GIACOVAZZO, C. (1976). *Acta Cryst.* A**32**, 74–82.
GRÖBNER, W. & HOFREITER, N. (1966). *Integraltafel* 2. *Bestimmte Integrale.* New York: Springer.
HAUPTMANN, H. & KARLE, J. (1953). *Acta Cryst.* **6**, 136–141.
HOWELLS, E. R., PHILLIPS, D. C. & ROGERS, D. (1950). *Acta Cryst.* **3**, 210–214.
KARLE, J. & GILARDI, R. D. (1973). *Acta Cryst.* A**29**, 401–407.
KARLE, J. & HAUPTMANN, H. (1953). *Acta Cryst.* **6**, 131–135.
KLUG, A. (1958). *Acta Cryst.* **8**, 515–543.
LIPSON, H. & WOOLFSON, M. M. (1952). *Acta Cryst.* **5**, 680–682.
MAIN, P., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J. P. & WOOLFSON, M. M. (1978). *A System of Computer Programs for Solution of Crystal Structures by Direct Methods.* Univ. of York, England.
MITRA, G. B. & BELGAUMKAR, J. (1973). *Proc. Indian Natl Sci. Acad. Part A*, **39**, 95–100.
NAYA, S., NITTA, I. & ODA, T. (1964). *Acta Cryst.* **17**, 421–433.
PARTHASARATHY, S. (1966). *Z. Kristallogr.* **123**, 27–50.
ROGERS, D. & WILSON, A. J. C. (1953). *Acta Cryst.* **6**, 439–449.
SHMUELI, U. (1979). *Acta Cryst.* A**35**, 282–286.
SHMUELI, U. & KALDOR, U. (1981). *Acta Cryst.* A**37**, 76–80.
SRINIVASAN, R. & PARTHASARATHY, S. (1976). *Some Statistical Applications in X-ray Crystallography.* Oxford: Pergamon Press.
WILSON, A. J. C. (1949). **2**, 318–321.
WILSON, A. J. C. (1952). *Research (London)*, **5**, 589–590.
WILSON, A. J. C. (1956). *Acta Cryst.* **9**, 143–144.
WILSON, A. J. C. (1978). *Acta Cryst.* A**34**, 986–994.
WILSON, A. J. C. (1980). *Acta Cryst.* A**36**, 945–946.

Acta Cryst. (1981). A**37**, 353–360

The Behaviour of Residuals with Respect to the Elimination of Low-Intensity Data

BY G. H. PETIT, A. T. H. LENSTRA AND J. F. VAN LOOCK

Department of Chemistry, University of Antwerp (UIA), Universiteitsplein 1, B-2610 Wilrijk, Belgium

(*Received* 14 *July* 1979; *accepted* 19 *November* 1980)

Abstract

Theoretical expressions for R_2 and R_2^N, indicator functions used in automated structure analysis, are given for space groups $P1$ and $P\bar{1}$ as a function of $\sigma_1^2 = \sum_{j=1}^P f_j^2/\sum_{j=1}^N f_j^2$ (P is the number of atoms in the tentative structural model) and of a threshold a on the intensity data. It appears that R_2 is only slightly dependent, R_2^N, however, is strongly dependent on a. If low-intensity reflections are eliminated from the data set, R_2, and not R_2^N, thus remains a useful control function in testing the reliability of atomic positions Theory and experiment are shown to be in good agreement.

1. Introduction

In automated structure analysis, one obviously needs criteria to discriminate between correct and incorrect structure models. As such a criterion we prefer the mathematical indicator function R_2 above chemical criteria. R_2 is defined as

$$R_2 = \frac{\sum\limits_H (I_N - I_P)^2}{\sum\limits_H I_N^2}, \tag{1}$$

where I_N represents the observed intensity of the N-atom structure looked for and I_P is the calculated intensity of a proposed structure model containing P atoms.

