(1980) for the same crystal and radiation were used to derive the relative magnitudes of the extinction corrections for the two sets of data. The reasonableness of using the same extinction tensor can be understood from the fact that the crystal had not been removed from its support between the two experiments; consequently the directions of the normals to the planes of diffraction for corresponding reflections remained unchanged. To obviate the need for making TDS corrections, the peak widths of the monochromator data were adjusted to be identical to their equivalents in the filtered data with peak maxima in identical positions relative to the peak widths. It was felt prudent not to include the {111} reflections in the determination of K since the absorption edge of the β filter fell beneath their peaks. No reflections were rejected on the basis of the difference between extinction corrections [maximum value of ypc was 0·0007 corresponding to $y_{ext} = 0{\cdot}8885$ (monochromator) and 0·8891 (filter) for the 220 reflection]. The value of K determined under these conditions was 0·805 (11).

In summary, our experimentally derived K for Cu $K\alpha$ radiation falls between the expected values for an ideally mosaic and perfect graphite monochromator. In contrast, the values we have obtained for Mo $K\alpha$ and Ag $K\alpha$ radiation fall outside this range. Moreover, and contrary to the results of Jennings (1968) and Le Page, Gabe & Calvert (1979), they are smaller than those expected for an ideally mosaic crystal.

References

BECKER, P. J. & COPPENS, P. (1974). *Acta Cryst.* A**30**, 129–153.

BECKER, P. J. & COPPENS, P. (1975). *Acta Cryst.* A**31**, 417–425.

DWIGGINS, C. W. (1975). *Acta Cryst.* A**31**, 395–396.

FLACK, H. D. & VINCENT, M. G. (1980). *Acta Cryst.* A**36**, 620–624.

HAMILTON, W. C. (1964). *Statistics in Physical Science.* New York: Ronald Press Co.

INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1978). *Acta Cryst.* A**34**, 159–160.

JENNINGS, L. D. (1968). *Acta Cryst.* A**24**, 472–474.

LEHMANN, M. S. (1975). *J. Appl. Cryst.* **8**, 619–622.

LEHMANN, M. S. & LARSEN, F. K. (1974). *Acta Cryst.* A**30**, 580–584.

LEPAGE, Y., GABE, E. J. & CALVERT, L. D. (1979). *J. Appl. Cryst.* **12**, 25–26.

MIYAKE, S., TOGAWA, S. & HOSOYA, S. (1964). *Acta Cryst.* **17**, 1083–1084.

NELMES, R. J. (1975). *Acta Cryst.* A**31**, 273–279.

RIGOULT, J. (1979). *J. Appl. Cryst.* **12**, 116–118.

STEVENS, E. D. (1974). *Acta Cryst.* A**30**, 184–189.

THORNLEY, F. R. & NELMES, R. J. (1974). *Acta Cryst.* A**30**, 748–757.

VINCENT, M. G. & FLACK, H. D. (1980). *Acta Cryst.* A**36**, 610–614.

VINCENT, M. G., YVON, K. & LEVINSON, M. (1980). In preparation.

WHITTEN, W. D., CHUNG, P. L. & DANIELSON, G. C. (1965). *J. Phys. Chem. Solids*, **26**, 49–56.

Acta Cryst. (1980). A**36**, 620–624

On the Polarization Factor for Crystal-Monochromated X-radiation. III. A Weighting Scheme for Products

BY H. D. FLACK AND M. G. VINCENT*

Laboratoire de Cristallographie aux Rayons X, Université de Genève, quai Ernest Ansermet 24, *CH* 1211 *Genève* 4, *Switzerland*

(*Received* 31 *December* 1979; *accepted* 11 *February* 1980)

Abstract

The minimum-variance weighting scheme for calculating the weighted mean of a quantity for the following unusual case is analysed. The quantity is derived from the product of two independent quantities for each of which a set of data is available. All the products between the two sets of data are taken into consideration. The application to measurements of the polarization state of X-radiation is outlined.

* Present address: Biozentrum der Universität Basel, Klingelbergstrasse 70, CH 4056 Basel, Switzerland.

