely R factor (for a structure with radius $\hat{r} = 10$ Å and c repeat 20 Å, for $\check{\rho} = 0$) with maximum resolution (——). The first approximations [(19) and (21)] are shown by the dashed lines and the second approximations [(25) and (27)] by the dotted lines. The upper curves are for 10_1 helix symmetry and the lower curves for 5_1 helix symmetry. (*b*) The same as (*a*) except that the minimum resolution is 10 Å ($\check{\rho} = 0.1$ Å^{-1}).

with the (v independent) approximations derived here.

The effect of different values of p on the R factor is shown in Fig. 5 for $1 < p < 3$. The first approximation is independent of p and the second approximation reproduces the dependence on p quite well. The value of p that is appropriate depends on what value of $J_n(x)$ is considered significant in contributing to the diffracted intensity (this value is smaller for larger p). This would depend in turn on the accuracy with which $I_1(R)$ is measured. The error level $E(p)$ introduced by neglecting the Fourier–Bessel terms based on (7) could be very approximately quantified by

$$E(p) = \max_n \{J_n(n-p)\} \tag{28}$$

which can be considered a relative (or normalized) error since $\max_{n,x}\{J_n(x)\} = J_0(0) = 1$. This error level is shown in Table 1 for the above range of values of p. Equating this to a probable (relative) precision of intensity measurement of, say, between 5 and 10%, a value of p between 1 and 3 would appear to be realistic. Fig. 5 shows that the choice of p within this range has minimal significant effect ($\Delta\mathcal{R} < 0.02$) on the calculated largest likely R factors.

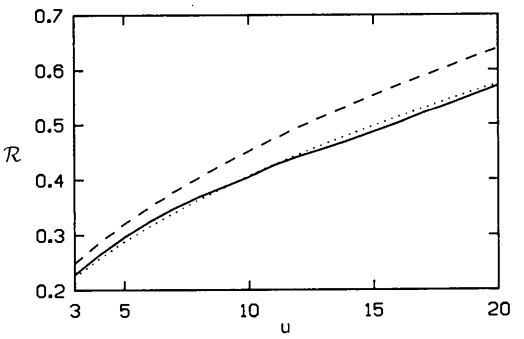

Fig. 3. Variation of the largest likely R factor with helix symmetry u_1 for a structure with $\hat{r} = 10$ Å and $c = 20$ Å, $\check{\rho} = 0$ and a maximum resolution of 3 Å (——). The first approximation (19) is shown by the dashed line and the second approximation (25) by the dotted line.

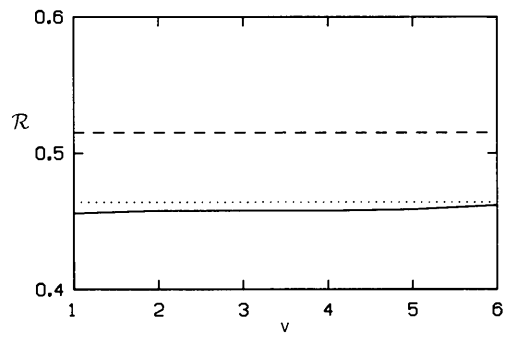

Fig. 4. Variation of the largest likely R factor for a structure with $\hat{r} = 10$ Å and $c = 20$ Å, $\check{\rho} = 0$ and a maximum resolution of 3 Å with v for 13_v helix symmetry (——). The first approximation (19) is shown by the dashed line and the second approximation (25) by the dotted line.

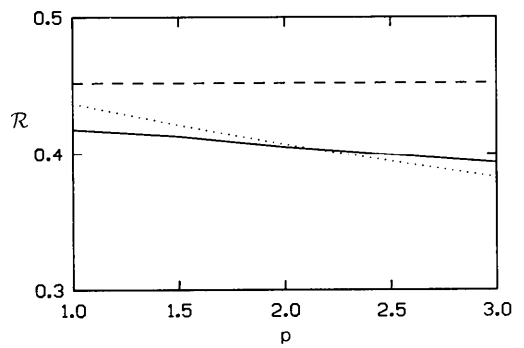

Fig. 5. Variation of the largest likely R factor for a structure with $\hat{r} = 10$ Å and $c = 20$ Å, $\check{\rho} = 0$ and a maximum resolution of 3 Å with p for 10_1 helix symmetry (——). The first approximation (19) is shown by the dashed line and the second approximation (25) by the dotted line.

Table 1. *Approximate error levels $E(p)$ resulting from neglecting Fourier–Bessel structure factors for which $|n| > 2\pi Rr - p$*

p	$E(p)$
1.0	0.132
1.5	0.090
2.0	0.068
2.5	0.054
3.0	0.045

The approximate expressions derived here were used to calculate largest likely R factors for some typical fiber structures and are compared with the values calculated exactly (Millane, 1989*b*). The structures used were a nucleic acid (Park, Arnott, Chandrasekaran, Millane & Campagnari, 1987), the helical virus TMV (Namba & Stubbs, 1985) and the bacteriophage Pf1 (Stark, Glucksman & Makowski, 1988). The first approximation (denoted by $\mathcal{R}_1$) given by (19) or (21), and the second approximation ($\mathcal{R}_2$) given by (25) or (27) are compared with the exact values ($\mathcal{R}$) in Table 2. These results show that the first approximation gives a useful guide to largest likely R factors although the errors can be as large as 0.05. The second approximation however is accurate to 0.01 in all cases, which is sufficient accuracy in practice.