The functional behaviour of R_2 is evaluated theoretically in many papers (Wilson, 1969, 1974, 1976; Lenstra, 1973, 1974, 1975, 1979; Parthasarathy & Parthasarathi, 1972; Parthasarathy, 1975; Van de Mieroop & Lenstra, 1978). Applying R_2 in realistic structure determinations, Van de Mieroop (1979) obtained encouraging results. In the procedure he tried, the structure is determined by adding the atoms one by one to the model. Each time the reliability of the enlarged model is checked, comparing its experimental R_2 value with the expected one. Clearly, the reliability check requires substantial amounts of computer time.

An acceleration of the reliability check was sought in a reduction of the numbers of reflections used to calculate R_2. Following the practice with rotation functions (Tollin & Rossmann, 1966), we decided to eliminate intensities for which $E^2 < a$ and to investigate the consequences on the path of R_2.

In this paper we report the theoretical derivation of R_2 as a function of the threshold a for the space groups $P1$ and $P\bar{1}$ and compare the results with corresponding experimental values. Parthasarathi & Parthasarathy (1975) have stated that the normalized residual R_2^N

0567-7394/81/030353-08$01.00

might be a better indicator function for small structure models than R_2.

R_2^N is defined as

$$R_2^N = \frac{\sum_H (I_N - I_P/\sigma_1^2)^2}{\sum_H I_N^2} \quad (2)$$

with

$$\sigma_1^2 = \sum_{j=1}^{P} f_j^2 \Big/ \sum_{j=1}^{N} f_j^2.$$

A comparison of R_2 and R_2^N is given.

2. The residual R_2 as a function of the threshold a

With the notation of Srinivasan & Parthasarathy (1976), (1) can also be written in terms of the normalized structure factor variables y_N and y_P as

$$R_2 = \frac{\langle (y_N^2 - y_P^2\,\sigma_1^2)^2\rangle}{\langle y_N^4\rangle}. \quad (3)$$

On simplification (3) yields

$$R_2 = \frac{\langle y_N^4\rangle + \sigma_1^4\langle y_P^4\rangle - 2\sigma_1^2\langle y_N^2 y_P^2\rangle}{\langle y_N^4\rangle}. \quad (4)$$

From a practical point of view, it is obvious that the threshold a is to be applied to the observed intensities only. For a model of the related type, *i.e.* for a correct structure model, y_N and y_P are dependent variables. Then the moments of y_P are a function of the threshold. So for the related case we find

$$R_2(a) = \frac{\langle y_N^4\rangle_a + \sigma_1^4\langle y_P^4\rangle_a - 2\sigma_1^2\langle y_N^2 y_P^2\rangle_a}{\langle y_N^4\rangle_a}. \quad (5)$$

When the structure model is completely wrong (*i.e.* the unrelated case), y_N and y_P are mutually independent. Therefore,

$$\langle y_N^2 y_P^2\rangle_a = \langle y_N^2\rangle_a \langle y_P^2\rangle = \langle y_N^2\rangle_a \quad (6)$$

because

$$\langle y_P^2\rangle = 1.$$

R_2 for the unrelated case is then given by

$$R_2(a) = \frac{\langle y_N^4\rangle_a + \sigma_1^4\langle y_P^4\rangle - 2\sigma_1^2\langle y_N^2\rangle_a}{\langle y_N^4\rangle_a}. \quad (7)$$

Equations (5) and (7) for the related and unrelated cases, respectively, are valid for any space group. The moments of y_N and y_P can be obtained by using the joint probability distribution function $P(y_N, y_P)$. This distribution, however, is space-group dependent. For brevity we confine ourselves to the space groups $P1$ and $P\bar{1}$. Let the structure contain N equal atoms, while the tentative structure model contains P atoms at their exact positions. The joint probability function is then given by

$$P(y_N, y_P) = \frac{4y_N y_P}{\sigma_2^2} \exp - \left[\frac{y_N^2 + y_P^2}{\sigma_2^2}\right] I_0 \left(\frac{2\sigma_1 y_N y_P}{\sigma_2^2}\right) \quad (8)$$

for a non-centrosymmetric crystal, and by

$$P(y_N, y_P) = \frac{2}{\pi\sigma_2} \exp - \left[\frac{y_N^2 + y_P^2}{2\sigma_2^2}\right] \cosh \left(\frac{\sigma_1 y_N y_P}{\sigma_2^2}\right) \quad (9)$$