0567-7394/80/040620-05$01.00

Introduction

In a previous paper (Vincent & Flack, 1980), a method is described for determining the polarization ratio, K, of X-radiation reflected from a monochromator. This value is estimated by taking the weighted mean of $M \times N$ individual values K_{ij}, *i.e.*

$$\hat{K} = \sum_{i=1}^{M} \sum_{j=1}^{N} a_{ij} K_{ij} \Big/ \sum_{i=1}^{M} \sum_{j=1}^{N} a_{ij}, \qquad (1)$$

where a_{ij} is the weight attributed to K_{ij}.

The K_{ij} are obtained from the relationship

$$K_{ij} = c_{ij} A_i B_j + d_{ij}, \quad (2)$$

where c_{ij} and d_{ij} are known constants and A_i, $(i = 1, 2, \ldots, M)$ and B_j, $(j = 1, 2, \ldots, N)$ are *independent* series of data (with e.s.d.'s) derived from intensity measurements. The subject of this paper is the choice of the 'proper' or 'best' weighting scheme, a_{ij}, to apply in the above-defined problem.

A simple choice for the weighting scheme would be to use

$$a_{ij} = 1/\sigma_{ij}^2,$$

where σ_{ij}^2 could be estimated from

$$\sigma_{ij}^2 = B_j^2 \sigma_i^2 + A_i^2 \sigma_j^2,$$

where σ^2 represents the population variance of the subscripted quantity (Hamilton, 1964). From a theoretical point of view this weighting scheme is unsatisfactory as it treats the $M \times N$ values of K_{ij}, derived from $M + N$ values of A_i and B_j, as independent observations whereas the error disturbances in the K_{ij} must be correlated.

From a practical point of view this simple weighting scheme was tested in a simulation with model data of A_i and B_j derived from a chosen value of K and with a Gaussian pseudo-random error added in. The value of $\hat{K}$ was frequently unreasonably different from the model value of K. The discrepancy between $\hat{K}$ and the model K increased with decreasing wavelength.

We were thus driven to determine a more satisfactory weighting scheme. The criterion to be applied (Hamilton, 1964) is that $\hat{K}$ should be an unbiased estimator of K and that its variance should be a minimum. The steps in the analysis, given below, to determine the a_{ij} are as follows. The error disturbances on the A_i and B_j, and the expected values of their first-, second-, third- and fourth-order moments are defined. It is then shown that, for any weighting scheme, $\hat{K}$ is an unbiased estimator of K. Next, the variance of $\hat{K}$ is determined and this value is minimized with respect to the a_{ij}. An expression for the weighting scheme in terms of the population variances of the A_i and B_j is thus obtained. We next suppose that estimates of these population variances are available but that these are only known within a scale factor. It is shown that the variance of $\hat{K}$ is proportional to this (as yet unknown) scale factor. It may, however, be estimated by considering the expected value of the weighted squared deviations of the K_{ij} about $\hat{K}$. Finally, a practical method for calculating the weighting scheme in a computer program is presented. As much of the algebra is straightforward, only the important intermediate equations are given in the text.

Analysis

Let us write

$$\gamma_{ij} = A_i B_j (i = 1, 2, \ldots, M; j = 1, 2, \ldots, N), \quad (3)$$

and K for the population mean of the desired quantity. In the following analysis it will be necessary to have multiple summations over the observations. As a matter of convention, i, k and m take values from 1 to M, and j, l and n take values from 1 to N. It is convenient also to choose weights a_{ij} so that

$$\sum_i \sum_j a_{ij} = 1. \quad (4)$$

Definition of errors

Let us assume that our A_i, B_j of (2) are subject to random errors. We may write

$$A_i = E_i + e_i \quad \text{and} \quad B_j = F_j + f_j, \quad (5)$$

where E_i and F_j are the population means and e_i and f_j represent the random errors on these quantities. Taking expected values, $\varepsilon\{A_i\} = E_i$ and $\varepsilon\{B_j\} = F_j$, we assume that

$$\varepsilon\{e_i\} = \varepsilon\{f_j\} = 0; \quad i = 1, \ldots, M, \quad j = 1, \ldots, N. \quad (6a)$$