It should be noted that the analysis described by Millane (1989*b*) assumed that the terms $G_{nl}(R)$ and $G_{-nl}(R)$ contain the same information, and so do not contribute separately to m. This is in fact not correct, except on the equator as described in § 3(*b*) above. This error had two effects. The actual R factors calculated there were slightly too large, although comparison of Table 2 of Millane (1989*b*) with Table 2 here shows that this is of little significance. For small values of u, the R factor alternated between two curves for u odd and even [Fig. 2(*b*) of Millane,

Table 2. *Exact ($\mathcal{R}$) and approximate ($\mathcal{R}_1$ and $\mathcal{R}_2$) largest likely R factors for three structures*

Molecule	Helix symmetry	Maximum radius (Å)	Minimum resolution (Å)	Maximum resolution (Å)	$\mathcal{R}$	$\mathcal{R}_1$	$\mathcal{R}_2$
DNA	10_1	10.0	∞	3.0	0.401	0.452	0.399
DNA	10_1	10.0	∞	2.5	0.374	0.412	0.370
TMV	49_3	90.0	10.0	5.0	0.373	0.392	0.382
TMV	49_3	90.0	10.0	3.0	0.307	0.319	0.316
Pfl	27_5	30.0	10.0	5.0	0.458	0.504	0.457
Pfl	27_5	30.0	10.0	3.0	0.381	0.410	0.389

$\mathcal{R}_1$ is calculated using equation (19) or (21), and $\mathcal{R}_2$ using equation (25) or (27).

1989*b*], whereas no such alternation is actually present (Fig. 3 here).

5. Concluding remarks

Analytical expressions have been derived that show quite clearly the dependence of the largest likely R factor in fiber diffraction on resolution, helix symmetry and molecular radius. These apply to noncrystalline specimens only - for polycrystalline specimens (9) is replaced by a sum over the Bragg reflections, and expressions such as (14) for m are not appropriate. The simplest expressions (19) and (21) show the main dependence on these parameters and, although they do not give accurate values in all cases, they represent the overall behavior quite well. More accurate expressions (25) and (27) have been derived in the form of corrections to the simpler expressions. These give quite accurate values in the cases examined here, and describe the behavior in terms of the different parameters very well. The utility of these results is that they provide insight into the behavior of fiber diffraction R factors, and the simple form of the expressions obtained may make them useful in further analysis. Although largest likely R factors can be calculated straightforwardly numerically (Millane, 1989*b*), calculation using (25) or (27) is much simpler and should be sufficiently accurate in most cases. The approximate analytical expressions for $m(R, Z)$ derived here may also be useful in other areas of fiber diffraction theory.

It is common practice to use (7) with $p = 2$ as the basis for determining which Fourier-Bessel terms contribute significantly to the diffracted intensity. Although this is a reasonable approach, it is clearly an approximation that is worthy of further study. The results in Table 1 for different values of p provide a little information in this regard. There is no difficulty when simply calculating diffracted intensities since inclusion of more terms than necessary does not affect the calculation. However, when considering the number of *significant* terms, the cutoff for n should be such that the error introduced by excluding the higher-order terms is constant over the diffraction pattern. It is not clear that a cutoff based on (7) achieves this.

I am grateful to the US National Science Foundation for support (DMB-8916477) and Deb Zerth for word processing.

APPENDIX
Approximation to $m(R, Z)$

The intent here is to derive a linear approximation to $m(R, Z)$ given by (14) that has no constant term, but gives an approximation to B in (9) that is better than that given by simply ignoring the constant term in (14). Since the approximation to $m(R, Z)$ used is independent of Z, consider the rectangular region $0 < Z < \hat{\rho}, 0 < R < C$, where C is a constant, rather than the quadrant $0 < \rho < \hat{\rho}$. For the case $\check{\rho} = 0$, requiring the areas of these two regions to be equal gives

$$C = \pi\hat{\rho}/4. \qquad (A1)$$

This allows the expression for B to be reduced to a single integral over R, and denoting (14) by $m = \alpha + \beta R$ and the required approximation by $m = \gamma R$, the value of γ is determined by requiring that

$$\int_0^C (\gamma R)^{1/2}\,\mathrm{d}R = \int_0^C (\alpha + \beta R)^{1/2}\,\mathrm{d}R. \qquad (A2)$$

Evaluating (*A*2) gives

$$\gamma = \beta[(1 + \alpha/\beta C)^{3/2} - (\alpha/\beta C)^{3/2}]^2 \qquad (A3)$$

and expanding to first order in α gives

$$\gamma = \beta(1 + 3\alpha/\beta C). \qquad (A4)$$

Substituting for α, β and C gives equation (23) for $m(R, Z)$.