for a centrosymmetric crystal (Srinivasan & Parthasarathy, 1976). Here I_0 stands for the modified Bessel function of the first kind of order zero and $\sigma_2^2 = 1 - \sigma_1^2$. In order to evaluate $R_2(a)$, one has to calculate

$$\langle y_N^n y_P^m\rangle_a = \frac{\int_{\sqrt{a}}^{\infty}\int_0^{\infty} y_N^n y_P^m\, P(y_N, y_P)\, \mathrm{d}y_P\, \mathrm{d}y_N}{\int_{\sqrt{a}}^{\infty}\int_0^{\infty} P(y_N, y_P)\, \mathrm{d}y_P\, \mathrm{d}y_N} \quad (10)$$

for $n = 0, 2, 4$ and $m = 4, 2, 0$.

The moments $\langle y_N^2\rangle_a$, $\langle y_N^4\rangle_a$ and the normalization factor can also be evaluated more simply using the marginal probability function $P(y_N)$ for the non-centrosymmetric and centrosymmetric cases. These moments can thus be obtained by the integration

$$\langle y_N^n\rangle_a = \int_{\sqrt{a}}^{\infty} y_N^n\, P(y_N)\, \mathrm{d}y_N \quad (11)$$

for $n = 0, 2$ and 4.

3. $R_2(a)$ for $P1$ and $P\bar{1}$

The expectation values of y_N^2, y_N^4, y_P^4 and $y_P^2 y_N^2$ as a function of the threshold a are calculated in Appendix A for $P1$ and in Appendix B for $P\bar{1}$. We then obtain

(i) for the related case in $P1$:

$$R_2(a) = \{a^2[\sigma_1^8 - 2\sigma_1^4 + 1] + 2a[-\sigma_1^8 + 2\sigma_1^6 - \sigma_1^4 - \sigma_1^2 + 1] + 2(1 - \sigma_1^2)\}(a^2 + 2a + 2)^{-1}; \quad (12)$$

(ii) for the unrelated case in $P1$:

$$R_2(a) = [a^2 + 2a(1 - \sigma_1^2) + 2(\sigma_1^4 - \sigma_1^2 + 1)] \times (a^2 + 2a + 2)^{-1}; \quad (13)$$

(iii) for the related case in $P1$:

$$R_2(a) = 1 + \left\{ -\sigma_1^2(2+\sigma_1^2)\,\mathrm{erfc}\sqrt{\frac{a}{2}} + \sigma_1^2[-2-4\sigma_1^2+6\sigma_1^4-3\sigma_1^6+\sigma_1^2(\sigma_1^4-2)a] \times \sqrt{\frac{2a}{\pi}}\,e^{-a/2} \right\} \times \left[3\,\mathrm{erfc}\sqrt{\frac{a}{2}} + (3+a)\sqrt{\frac{2a}{\pi}}\,e^{-a/2}\right]^{-1}; \quad (14)$$

(iv) for the unrelated case in $P\bar{1}$:

$$R_2(a) = 1 + \left[-2\sigma_1^2\,\mathrm{erfc}\sqrt{\frac{a}{2}} - 2\sigma_1^2\sqrt{\frac{2a}{\pi}}\,e^{-a/2} + 3\sigma_1^4\,\mathrm{erfc}\sqrt{\frac{a}{2}}\right] \times \left[3\,\mathrm{erfc}\sqrt{\frac{a}{2}}\,(3+a)\sqrt{\frac{2a}{\pi}}\,e^{-a/2}\right]^{-1}. \quad (15)$$

Graphs of the functions are shown in Figs. 1, 2, 3 and 4 respectively.

Except for the related cases, the behaviour of the discrepancy index $R_2(a)$ is strongly influenced by the elimination of low-intensity data from the calculations. The behaviour of $R_2(a)$ is characterized by two different regions, as can be seen in Figs. 1 and 3. For the non-centrosymmetric case, $R_2(a)$ decreases in a monotonic way in the region $\sigma_1^2 > 0{\cdot}7$, whereas for $\sigma_1^2 < 0{\cdot}7$ the residual contains an additional minimum with respect to the threshold. This minimum moves to the right as σ_1^2 increases and is located at $a = \infty$ when σ_1^2 reaches the value 0·7. Similar considerations can be made for the centrosymmetric case, namely for $\sigma_1^2 > 0{\cdot}5$, $R_2(a)$ decreases monotonically and for $\sigma_1^2 < 0{\cdot}5$ a minimum on the R_2 path occurs.

