Acta Cryst. (1971). A**27**, 540

Dispersion Analysis of Single-Crystal Diffractometer Measurements. II. Variance Analysis of a Set of Intensity Measurements

BY H. SCHULZ

Institut für Kristallographie und Petrographie of the Eidgenössische Technische Hochschule, Sonneggstrasse 5, 8006 *Zürich, Switzerland*

(*Received* 27 *July* 1970 *and in revised form* 23 *November* 1970)

A single crystal of β-eucryptite ($Li_2O.\ Al_2O_3.\ 2SiO_2$), shaped as a sphere, has been used to collect reflexion intensities. A variance analysis has been carried out using 830 groups of symmetry-equivalent reflexions formed from 3110 single measurements. Each group contained at least 2, but not more than 6 reflexions. The groups were divided into ranges corresponding to their diffraction angle θ and their intensity mean value. The long time drift of the diffractometer was checked and corrected by a control reflexion. After this correction, the estimated variances of the lower intensity ranges agreed with the expected variances calculated from the counting-statistics. The groups with larger intensities showed further effects. The estimated standard deviation of the group with the largest intensities is 5 times larger than the standard deviation expected from the counting statistics.

1. Experimental features

β-Eucryptite ($Li_2O.\ Al_2O_3.\ 2SiO_2$) has a high quartz-like structure. It crystallizes in space group $P6_222$ (Winkler, 1948; Tscherry & Laves, 1970; Tscherry & Schulz, 1970). The crystal sphere used in the experiment had a diameter of $d=225\ \mu$ with an estimated standard deviation of a single measurement $\sigma(d)=3{\cdot}6\ \mu$.

The diffractometer (*cf.* part I, §4) selects measuring time and the weakening filter in order to measure the peak intensity of the reflexions with approximately the same count rate (40,000 counts) and to get a dead-time loss of less than 1%. For weak reflexions a maximum value for the measuring time per step can be present.

After every 15 intensity measurements the intensity of a control reflexion was measured. Other details are listed in Table 1.

About 5000 single measurements were grouped into symmetry-equivalent reflexions. Only groups with at least two single measurements and an unweighted intensity mean value larger than its threefold expected standard deviation were used in analysing the variance. A total of 830 groups with 3110 single measurements was obtained.

2. Equations for the analysis of the variance

Reflexion intensities (I) are calculated using linear combinations of pulse rates R_n [*cf.* part I, §1, equation (1)]. In general, the intensities (I) must often be multiplied by factors (K_m) in order to correct for a long time drift of the equiment or for the influence of weaking filters (Abrahams 1969). Therefore, the correct intensities IC and their expected variance $V(IC)$ are calculated with equations (1) and (2):

$$IC=\prod_m K_m \sum_n a_n R_n = I \prod_m K_m \tag{1}$$

$$\begin{aligned} V(IC) &= \prod_m K_m^2 \sum_n a_n^2 V(R_n) + (\sum a_n R_n)^2 \sum_m V(K_m) \\ &\simeq \prod_m K_m^2 \sum_n a_n^2 R_n + (\sum a_n R_n)^2 \sum_m V(K_m) \\ &= \prod_m K_m^2 \sum_n a_n^2 R_n + I^2 \sum_m V(K_m) \end{aligned} \tag{2}$$

Table 1. *Details of the intensity measurements.*

Measuring section	θ range	Measured reflexions	Parts of the reciprocal lattice used for the measurement	Maximum measuring time/measuring step	Max. deviation between the control reflexions	Expected maximum deviation $6\sigma_r(CR)100$
I	0–20°	Main reflexions ($h=2n$, $k=2n$, $l=2n$) Superstructure reflexions	$k\geq0$ $l\geq0$	0·6 sec/0·01 °θ	6·0%	4·8%
II	20–40	Main reflexions Superstructure reflexions	$h\leq0$ $k\geq0$ $l\leq0$	0·6 sec/0·01 °θ	10·0	4·5
III	40–70	Main reflexions	$h\leq0$ $k\geq0$ $l\geq0$	0·6 sec/0·01 °θ	6·8	7·2

Based on equations (1) and (2) formulae are given in the literature (Busing & Levy, 1957; Cetlin & Abrahams, 1963; Smith & Alexander, 1963; Abrahams & Bernstein, 1965; Killean, 1967), whereby an approximately normal distribution of each correction factor has been assumed.