Further, we will assume that the individual error terms have a finite variance

$$\varepsilon\{e_i^2\} = \sigma^2(E_i) \quad \text{and} \quad \varepsilon\{f_j^2\} = \sigma^2(F_j) \quad (6b)$$

and that the error disturbances are independent one from another, *i.e.*

$$\begin{aligned} \varepsilon\{e_i f_j\} &= 0, \quad \varepsilon\{e_i e_k\} = 0 \text{ for } i \neq k, \\ \varepsilon\{f_j f_l\} &= 0 \text{ for } j \neq l, \\ \varepsilon\{e_i^2 f_j^2\} &= \sigma^2(E_i)\sigma^2(F_j), \\ \varepsilon\{e_i f_j e_k f_l\} &= 0 \text{ for } i \neq k \text{ or } j \neq l, \\ \varepsilon\{e_i f_j e_k\} &= 0 \quad \text{and} \quad \varepsilon\{e_i f_j f_l\} = 0. \end{aligned} \quad (6c)$$

The expected value of $\hat{K}$

Substituting (5) into (2) and the result into (1), with (4), we obtain

$$\hat{K} = K + \sum_i \sum_j a_{ij} c_{ij} |E_i f_j + F_j e_i + e_i f_j|. \quad (7)$$

Taking expected values and applying (6) it follows that $\varepsilon\{\hat{K}\} = K$. *i.e.* $\hat{K}$ is an unbiased estimator of K.

The variance of $\hat{K}$

Squaring (7) we obtain

$$(\hat{K}-K)^2=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N}a_{ij}a_{kl}c_{ij}c_{kl}[E_iE_kf_jf_l + F_jF_le_ie_k + e_ie_kf_jf_l + 2E_iF_lf_je_k + 2F_je_ie_kf_l + 2E_ie_kf_jf_l]. \quad (8)$$

Taking expected values in (8) and applying (6), one finds

$$\varepsilon\{(\hat{K}-K)^2\}=\sum_i\sum_j a_{ij}^2c_{ij}^2[E_i^2\,\sigma^2(F_j) + F_j^2\,\sigma^2(E_i) + \sigma^2(E_i)\sigma^2(F_j)] + \sum_i\sum_j\sum_{k\neq i}^{}a_{ij}a_{kj}c_{ij}c_{kj}E_iE_k\,\sigma^2(F_j) + \sum_i\sum_j\sum_{l\neq j}^{}a_{ij}a_{il}c_{ij}c_{il}F_jF_l\,\sigma^2(E_i), \quad (9)$$

which is an expression for the variance of $\hat{K}$.

Minimizing the variance of $\hat{K}$

Let us now minimize $\varepsilon\{(\hat{K}-K)^2\}$ subject to the restriction (4) that $\sum_i\sum_j a_{ij}=1$. This may be achieved by differentiation of (9) including a Lagrangian multiplier. Set

$$Q=\varepsilon\{(\hat{K}-K)^2\}+\lambda\left(\sum_i\sum_j a_{ij}-1\right), \quad (10)$$

and the conditions $\partial Q/\partial a_{ij}=0$ produce

$$a_{ij}=-\eta_{ij}[\lambda + c_{ij}\,u_jE_i\,\sigma^2(F_j) + c_{ij}\,v_iF_j\,\sigma^2(E_i)], \quad (11)$$

where

$$\eta_{ij}=1/c_{ij}^2[E_i^2\,\sigma^2(F_j) + F_j^2\,\sigma^2(E_i) + 2\sigma^2(E_i)\,\sigma^2(F_j)], \quad (12)$$

$$u_j=\sum_k a_{kj}c_{kj}E_k \quad \text{and} \quad v_i=\sum_l a_{il}c_{il}F_l. \quad (13)$$

λ in (11) may be eliminated by imposing (4). One obtains

$$a_{ij}=\eta_{ij}\left\{(1+\psi)\Big/\sum_k\sum_i\eta_{kl} - c_{ij}[u_jE_i\,\sigma^2(F_j) + v_iF_j\,\sigma^2(E_i)]\right\}, \quad (14)$$

where

$$\varphi=\sum_i\sum_j c_{ij}\,\eta_{ij}[u_jE_i\,\sigma^2(F_j) + v_iF_j\,\sigma^2(E_i)]. \quad (15)$$

Thus, (14) defines the weighting scheme that we need although it is in an implicit form as the a_{ij} also appear on the right-hand side of (14) through the u_j and v_i in (13).