For the unrelated cases, however, the influence on R_2 of the threshold is more outspoken as is seen in Figs. 2 and 4. The residual shows always a minimum and, for $a \to \infty$, the value of R_2 reaches the limiting value 1 for all σ_1^2 values.

4. Experimental verification

(a) The non-centrosymmetric case

The theory was checked for space group $P1$ with the experimental intensity data and the refined structural parameters of ammonium hydrogen 1-malate, (R = 0·024; Versichel, Van de Mieroop & Lenstra, 1978). The compound, with empirical formula $C_4H_9NO_5$, was regarded as an equal-atom structure. Moreover, it was shown to be a good test example for space group $P1$ (Van de Mieroop & Lenstra, 1978).

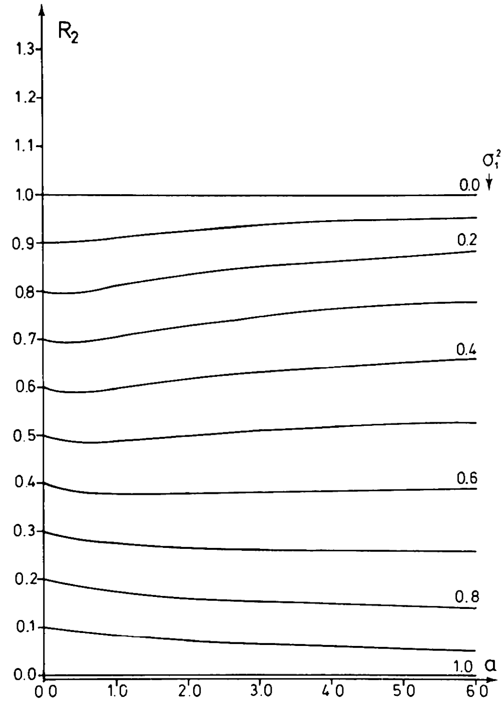

Fig. 1. R_2 for the related case in space group $P1$.

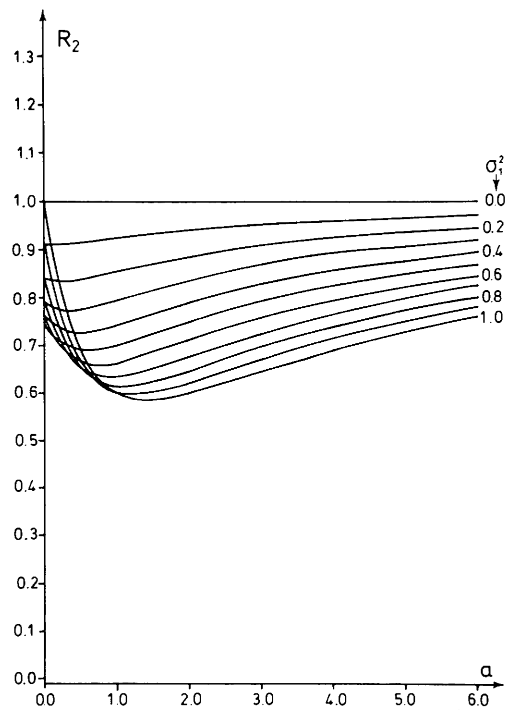

Fig. 2. R_2 for the unrelated case in space group $P1$.

R_2 values are enumerated for the related case only, using the observed intensities instead of the normalized ones. The good agreement between theory and experiment is shown in Table 1, in which some experimental R_2 values and the corresponding theoretical ones are given as a function of the threshold a.