The diffractometer used in the experiment measured the peak and background intensities P and B within equal time. Therefore, the reflexion intensities (I) equal the difference of these two pulse rates [*cf.* part I, §2, equation (8)], and equations (1) and (2) result in equations (3) and (4):

$$IC = STC(F)^q AD(P-B)\,, \tag{3}$$

$$\frac{V(IC)}{IC^2} = \frac{P+B}{(P-B)^2} + \frac{V(C)}{C^2} + \frac{q^2V(F)}{F^2} + \frac{V(A)}{A^2} + \frac{V(D)}{D^2}\,, \tag{4}$$

where

S = scale factor $V(S)=0$, constant for the whole data set,

T = measuring time factor, $V(T)=0$,

IC = intensity (corrected),

C = control reflexion factor,

F = weakening factor of one filter,

q = number of weakening filters, used

A = factor for absorption correction, and

D = correction factor for other measuring errors.

In the following equations the term IC is used, if at least one correction factor K is applied to I.

The scale factor was used during data reduction to avoid a storage overflow. The measuring time factor compensates for the different measuring times of the intensities (*cf.* §1).

C equalizes the long time drift of the diffractometer. C and $V(C)$ are calculated with equations (5) and (6):

$$C(j) = CR(1)/CR(j), \tag{5}$$

$$\begin{aligned} V[C(j)] &= C(j)^2V[CR(j)]/CR(j)^2,\\ &= C(j)^2\sigma_r^2[CR(j)] \end{aligned} \tag{6}$$

where $CR(j)$ is the intensity of the control reflexion j and σ_r is the relative standard deviation.

Intensities measured between $CR(j)$ and $CR(j+1)$ are multiplied by $C(j)$. But this simple correction can be applied only if the deviations of the control reflexion from their mean value have the same order of magnitude as the expected standard deviation $\sigma(CR)$. From equations (4) and (6) follows equation (7). The expressions $\sigma_r(x)$ in equation (7) hold for the relative standard deviation $\sigma(X)/X$ of any value X:

$$\begin{aligned} \sigma_r^2(IC) = \sigma_r^2(I) + C(j)^2\sigma_r^2[CR(j)]\\ + q^2\sigma_r^2(F) + \sigma_r^2(A) + \sigma_r^2(D). \end{aligned} \tag{7}$$

3. Short time drift of the diffractometer

The distribution function of the diffractometer for large pulse rates was determined by repeated measurements of the same reflexion intensity (*cf.* part I, §5). The results showed that the expected variance $V''(I)$ (*cf.* part I, §1), calculated from the counting statistics, is probably too small. The 179 intensities used for this statement were measured within two hours, at a time during which the diffractometer probably had the smallest drift. Nevertheless, the discrepancy between expected and estimated variance ($V''(I)$ and $V'(I)$) may be caused by a long time drift, which could be corrected. To check this, the influence of a pure short time drift was determined from the pulse rates of the peak (P), the background (B) and the reflexion intensity (I). A procedure in which the differences (8) were calculated was used. The expected values of these differences and their expected variances are given by equations (9) and (10).

$$\begin{aligned} DP_{n,m} &= P_n - P_{n+m} \qquad 1 \le n \le 179-m\\ DB_{n,m} &= B_n - B_{n+m} \qquad m = 1,2,\ldots\\ DI_{n,m} &= I_n - I_{n+m} \end{aligned} \tag{8}$$

$$EW(DP) = EW(DB) = EW(DI) = 0$$

$$\begin{aligned} V''(DP) &= 2\bar{P}\\ V''(DB) &= 2\bar{B}\\ V''(DI) &= 2(\bar{P}+\bar{B}) \end{aligned} \tag{10}$$

The $\chi^2(DP_m)$, $\chi^2(DB_m)$ and $\chi^2(DI_m)$ are now calculated according to equation (12), part I, §2. These χ^2 values have a degree of freedom equal to $(179-m-1)$.