Use of sample variances of A_i and B_j

The weighting scheme defined by (14) requires values of $\sigma^2(E_i)$ and $\sigma^2(F_j)$ to be known. We will now suppose that we have experimental values of these quantities available $s^2(E_i)$, $s^2(F_j)$ but that they are only known to within an unknown scale factor, *i.e.*

$$\sigma^2(E_i)=\beta s^2(E_i) \quad \text{and} \quad \sigma^2(F_j)=\beta s^2(F_j). \quad (16)$$

In order to estimate β, it is useful to substitute (16) into (12)—(15), omit all terms in s^4 and of higher order and replace a_{ij} by a proportional quantity w_{ij}. We obtain

$$x_{ij}=\beta\eta_{ij}=1/c_{ij}^2[E_i^2\,s^2(F_j)+F_j^2\,s^2(E_i)], \quad (17)$$

$$w_{ij}=x_{ij}\left\{1+\varphi-c_{ij}[u_jE_i\,s^2(F_j) + v_iF_j\,s^2(E_i)]\left(\sum_k\sum_l x_{kl}\right)\right\}, \quad (18)$$

$$u_j=\frac{\sum_k w_{kj}c_{kj}E_k}{\sum_i\sum_l w_{il}} \quad \text{and} \quad v_i=\frac{\sum_l w_{il}c_{il}F_l}{\sum_k\sum_j w_{kj}}, \quad (19)$$

$$\varphi=\sum_i\sum_j c_{ij}x_{ij}[u_jE_i\,s^2(F_j)+v_iF_j\,s^2(E_i)], \quad (20)$$

$$\hat{K}=\sum_i\sum_j w_{ij}K_{ij}\Big/\sum_i\sum_j w_{ij} \quad (21)$$

and

$$\text{variance}\,(\hat{K})=\varepsilon\{(\hat{K}-K)^2\}=\beta\left[1+\varphi-\sum_i\sum_j\frac{w_{ij}^2}{x_{ij}}\Big/\sum\sum x_{ij}\right]\Big/\sum_i\sum_j w_{ij}. \quad (22)$$

The above relationships are readily derived when it is realized that

$$\sum_i\sum_j x_{ij}=\sum_i\sum_j w_{ij}. \quad (23)$$

Estimation of β

Consider $\sum_i\sum_j w_{ij}(K_{ij}-\hat{K})^2$ and its expected value. From (7), (2) and (5) one finds that

$$K_{kl}-\hat{K}=c_{kl}E_kf_l+c_{kl}F_le_k-\sum_i\sum_j w_{ij}c_{ij}E_if_j\Big/\sum_i\sum_j w_{ij}-\sum_i\sum_j w_{ij}c_{ij}F_je_i\Big/\sum_i\sum_j w_{ij}. \quad (24)$$

Taking expected values in (24) one finds

$$\varepsilon\left\{\sum_k\sum_l w_{kl}(K_{kl}-\hat{K})^2\right\}=\beta\left[\sum_k\sum_l (w_{kl}/x_{kl})-1-\varphi+\sum_i\sum_j (w_{ij}^2/x_{ij})\Big/\sum_i\sum_j x_{ij}\right], \quad (25)$$

which allows the value of β to be estimated by assuming that $\sum_k \sum_l w_{kl}(K_{kl}-\hat{K})^2$ has attained its expected value.