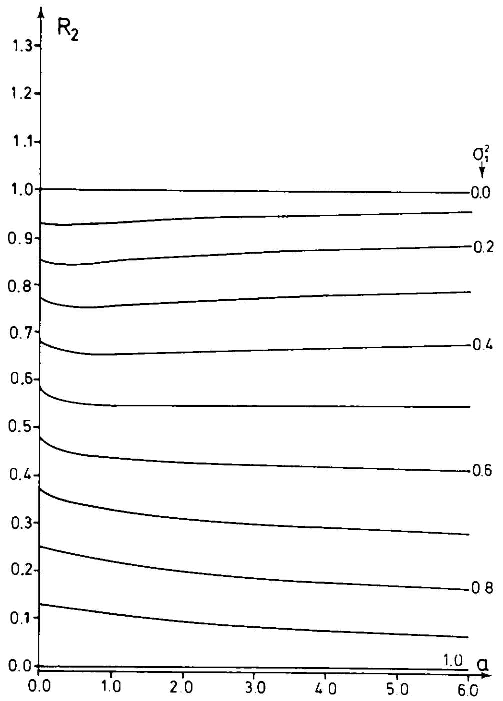

Fig. 3. R_2 for the related case in space group $P\bar{1}$.

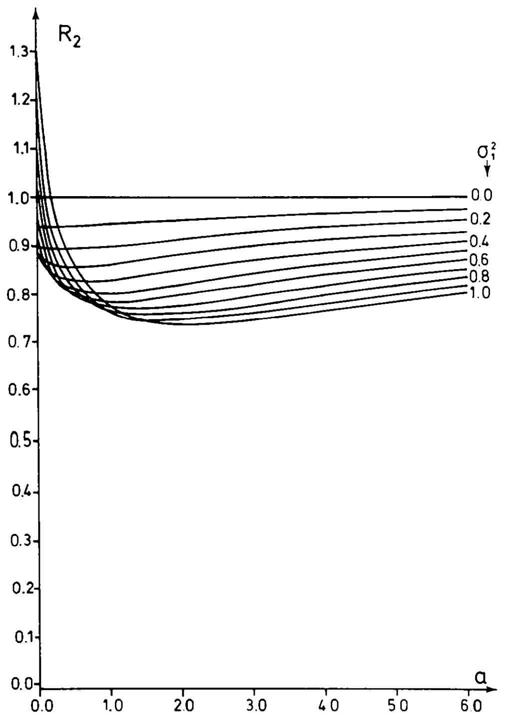

Fig. 4. R_2 for the unrelated case in space group $P\bar{1}$.

(*b*) *The centrosymmetric case*

An artificial equal-atom structure, produced from the structure of *cis-cis*-4,6-dimethyl-trimethylene sulphite (Petit, Lenstra & Geise, 1978) by replacing all carbon and sulphur atoms by oxygen, was used as test structure for $P\bar{1}$. The compound with empirical formula $C_5H_{10}O_3S$ crystallizes in space group $P2_1/c$ and has $Z=4$.

The residual values in Table 2 are 'averaged' ones, *i.e.* the average was calculated by taking each individual structure model with say $\sigma_1^2=0{\cdot}56$ (equivalent to five independent atoms out of the total of nine atomic positions possible) and then averaging over all possible permutational structure models.

Table 2 shows an even better agreement between theory and experiment for the centrosymmetric structure than was obtained in the non-centrosymmetric structure (Table 1). In the latter case no averaging procedure was applied. Since we deal then with one single structure model the test might be biased.

For both $P1$ and $P\bar{1}$ the agreement is sufficient to state that R_2 in the presence of a threshold remains a proper indicator function. Moreover, the computing time required to check the reliability of a tentative structure model can be reduced to about 20% of the original time by eliminating all E^2 values below $a=2$.

The usefulness of R_2 as an indicator function in structure evaluation procedures is obviously not only

Table 1. *Comparison of experimental R_2 values with the theoretical ones as a function of the threshold a for a non-centrosymmetric structure*

σ_1^2 measures the size of the known, correct structure fragment. Here $\sigma_1^2=0{\cdot}9$ corresponds to a model with 9 out of 10 atoms in the molecule. The asterisk indicates a local minimum in R_2 as a function of a.