The estimated values of DP_m, DB_n and DI_m lie between -4 and $+5$ up to $m=15$. For $m>15$, systematic deviations between -20 and $+20$ appear. Therefore, the first 11 $\chi^2(DP_m)$, $\chi^2(DB_m)$, $\chi^2(DI_m)$ were summed up. Their values are $\chi^2(DP)=2400$, $\chi^2(DB)=2290$ and $\chi^2(DI)=2380$; their degree of freedom is $G=1892$; therefore $\pi(\chi^2)=61{\cdot}5$ [*cf.* part I, §3, equations (17) and (18)]. The value of $|\chi^2-G|$ is larger than $3\sigma(\chi^2)$ in all three cases; this means the denominator of the χ^2 values, the expected variance $V''(I)$, is too small.

The $|\chi^2-G|$ values are about in the same range, although the counting rate of the peak is 6 times larger than that of the background. Due to the summation of 11 χ^2 values, 11 counting rates measured within 10 minutes are compared with each other. The intensity of a control reflexion, which can be used to correct the long time drift of the diffractometer, was measured every second hour only. Deviations from the variance caused by a short time drift can be corrected only by increasing this variance. A correction factor was chosen which decreases the χ^2 values to the 'aim value' $G+2\sigma(\chi^2)$. It is equal to 1·19 for the peak, 1·14 for the background and 1·18 for the intensity. A factor of 1·15 is used in the calculations that follow in §4.

4. Analysis of the variance of a set of intensity measurements

1. *Arrangement of the 'property groups' and 'aim ranges'*

The symmetry-equivalent reflexions were divided into property groups (*P* groups, *cf.* part I, §3) according to their diffraction angle θ and their intensity mean value $\bar{IC}$. Some characteristics of the *P* groups are listed in Table 2. The intensity ranges are designated as *I*1–*I*6, the θ ranges as $\theta 1$–$\theta 3$.

The analysis is carried out in steps to determine the influence of each correction factor [method (*b*), part I, §3]. The degrees of freedom of the $\chi^2(P)$ values equal the number of used reflexions minus the number of corresponding *S* groups (Table 2).

The expected value of a $\chi^2(P)$ value with a degree of freedom of *G* is approximately equal to *G*. The quotient $Q=\chi^2(P)/G$ should be equal to about 1. Value $\chi^2(P)$ belongs to a normal distributed totality with the variance $2G$ (*cf.* part I, §3). Therefore, *Q* may deviate from 1, but it should lie only with a probability of 0·06 outside the limiting values Q_l given by equation (11):

$$Q_{\pm l}=(G\pm 2\sigma[\chi^2(P)])/G = 1\pm 2\sigma_r[\chi^2(P)]\,. \quad (11)$$

These limiting values are listed in Table 3, last subline.

4·2. *Initial values*

The *Q* values, calculated from the intensities and their expected variances determined from the pulse rates without application of correction factors, are listed in Table 3, subline 1. Only the *Q* values of the *I*1, *I*2 ranges are nearly equal to 1. Some *Q* values lie within the aim range. Value $Q(I3,\theta 2)$ and all *Q* values of higher intensity ranges are larger than 3.

4·3. *Control reflexion correction*

Note that $Q(I3,\theta 2)$ is much larger than $Q(I3,\theta 1)$ and $Q(I3,\theta 3)$. These differences are explained easily by the very different spread of the control reflexion intensities (Table 1, last 2 columns).

Therefore, a control-reflexion correction corresponding to equations (3) to (6) has been applied to the data set. After this correction, the *Q* values were recalculated (Table 3, subline 2). The correction's in-

Table 2. *Details of the property groups.*

		Intensity ranges					
		*I*1	*I*2	*I*3	*I*4	*I*5	*I*6
		10^1–10^2	10^2–10^3	10^3–10^4	10^4–10^5	10^5–2·5 10^5	2·5 10^5–8·7 10^5
θ ranges	Subline						
	1	105	325	155	77	33	11
θ_1	2	18	61	35	16	8	3
(0–20°)	3	87	264	120	61	25	8
	4	74	420	2,700	46,000	150,000	510,000
	1	482	657	298	87	–	–
θ_2	2	133	198	89	26	–	–
(20–40°)	3	349	459	209	61	–	–
	4	74	300	3,900	28,000	–	–
	1	62	702	246	–	–	–
0_3	2	16	205	75	–	–	–
(40–70°)	3	46	497	171	–	–	–
	4	77	420	2,000	–	–	–

Subline 1: Number of single reflexions.
Subline 2: Number of *S* groups
Subline 3: Degree of freedom.
Subline 4: Mean value of the intensity.