For the case $\check{\rho} \neq 0$, the analysis is identical except that

$$C = \pi\hat{\rho}[1 - (\check{\rho}/\hat{\rho})^2]/4 \qquad (A5)$$

and, for $\check{\rho} \ll \hat{\rho}$, using this gives (26).

References

ABRAMOWITZ, M. & STEGUN, I. A. (1972). *Handbook of Mathematical Functions.* New York: Dover.
KLUG, A., CRICK, F. H. C. & WYCKOFF, H. W. (1958). *Acta Cryst.* **11**, 199-213.
MAKOWSKI, L. (1982). *J. Appl. Cryst.* **15**, 546-557.
MILLANE, R. P. (1989*a*). *Acta Cryst.* A**45**, 258-260.
MILLANE, R. P. (1989*b*). *Acta Cryst.* A**45**, 573-576.

MILLANE, R. P. (1990). *Acta Cryst.* A**46**, 68-72.
NAMBA, K. & STUBBS, G. (1985). *Acta Cryst.* A**41**, 252-262.
PARK, H. S., ARNOTT, S., CHANDRASEKARAN, R., MILLANE, R. P. & CAMPAGNARI, F. (1987). *J. Mol. Biol.* **197**, 513-523.
STARK, W., GLUCKSMAN, M. J. & MAKOWSKI, L. (1988). *J. Mol. Biol.* **199**, 171-182.
STUBBS, G. (1989). *Acta Cryst.* A**45**, 254-258.
WILSON, A. J. C. (1950). *Acta Cryst.* **3**, 397-399.

Acta Cryst. (1992). A**48**, 215-221

A Density-Matrix Approach to Coherence in High-Energy Electron Diffraction

BY A. G. WRIGHT* AND D. M. BIRD

School of Physics, University of Bath, Bath BA2 7AY, England

(*Received* 17 *May* 1991; *accepted* 29 *August* 1991)

Abstract

It is shown that a density-matrix formalism may be used to analyse various aspects of coherence in high-energy electron diffraction theory. This approach is demonstrated by two examples: an analysis of the coherence between two Bloch waves generated by diffuse scattering as a function of crystal thickness (the dependent-to-independent Bloch-wave 'transition') and an analysis of the coherence between electrons which are diffusely scattered in different directions and their contribution to high-resolution images.

1. Introduction

Using density-matrix theory one can fully describe a quantum or statistical system even if one cannot construct the exact wave function for the system (*e.g.* Blum, 1981). To quote Ziman (1969), 'we cannot learn more than is given by the density matrix; it is all we know and all we need to know about the "state" of the system'. One important feature of the density matrix is that it contains information about the correlation or coherence between the states of a quantum system; this is manifested in the off-diagonal elements of the relevant density matrix. The effects of interactions on the system are described by the Liouville equation which governs the evolution of the density matrix. Physical observables are then readily found by operating on the resulting matrix.

There are many different kinds of scattering experiments where a density-matrix description provides a convenient framework in which to describe the coherence among the states involved in the scattering. These include photon (*e.g.* Loudon, 1983), electron (*e.g.* Blum, 1981) and neutron (*e.g.* Balcar & Lovesey, 1989) scattering. The question of coherence arises in several contexts in high-energy electron diffraction theory. One obvious case is in high-resolution imaging, where the coherently diffracted beams are reconstructed (albeit with the effects of lens aberrations) to form an image. A more difficult question concerns the role of diffusely scattered electrons in high-resolution imaging (Cowley, 1988). Do these simply provide a uniform background, or do they also contribute to form a 'background' lattice image? The answer must lie in the degree of coherence between electrons which have been diffusely scattered to different points in the diffraction plane. A second area where coherence has been widely discussed is in the detailed mechanism of inelastic and diffuse scattering processes. A number of questions arise here, such as whether diffusely and inelastically scattered electrons preserve diffraction image contrast (*e.g.* Howie, 1963; Rez, Humphreys & Whelan, 1977) and whether dependent or independent Bloch-wave models provide the most appropriate description (*e.g.* Cherns, Howie & Jacobs, 1973; Wright & Bird, 1989, and references therein). There has been some confusion about what exactly is meant by coherence in this case: what precisely is coherent (or otherwise) with what? Our aim in this paper is to show that a density-matrix approach provides a unifying description which encompasses all these areas where coherence is the issue. The use of density matrices is not new in diffraction and channelling theory. Rez (1977) and Dudarev & Ryazanov (1988) set up a formalism to analyse multiple elastic and inelastic scattering but they do not attempt to analyse detailed questions of coherence. Kagan & Kononets (1973, 1974) discuss the use of density matrices in the theory of particle channelling and analyse the damping of the off-diagonal elements of the matrix (see below) due to inelastic scattering. However, they do not discuss aspects of the theory which are specific to high-energy electron scattering. The results we derive are not new and can be found without density-matrix theory. Nevertheless, our analysis shows how correlation and coherence can be discussed in a rigorous

* Present address: School of Physical Sciences, Flinders University, Bedford Park 5042, South Australia, Australia.

0108-7673/92/020215-07$03.00