	$\sigma_1^2=0{\cdot}9$		$\sigma_1^2=0{\cdot}5$	
a	100 × R_2 (exp)	100 × R_2 (theor)	100 × R_2 (exp)	100 × R_2 (theor)
0·00	9·44	10·00	54·10	50·00
0·20	9·26	9·76	52·91	49·08
0·40	9·11	9·44	52·80	48·65
0·60	8·95	9·08	52·38*	48·53*
0·80	8·94	8·73	52·62	48·58
1·00	8·74	8·40	52·78	48·75
1·20	8·43	8·09	53·48	48·97
1·40	7·49	7·81	54·31	49·22
1·60	7·45	7·56	54·43	49·48
1·80	7·44	7·33	54·65	49·75
2·00	7·30	7·12	54·93	50·00
3·00	6·10	6·33	52·92	51·10
4·00	5·66	5·82	52·00	51·92
5·00	5·56	5·46	52·34	52·53
6·00	5·81	5·20	55·00	53·00

Table 2. *Comparison of experimental R_2 values with the theoretical ones as a function of the threshold a for a centrosymmetric structure*

σ_1^2 measures the size of the known, correct structure fragment, *e.g.* $\sigma_1^2 = 0{\cdot}11$ corresponds here to a model with 1 out of the 9 atoms in the molecule. NO represents the number of reflections taken in the calculation of R_2. Standard deviations of the 'experimental' R_2 values for $\theta_{max} = 30°$ are shown in parentheses.

		$\sigma_1^2 = 0{\cdot}11$		$\sigma_1^2 = 0{\cdot}56$	
a	NO	100 × R_2 (exp)	100 × R_2 (theor)	100 × R_2 (exp)	100 × R_2 (theor)
0·0	1561	93·92 (1·25)	92·18	50·86 (9·25)	52·68
0·2	957	93·83 (1·24)*	91·98*	49·45 (9·14)	50·60
0·4	753	93·87 (1·24)	92·07	49·30 (9·12)	49·87
0·6	621	93·98 (1·24)	92·23	49·08 (9·24)	49·40
0·8	532	94·20 (1·21)	92·40	48·83 (9·25)	49·05
1·0	455	94·86 (1·32)	92·59	48·26 (9·30)	48·79
1·2	387	95·00 (1·34)	92·78	48·16 (9·42)	48·59
1·4	344	95·02 (1·35)	92·96	48·15 (9·45)	48·43
1·6	308	95·65 (1·42)	93·13	47·86 (10·02)	48·30
1·8	264	95·71 (1·47)	93·30	47·84 (10·19)	48·19
2·0	240	95·83 (1·45)	93·45	47·78 (10·34)	48·10
3·0	149	96·30 (1·56)	94·12	47·70 (11·34)	47·84
4·0	86	96·61 (1·77)	94·61	47·68 (12·92)	47·72
5·0	58	96·64 (1·99)	95·00	47·68 (14·21)	47·66

determined by the predictability of the residual value itself. Another important factor is the standard deviation on R_2. Unfortunately the present theory does not allow us to predict $\sigma(R_2)$ because our formulae are derived from distribution functions which are, strictly speaking, only valid if the number of observations is infinite. We will tackle this problem in a forthcoming paper (Van Havere & Lenstra, 1981). Our averaging procedure gave us, however, some empirical and most useful information on the behaviour of $\sigma(R_2)$ as a function of both the size of the proposed correct structure model and the threshold which is depicted in Fig. 5. Surprisingly the introduction of small values for the threshold does not influence $\sigma(R_2)$ much. At $a = 1$ hardly any increase in $\sigma(R_2)$ is noted even though approximately 70% of the observations involved in the R_2 calculations are ignored.

5. $R_2^N(a)$ for $P1$ and $P\bar{1}$

Results for the normalized index $R_2^N(a)$ can be obtained by following exactly the same procedure as before. We obtain

(i) for the related case in $P1$:

$$R_2^N(a) = (1-\sigma_1^2)[(1-\sigma_1^2)a^2 + 2\sigma_1^2 a + 2] \times [a^2 + 2a + 2]^{-1}; \quad (16)$$

