Acta Cryst. (1965). **19**, 683

The correlation of zones of X-ray intensity data from common reflexions.

By A. D. Rae*, *Chemistry Department, University of Auckland, New Zealand*

(*Received* 22 *December* 1964 *and in revised form* 1 *March* 1965)

The intensities of X-ray reflexions in a single zone, obtained by either multifilm Weissenberg or precession-camera techniques, are easily reduced to a single scale with little error. An exact, readily computable, least-squares procedure has been developed to accomplish the next stage which is the correlation of the individual zones through reflexions common to two or more of the zones.

The usual least-squares minimization of the sum of the weighted residuals, $R=\sum\limits_i w_i\Delta_i^2$, with respect to the parameters u_r, results in a set of linear simultaneous equations, $\sum\limits_i w_i\Delta_i\partial\Delta_i/\partial u_r=0$ ($r=1$ to N) provided the residuals Δ_i are linear in the parameters u_r. Rollett & Sparks (1960) used residuals $\Delta_{hij}=k_iF_{hi}^2-k_jF_{hj}^2$ and minimized $R=\sum\limits_{h,i,j} w_{hij}\Delta_{hij}^2$ where F_{hi}^2 is the value of $|F_{hkl}|^2$ in the ith zone. ($w_{hij}=0$ if $i=j$ or if there is no correlation observed between F_{hi}^2 and F_{hj}^2). The resulting equations are homogeneous and a further criterion upon the scale constants, k_i, is necessary for a non-trivial solution of the equations. Rollett & Sparks used a latent vector method incorporating the normalizing condition $\sum\limits_{i=1}^N k_i^2=1$, the method being a multidimensional analogue of that described by Schomaker, Waser, Marsh & Bergman (1959) for the similar problem of fitting least-squares lines and planes to a given set of points.

However, some other normalizing condition $\sum\limits_{i=1}^N c_ik_i^2=1$ could well be the correct condition, and their method of solution is no more justified than if a single scale factor were held constant. A further difficulty is that the weights used are dependent on the scale factors. Hamilton, Rollett & Sparks (1965) recognized these faults and used a method minimizing $\sum\limits_{h,i}\varphi_{hi}^2$ where $\varphi_{hi}^2=w_{hi}(F_{hi}^2-G_iF_h^2)^2$. $\sqrt{w_{hi}}$ varies as the reciprocal of the standard deviation of F_{hi}^2, $G_i=k_i^{-1}$ and F_h^2 is the best least-squares value for the reflexion and is a function of the G_i. Since the residuals are no longer linear in the G_i an iterative least-squares procedure must be used to solve the equations.

The necessity for iterative procedures can be avoided if residuals linear in the logarithms of the scale factors are considered. Either $\Delta'_{hij}=\log(k'_iF_{hi}^2)-\log(k'_jF_{hj}^2)$ or $\Delta_{hij}=\log(k_iF_{hi})-\log(k_jF_{hj})$ may be used as residuals in the minimization of $\sum\limits_{\substack{\text{all pairs}\\ h,i,j}} w_{hij}\Delta_{hij}^2$, the two procedures being equivalent since $\Delta'_{hij}=2\,\Delta_{hij}$, giving $\log k'_i=2\log k_i$ using the same values for w_{hij} in each case since ideally $\sqrt{w_{hij}}$ varies as the reciprocal of $\sigma(\Delta_{hij})$, the standard deviation of Δ_{hij}.

Consider the case of $\Delta_{hij}=\log(k_iF_{hi})-\log(k_jF_{hj})=l_i-l_j+\log(F_{hi}/F_{hj})$ where l_i is written for $\log k_i$. The least-squares equations are set up, in matrix notation $\mathbf{AL}=\mathbf{B}$, with matrix elements $a_{ij}=a_{ji}=-\sum\limits_h w_{hij}$; $a_{ii}=\sum\limits_{\substack{h,j\\ i\neq j}} w_{hij}$ and $b_i=-\sum\limits_{h,j} w_{hij}\log(F_{hi}/F_{hj})$. It is seen that $\sum\limits_{i=1}^N b_i=0$ and $\sum\limits_{i=1}^N a_{ij}=0$ showing that the equations are redundant. This implies that any one of the equations, say the first, may be discarded and the equations solved provided the correct additional condition upon the values of l_i is known. It is to be noted that if the equations were solved in terms of l_1, solutions of the type $l_i=r_il_1+s_i$ would be obtained, giving different values for the scale constants k_i unless all the values of r_i were fortuitously equal, in which case the logarithmic property of the l_i puts the solutions on a scale dependent on the initial choice of l_1.