Table 3. *Q values of the property groups*

	Subline	*I*1	*I*2	*I*3	*I*4	*I*5	*I*6
	1	1·05	1·34	1·32	6·0	17·6	81·0
	2	1·06	1·32	1·46	3·0	7·1	36·8
$\theta 1$	3	0·92	1·15	1·27	2·6	6·2	32·0
	4	–	–	1·17	2·1	5·0	25·9
	5						13·7
	6	0·69–1·31	0·83–1·17	0·74–1·26	0·64–1·36	0·44–1·56	0–2·0
	1	1·30	1·57	3·30	8·91		
	2	1·32	1·08	1·37	2·04		
$\theta 2$	3	1·15	0·94	1·19	1·78		
	4	–	–	1·05	1·43		
	5	0·85–1·15	0·86–1·14	0·79–1·21	0·64–1·36		
	1	0·85	1·06	1·09			
	2	0·86	1·06	1·12			
$\theta 3$	3	0·75	0·92	0·97			
	4	–	–	0·92			
	5	0·58–1·42	0·88–1·12	0·79–1·21			

fluence is clearly evident from the $Q(I3, \theta 2)$ and $Q(I4, \theta 2)$ values, which, when corrected are smaller than $Q(I3, \theta 1)$ and $Q(I4, \theta 1)$.

4·4. *Increasing the expected variance*

It was shown in §3, that it is necessary to increase the expected variance $V''(I)$, calculated from the counting statistics, by a factor of about 1·15 due to short time drifts of the diffractometer. Therefore, the Q values calculated after the control reflexion correction (Table 3, subline 2) must be divided by this factor (Table 3, subline 3).

The Q values of the $I1$–$I3$ classes except for $Q(I3, \theta 1)$ now lie within the aim range and spread around 1. Value $Q(I3, \theta 1)$ reaches nearly the upper limit Q_{+l}.

4·5. *Absorption correction*

The radius of the crystal sphere has an estimated relative standard deviation $\sigma_r(r) = 0{\cdot}016$ (*cf.* §1), which gives corresponding errors of the absorption correction. For $\mu r = 0{\cdot}087$, $A = 1$ and

$$\sigma_r(A) = 3{\cdot}8 * \mu r * \sigma_r(r) = 0{\cdot}0053 \ .$$

This equation is independent of θ (Jeffrey & Rose 1964, Fig. 4).

$V(IC)$ must be recalculated using equation (7). But it is not necessary to calculate a new χ^2 value for each S group. A sufficiently good determination of these new Q values is possible using the present Q values. Up to now, $\sigma_r^2(IC)$ has been calculated only with the first and second terms of equation (7):

$$\sigma_r^2(IC) = \sigma_r^2(I) + C(j)^2 \sigma_r^2[CR(j)]$$

Fig. 1 shows the relative standard deviation $\sigma_r^2(IC)$ as a function of IC after the control reflexion correction [(*cf.* equation (5)]. For the three largest intensity classes, $\sigma_r^2(IC)$ is nearly constant and equal to 0·010. The value of $\sigma_r^2(IC)$ increases if the intensity is lower than a characteristic intensity, which depends on the maximum measuring time/measuring step (*cf.* §1).