(ii) for the unrelated case in $P1$:

$$R_2^N(a) = (a^2+2)(a^2+2a+2)^{-1}; \quad (17)$$

(iii) for the related case in $P\bar{1}$:

$$R_2^N(a) = \left\{4(1-\sigma_1^2)\operatorname{erfc}\sqrt{\frac{a}{2}} + \left[1 + (1-\sigma_1^2)^2 a + 2\sigma_1^2 - 3\sigma_1^4\right]\sqrt{\frac{2a}{\pi}}\, e^{-a/2}\right\} \times \left[3\operatorname{erfc}\sqrt{\frac{a}{2}} + (3+a)\sqrt{\frac{2a}{\pi}}\, e^{-a/2}\right]^{-1}; \quad (18)$$

(iv) for the unrelated case in $P\bar{1}$:

$$R_2^N(a) = \left[4\operatorname{erfc}\sqrt{\frac{a}{2}} + (1+a)\sqrt{\frac{2a}{\pi}}\, e^{-a/2}\right] \times \left[3\operatorname{erfc}\sqrt{\frac{a}{2}} + (3+a)\sqrt{\frac{2a}{\pi}}\, e^{-a/2}\right]^{-1}. \quad (19)$$

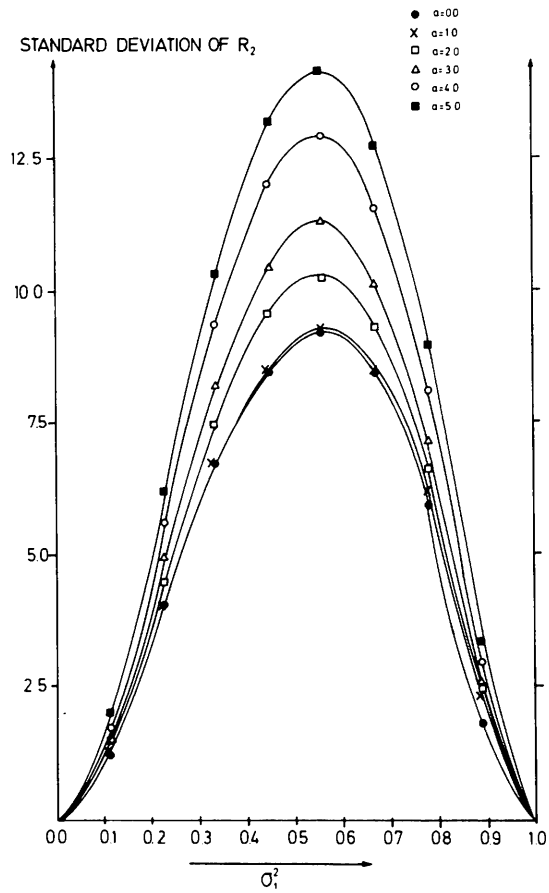

Fig. 5. Variation of $\sigma(R_2)$ with the size of the proposed correct structure model and the threshold. ● $a = 0{\cdot}0$, × $a = 1{\cdot}0$, □ $a = 2{\cdot}0$, △ $a = 3{\cdot}0$, ○ $a = 4{\cdot}0$, ■ $a = 5{\cdot}0$.

Graphs of these functions are shown in Figs. 6, 7, 8 and 9 respectively.

We note that the normalized index in both unrelated cases is independent of the size of the model σ_1^2. On the other hand, there is a quite strong dependence of R_2^N upon the threshold a. Particularly at low values of a the normalized index decreases very rapidly. This feature is found for the related and for the unrelated cases in any space group.

A comparison of R_2 and R_2^N is the most interesting with respect to their behaviour in the related case, *i.e.* when the tentative structure model is correct. Starting from a zero-atom model R_2^N decreases faster than R_2 for relatively small models. Because of this aspect Parthasarathi & Parthasarathy (1975) prefer, for small structure models, R_2^N above R_2, thereby overlooking two disadvantages. Firstly, for a zero-atom model $R_2 = 1{\cdot}00$ under all conditions. The starting value of R_2^N, however, depends on the actual intensity distribution of the structure looked for. In the case of an 'ideal' $P1$ structure $R_2^N = 1{\cdot}00$, whereas for an 'ideal' $P\bar{1}$ structure $R_2^N = 1{\cdot}33$. In practical circumstances R_2^N will have some unknown, inbetween starting value. Secondly, owing to a non-ideal signal-to-noise ratio the measured intensity for a number of reflections will turn out to be zero or negative. Usually these reflections are omitted from the analysis or given artificially an intensity of zero. Such an action is equivalent to the introduction of a small, but unknown, threshold in the data set. Its influence on R_2 (see Figs. 1 and 3) can be ignored, but the impact on R_2^N (Figs. 6 and 8) for small structure models is substantial.