The correct condition may be shown to be $\sum\limits_{j=1}^N a_{jj}l_j=0$. In the equations $\mathbf{AL}=\mathbf{B}$ only the elements of $\mathbf{B}$ involve the values of F_{hi}. This implies that a rescaling of the data in a second cycle of calculations would generate the equations $\mathbf{A}\Delta\mathbf{L}=\mathbf{B}'$ having the solution $\Delta\mathbf{L}=0$ provided the elements b'_i of $\mathbf{B}'$ were all zero. This then is the necessary condition for the original solution to be correct.

It is seen that $b'_i=-\sum\limits_{h,j} w_{hij}\log(k_iF_{hi}/k_jF_{hj})=b_i-\sum\limits_{j=1}^N a_{ij}l_j$.

From the condition $\sum\limits_{j=1}^N a_{jj}l_j=0$ used in the evaluation of the l_i, b'_i may be expressed as

$$b'_i=b_i+\sum_{j=1}^N(a_{jj}-a_{ij})l_j\,.$$

From the jth of the original equations

$$a_{jj}l_j=b_j-\sum_{\substack{i=1\\ i\neq j}}^N a_{ji}l_i$$

and thus

$$b'_i=\sum_{j=1}^N b_j-\sum_{\substack{i,j=1\\ i\neq j}}^N a_{ji}l_i\,.$$

Since $\sum\limits_{j=1}^N b_j=0$ and $\sum\limits_{j=1}^N a_{ji}=0$

$$b'_i=\sum_{i=1}^N a_{ii}l_i=0\,.$$

It has then been established that it is possible to evaluate, without iteration, the scale constants of the zones, correlating each reflexion with its equivalent reflexions in any of the remaining zones, each such correlation being individually weighted.

The following procedure was adopted in this laboratory for an IBM 1620 computer. Since $\Delta_{hij}=l_i-l_j-\log(F_{hi}/F_{hj})$, the weights used, $w_{hij}=\sigma^2(\Delta_{hij})=\sigma^2\{\log(F_{hi}/F_{hj})\}$, are independent of the scale factors and the values used were $w_{hij}=1/(m_{hi}^2+m_{hj}^2)$, expressing the standard deviation of F_{hi} as $m_{hi}F_{hi}$ (or of F_{hi}^2 as $2\,m_{hi}F_{hi}^2$). Values of m_{hi}^2 were estimated from the observed intensities as follows;

---

* Present address: Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England.

$$I > I_1 \quad m_{hi}^2 = 1$$
$$I_1 > I > I_2 \quad m_{hi}^2 = I_1/I$$
$$I < I_2 \quad m_{hi}^2 = (I_2/I)^{\frac{1}{2}} I_1/I\,.$$

A percentage error was assumed for intensities greater than I_1, taken to be the highest intensity that can be accurately measured from the top intensity film of a Weissenberg multipack. A constant error was assumed for intensities I_2 and I_1, the intensity region of the top film for which further measurement could be obtained from the second film of the multipack. An arbitrary factor $(I_2/I)^{\frac{1}{2}}$ was applied below I_2 to account for difficulty in accurately assessing these reflexions.

The reflexions were punched on cards together with the appropriate values of m_{hi}^2 and i. The cards were sorted to put all the cards of the same hkl indices together so that the matrix elements of the equations $\mathbf{AL} = \mathbf{B}$ could be built up for each set of indices without prolonged storage of the individual reflexions. The equations were solved by replacing the first of the N normal equations by the equation $\sum_{j=1}^{N} a_{jj} l_j = 0$. The values for the scale constants were punched out on cards and were used in a subsidiary program to obtain final values for F_{hkl} and mF_{hkl}, the standard of F_{hkl}, from the expressions

$$F_{hkl} = \frac{\sum_i m_{hi}^{-2} k_i F_{hi}}{\sum_i m_{hi}^{-2}} \qquad F_2^{hkl} m^2 = \frac{\sum_i m_{hi}^{-2} k_i^2 F_{hi}^2}{(\sum_i m_{hi}^{-2})^2}\,.$$

It was found advantageous to recalculate the scale constants omitting those correlations for which $|\log (k_i F_{hi}/k_j F_{hj})|$ exceeded an arbitrary limit as a check on the accuracy of the data measurement and to exclude data containing extinction errors.