For each P group one $\sigma_r(IC)$ can be estimated using Fig. 1 with the intensity mean value of a P group. These intensity mean values are listed in Table 2, subline 4. The standard deviation, which belongs to this intensity mean value, can be read from Fig. 1. The expression $\tilde{\sigma}_r$ is used for this sort of an estimate of one relative standard deviation for one P group. The standard deviations read from Fig. 1. correspond to

$$\tilde{\sigma}_r^2(IC) = \tilde{\sigma}_r^2(I) + C(j)^2 \tilde{\sigma}_r^2[CR(j)] \ .$$

This expression changes into equation (12) after increasing the expected variances as described in §4·4.

$$\tilde{\sigma}_r^2(IC) = 1{\cdot}15 \ \{\tilde{\sigma}_r^2(I) + C(j)^2 \tilde{\sigma}_r^2[CR(j)]\} \ . \qquad (12)$$

The new mean values of the variances, taking into consideration the absorption influence, can be calculated with equation (13):

$$\tilde{\sigma}_r^2(IC) = 1{\cdot}15 \ \{\tilde{\sigma}_r^2(I) + C(j)^2 \tilde{\sigma}_r^2[CR(j)]\} + \pi_r^2(A) \ . \qquad (13)$$

The Q values are decreased corresponding to the ratio of the variances (12) and (13); this means that the Q values have to be divided by the ratio of equations (13)/(12). The new Q values are listed in Table 2, subline 4.

The Q values of the $I1$–$I3$ ranges now fall into the aim range, and $Q(I4, \theta 2)$ has become smaller than $1 + 3\sigma_r[\chi^2(P)]$.

4·6. *Weakening filter correction*

The filter factors $(F)^q$ [*cf.* equation (3)] were measured three times with pulse rates of about 10^6 counts. The deviations between the three measurements increased rapidly with the number of filters.

Therefore, F and $\sigma(F)$ were estimated directly from the measured filter factors $(F)^q$:

$$F = 2{\cdot}160 \qquad \sigma(F) = 0{\cdot}011 \ .$$

Only the intensities of the $P(I6, \theta 1)$ group have been measured with 3 or 4 filters. This group was split into two groups of intensities measured with the same filter. Then a new $\chi^2(P)$ was calculated separately for each group accounting for $\sigma(F)$ and using the procedure described in §4·5. These two $\chi^2(P)$ values were added and then the final $Q(I6, \theta 1)$ value was calculated (Table 3, subline 5).

4·7. *Calculation of* $V(D)$

The Q values of the $(I4$–$I6, \theta 1)$ ranges are larger than two in spite of the applied correction used so far. They increase with increasing intensity range.

Fig. 1 shows that after the control-reflexion correction $\sigma_r(IC)$ is approximately equal for the three largest intensity classes. Therefore, it is not possible to explain the increase of the Q values by decreasing variances. By means of $V(D)$ [*cf.* equations (3) and (4)] it is quite easy to decrease $Q(I4, \theta 1)$, $Q(I5, \theta 1)$ and $Q(I6, \theta 1)$ to their corresponding Q_{+l} values. $V(D)$ shows the deviation from the symmetry equivalence which could not be removed by the described corrections.

It follows from §4·5 and equations (7) and (13) that Q decreases to Q_{+l}, if $\tilde{\sigma}_r^2(D)$ given by equation (14) is added to the present $\tilde{\sigma}_r^2(IC)$ value:

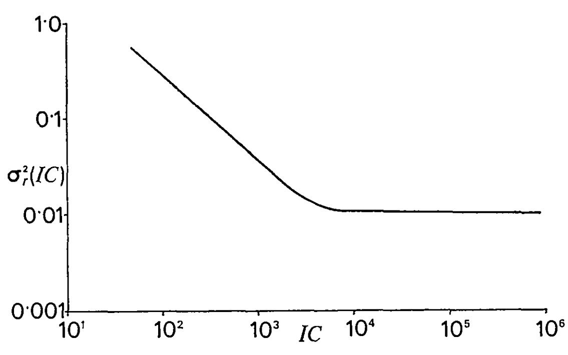

Fig. 1. Relative standard deviations of the S groups after the control-reflexion correction.

$$\tilde{\sigma}_r^2(D) = \frac{Q-Q_{+l}}{Q_{+l}} \{1 \cdot 15\,[\tilde{\sigma}_r^2(I)+C(j)^2\tilde{\sigma}_r^2[CR(j)]] + \sigma_r^2(A)+q^2\sigma_r^2(F)\} \quad (14)$$

Only for the $P(I6,\theta1)$ group, which contains the measurements with filters, is $q^2\sigma_r^2(F)$ unequal to zero.