Evaluation of w_{ij}

The equations (18) for w_{ij} represent a $MN \times MN$ system of linear equations in MN unknowns. Rather than solve directly such a system by matrix inversion, we have preferred an iterative procedure which progressively improves the weights in a series of cycles. Suppose we have an 'old' system of weights wo_{ij} and that a 'new' better system is required by adding in a shift Δw_{ij}, *i.e.*

$$wn_{ij}=wo_{ij}+\Delta w_{ij}. \quad (26)$$

wn_{ij} will be made to obey (18) approximately, keeping in mind that (23) needs to hold, *i.e.*

$$\sum_i\sum_j wn_{ij}=\sum\sum wo_{ij}=\sum\sum x_{ij}$$

or

$$\sum\sum \Delta w_{ij}=0. \quad (27)$$

uo_i, vo_j and un_i, vn_j are the values defined by (19) with the 'old' and 'new' system of weights respectively. Likewise, φ_o and φ_n are defined from (20). Substituting (26) into (18) to make wn_{ij} obey (18) exactly one finds,

$$\begin{aligned}\Delta w_{ij}=&-wo_{ij}+x_{ij}+x_{ij}\varphi_o\\&+x_{ij}\sum_k\sum_l c_{kl}x_{kl}\left[s^2(F_i)E_k\left(\sum_m \Delta w_{ml}c_{ml}E_m\right)\right.\\&\left.+s^2(E_k)F_l\left(\sum_n \Delta w_{kn}c_{kn}F_n\right)\right]\Big/\sum_k\sum_l x_{kl}\\&-c_{ij}x_{ij}[uo_jE_is^2(F_j)+vo_iF_js^2(E_i)]\sum_k\sum_l x_{kl}\\&-c_{ij}x_{ij}\left[E_is^2(F_j)\left(\sum_m \Delta w_{mj}c_{mj}E_m\right)\right.\\&\left.+F_js^2(E_i)\left(\sum_n \Delta w_{in}c_{in}F_n\right)\right]. \quad (28)\end{aligned}$$

To simplify (28) we make a diagonal approximation of the $MN \times MN$ matrix defining Δw_{ij} by taking only constants and terms in Δw_{ij}. This is tantamount to assuming that all weights other than one individual w_{ij} are correct. This approximation leads to

$$\begin{aligned}&\Delta' w_{ij}\{2-c_{ij}x_{ij}[u'_js^2(F_j)E_i+v'_is^2(E_i)F_j]\}\\&=-wo_i+x_{ij}+x_{ij}\varphi_o-c_{ij}x_{ij}[uo_js^2(F_j)E_i\\&\quad+vo_is^2(E_i)F_j]\sum_k\sum_l x_{kl}, \quad (29)\end{aligned}$$

where

$$u'_j=\frac{\sum_k c_{kj}x_{kj}E_k}{\sum_k\sum_l x_{kl}},\qquad v'_i=\frac{\sum_l c_{il}x_{il}F_l}{\sum_k\sum_l x_{kl}}.$$

The second term in the braces $\{\}$ on the left-hand side of (29) is small in comparison to 2 and may be neglected. It is also desirable to damp the shift $\Delta' w_{ij}$ by a factor δ in a practical application. Hence we obtain

$$\begin{aligned}wn_{ij}=&(1-\delta/2)\,wo_{ij}+\delta/2\left(x_{ij}\left\{1+\varphi_0\right.\right.\\&\left.\left.-c_{ij}[uo_js^2(F_j)E_i+vo_is^2(E_i)F_j]\sum_k\sum_l x_{kl}\right\}\right). \quad (30)\end{aligned}$$

A suitable starting set of weights are the x_{ij}. In principle, it is the (unknown) population means E_i and F_j of the observations which enter into (30). In practice it is adequate to use the observed values A_i and B_j in place of E_l and F_j.

Application to polarization measurements

Details of the measurements and the technique are given in Vincent & Flack (1980). Intensity measurements are available from the same crystal for both monochromated and filtered radiation. These intensity measurements are divided into two sets, which are mutually exclusive, based on values of θ. We may write

$$\begin{aligned}A_i&=[(1+\cos^2 2\theta_i)/2]\,I_i\,(\text{monochromator})/I_i\,(\text{filter}),\\&\quad i=1,2,\ldots,M,\\B_j&=[2/(1+\cos^2 2\theta_j)]\,I_j\,(\text{filter})/I_j\,(\text{monochromator}),\\&\quad j=1,2,\ldots,N,\end{aligned} \quad (31)$$

where I is the intensity measurement for a particular reflection under the stated conditions. From (3), $\gamma_{ij}=A_iB_j$, the polarization state of the radiation may be written as

$$K_{ij}=(\cos^2 2\theta_i-\gamma_{ij}\cos^2 2\theta_j)/(\gamma_{ij}-1). \quad (32)$$