The basic difference between the present method and that of Hamilton, Rollett & Sparks is aptly demonstrated by considering the first example given by them.

Example 1. Consider the following set of data

	Layer 1	Layer 2	Layer 3
F_1^2	100	2	
F_2^2		1	2
F_3^2	100	3	

Both methods give $k_2/k_3 = 2{\cdot}00$ independent of the weights used though the method of Hamilton, Rollett & Sparks gives a value of

$$\frac{k_1}{k_2} = \frac{F_{12}^2 + F_{32}^2}{F_{11}^2 + F_{31}^2} = \frac{0{\cdot}050}{2{\cdot}000}$$

if all the weights are the same, that is there exists a constant absolute error in the F_{hi}^2 values, whereas the present method gives a value of

$$\frac{k_1}{k_2} = \frac{\sqrt{(F_{12}^2 \,.\, F_{32}^2)}}{\sqrt{(F_{11}^2 \,.\, F_{31}^2)}} = \frac{0{\cdot}049}{2{\cdot}000}$$

if all the weights are the same, that is there exists a constant percentage error in the F_{hi}^2 values. For a constant absolute error in the F_{hi}^2 values the present method gives a value of $k_1/k_2 = 0{\cdot}053/2{\cdot}000$, giving a larger weight to the percentagewise more reliable correlation of F_3^2. It seems more realistic to use the present method of correlation depending on the percentage error of individual correlations, though with a large number of data little difference between the values given by the two methods should result.

References

Rollett, J. S. & Sparks, R. A. (1960). *Acta Cryst.* **13**, 273.

Schomaker, V., Waser, J., Marsh, R. E. & Bergman, G. (1959). *Acta Cryst.* **12**, 600.

Hamilton, W. C., Rollett, J. S. & Sparks, R. A. (1965). *Acta Cryst.* **18**, 129.

Acta Cryst. (1965). **19**, 684

Refinement of the crystal structure of black phosphorus.

By Allan Brown and Stig Rundqvist, *Institute of Chemistry, University of Uppsala, Uppsala, Sweden*

(*Received* 26 *February* 1965)

Elementary phosphorus occurs in a number of modifications, but so far only the crystal structure of the black, orthorhombic modification has been determined. On the basis of powder diffraction data, Hultgren, Gingrich & Warren (1935) reported the following structure: Space group *Cmca*, $Z = 8$, unit-cell dimensions $a = 3{\cdot}31$, $b = 10{\cdot}50$, $c = 4{\cdot}38$ kX; 8P in 8(f) with $y = 0{\cdot}098$, $z = 0{\cdot}090$.

In the course of recent studies of the polymorphism of phosphorus, it was found that the black modification can be prepared in a well-crystallized condition from solutions of phosphorus in liquid bismuth. The crystals, obtained as a residue after dissolving the bismuth matrix with 1:1 nitric acid, were generally needle shaped with the needle axis coincident with the crystallographic a axis. Full details of the preparation of black phosphorus from liquid bismuth will be reported elsewhere. It may, however, be mentioned here that the sample used in the present investigation contained only 0·055 at. % Bi as determined by chemical analysis. Some of this bismuth may possibly be dissolved in the black phosphorus crystals, but there are strong reasons for believing that most of this residual bismuth is present only as a mechanical impurity left by incomplete dissolution in nitric acid.

Preliminary X-ray investigations of the black phosphorus product indicated that the structure proposed by Hultgren *et al.* is substantially correct. Since this structure determination is not very accurate by present standards, a refinement of the structure by single-crystal methods was thought to be worth while.

Accurate unit-cell dimensions were determined by powder diffraction methods. Powder patterns were recorded at room temperature (22 °C) both in a Guinier–Hägg camera using strictly monochromatized Cr $K\alpha_1$ radiation and silicon ($a = 5{\cdot}43054$ Å) as internal calibration standard, and in a Philips Debye–Scherrer camera with Cr K radiation. The dimensions obtained from measurements of the Guinier films are $a = 3{\cdot}3138$, $b = 10{\cdot}477_6$, $c = 4{\cdot}3759$ Å, while those from measurements of the Debye–Scherrer films are $a = 3{\cdot}3136$, $b = 10{\cdot}477_8$, $c = 4{\cdot}3766$ Å (Nelson–Riley extrapolation from lines with $54° < \theta < 85°$). These results yield the following