$\tilde{\sigma}_r(D)$ and the final $\tilde{\sigma}_r(IC)$ of all intensity ranges are listed in Table 4. The measured error of the largest intensity range is 5 times larger than the one calculated from the pulse rates. Measurements of the $I3$, $I4$, and $I5$ ranges have the smallest errors with $\tilde{\sigma}_r(IC)$ values between 0·013 and 0·024.

Table 4. *Estimated mean values of relative standard deviations*

	Subline	Intensity ranges I1	I2	I3	I4	I5	I6
θ1	1	–	–	–	1·0	2·0	5·0
	2	45.	9·2	1·8	1·6	2·4	5·5
θ2	1	–	–	–	–	–	–
	2	45.	13·0	1·6	1·3	–	–
θ3	1	–	–	–	–	–	–
	2	45·	9·2	2·4	–	–	–

Subline 1: $\tilde{\sigma}_r(D)\ 10^2$.
Subline 2: $\tilde{\sigma}_r(IC)\ 10^2$.

$\tilde{\sigma}_r(D)$ increases from about 0·01 in the $I4$ range, over 0·02 in the $I5$ range, to 0·05 in the $I6$ range. These additional standard deviations are probably caused by the measuring procedure. The intensity of the peak was measured with the θ–2θ scan method, the background being in a fixed position on the right and left sides of the peak. This measuring method is applicable if the background intensity is a linear function (Fischer & Hahn, 1961). A drift of the peak within this measuring window causes measuring error whose magnitude increases with decreasing θ value and increasing intensity of the reflexion.

The measurements of the $P(I4,\theta1)$ group were used to confirm this hypothesis. The S groups were divided into two groups. One group contained all reflexions with θ values between 10 and 16°. The other group contained all reflexions with θ values larger than 16°. The Q values of both groups were calculated. It was found for the first group that $Q=2\cdot41$ with $Q_{+l}=1\cdot51$ and for the second group $Q=1\cdot35$ with $Q_{+l}=1\cdot53$. The reflexions of the first group with small θ values do not fall into the aim range, but the reflextions of the second group do fall into the aim range. This result confirmed the proposed reason for the existence of $\tilde{\sigma}_r(D)$.

5. Summary

The method described in this paper can be used to test the agreement between the estimated and expected variances of any subset of intensity measurements. It can be utilized to:

(*a*) Determine realistic weights from the measurements themselves.
(*b*) Compare estimated and expected deviations of intensity measurements from their mean values.
(*c*) Check a correction that influences symmetry-equivalent reflexions in different ways (*e.g.* absorption correction).
(*d*) Decide between two symmetries that a crystal might have, if the groups of symmetry-equivalent reflexions do not agree for the two symmetries.

I thank Prof. Hoffmann and Prof. Linder for many helpful discussions and Mr Hirsiger for his help in evaluating the measurements. The diffractometer that was used was made available by the Owens Illinois Company.

References

Abrahams, S. C. (1969). *Acta Cryst.* A**25**, 165.
Abrahams, S. C. & Bernstein, J. L. (1965). *Acta Cryst.* **18**, 926.
Busing, W. R. & Levy, H. A. (1957). *J. Chem. Phys.* **26**, 563.
Cetlin, B. B. & Abrahams, S. C. (1963). *Acta Cryst.* **16**, 943.
Fischer, K. & Hahn, Th. (1961). *Z. Kristallogr.* **116**, 27.
Jeffrey, J. W. & Rose, K. M. (1964). *Acta Cryst.* **17**, 343.
Killean, R. C. G. (1967). *Acta Cryst.* **23**, 54.
Smith, G. S. & Alexander, L. E. (1963). *Acta Cryst.* **16**, 462.
Tscherry, V. & Laves, F. (1970). *Naturw.* **57**, 1964.
Tscherry, V. & Schulz, H. (1970). *Naturw.* **57**, 194.
Winkler, H. G. F. (1948). *Acta Cryst.* **1**, 27–34.