To apply the weighting scheme described above it is necessary to linearize (32). This is achieved by

expanding (32) to first order in a Taylor series about some 'guessed' value of K, K_g. One obtains the following values to be used in (2).

$$c_{ij} = \frac{(\cos^2 2\theta_j + K_g)^2}{\cos^2 2\theta_j - \cos^2 2\theta_i},$$

$$d_{ij} = K_g - \frac{(\cos^2 2\theta_i + K_g)(\cos^2 2\theta_j + K_g)}{\cos^2 2\theta_j - \cos^2 2\theta_i}. \quad (33)$$

The weighting scheme was implemented in a Fortran program. The weights were refined in cycles according to (30) until the maximum relative shift was less than a specified quantity (a value of 0·5% was chosen). It was found that δ had to be small (0·03) to avoid divergence. A considerable number of cycles (~650) were necessary in order for the weights to converge but the calculation nevertheless is extremely rapid on a modern computer. We have not attempted to improve the rate of convergence. The value of K_g initially chosen was for a kinematical monochromator and the whole calculation was rerun with K_g equal to $\hat{K}$ obtained from the first run. The value of $\hat{K}$ determined is highly insensitive to the value of K_g.

Some simulations were carried out to test the method and the program. These were based on the data used in paper II of this series but with intensities calculated from a model value of K, and with a pseudo-random Gaussian error added in to match the counting statistics of the experiment. From these simulations it is clear that the weighting scheme produces values of $\hat{K}$ closer to K than the simple weighting scheme of x_{ij} and that the e.s.d. of $\hat{K}$ is also reduced.

References

HAMILTON, W. C. (1964). *Statistics in Physical Science.* New York: Ronald Press Co.

VINCENT, M. G. & FLACK, H. D. (1980). *Acta Cryst.* A**36**, 614–620.

Acta Cryst. (1980). A**36**, 624–632

On the Theory and Estimation of the Three-Phase Structure Seminvariant in $P\bar{1}$

BY HERBERT HAUPTMAN

Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, New York 14203, USA

(*Received* 16 *October* 1979; *accepted* 8 *February* 1980)

Abstract

By embedding the three-phase structure seminvariant T and its three symmetry-related variants in suitable quintets Q (five-phase structure invariants) one obtains the extensions Q of T to which T is related *via* the space-group symmetries. The neighborhoods of T are then defined in terms of the neighborhoods of its extensions. The conditional probability distribution of T, given the seven magnitudes $|E|$ in its first neighborhood, is derived. The distribution yields a reliable estimate (0 or π) for T in the favorable case that the variance of the distribution happens to be small.

1. Introduction

In recent years the basic concepts and mathematical formalism needed for the development of the probabilistic theory of the structure seminvariants have been elucidated. For example, it has long been known that, for fixed enantiomorph, the collection of observed magnitudes $|E|$ determines, in general, the values of all the structure seminvariants. A major recent advance is the neighborhood principle: For fixed enantiomorph, the value of any structure seminvariant T is primarily determined, in favorable cases, by the values of one or more small sets of magnitudes $|E|$, the neighborhoods of T, and is relatively insensitive to the values of the great bulk of remaining magnitudes (Hauptman, 1975). The conditional probability distribution of T, given the magnitudes in any of its neighborhoods, yields an estimate for T which is particularly good in the favorable case that the variance of the distribution happens to be small. With the identification of systems of neighborhoods for the structure invariants, the probabilistic theory of the structure invariants developed rapidly, especially in space groups $P1$ and $P\bar{1}$, but a great deal of work still remains to be done in deriving accurate and readily computable probability distributions, particularly for the space-group-special structure invariants.

Again, with the formulation of the extension concept [Hauptman (1977*b*, 1978); but see Giacovazzo (1977) for the equivalent concept called representation theory], the probabilistic theory of the structure seminvariants was reduced to that of the structure invariants, which is more highly developed. In particular, the

0567-7394/80/040624-09$01.00