On the other hand, one should not over-rate the importance of these drawbacks. For instance, a

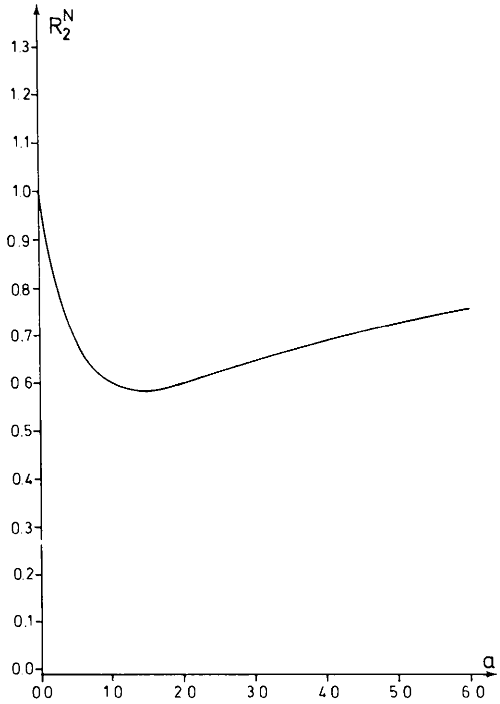

Fig. 7. R_2^N for the unrelated case in space group $P1$.

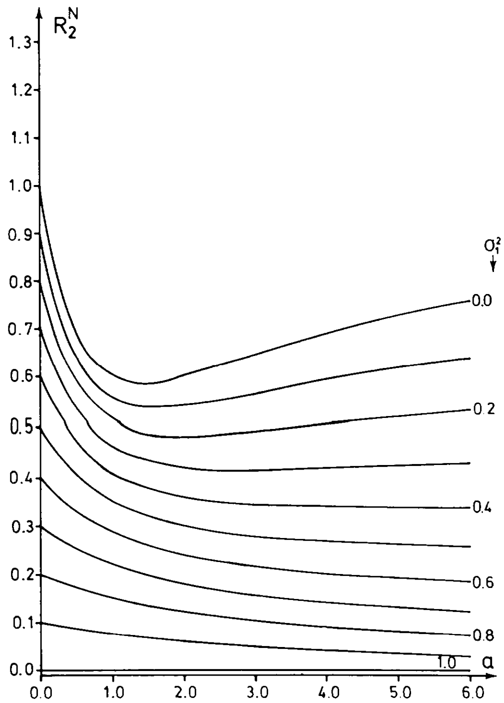

Fig. 6. R_2^N for the related case in space group $P1$.

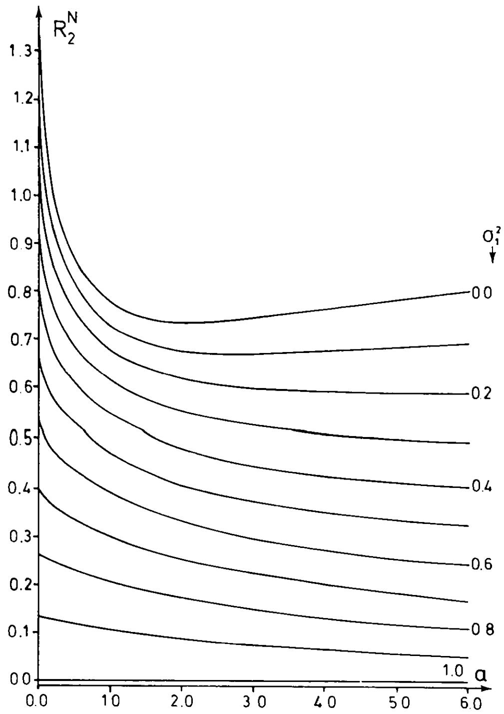

Fig. 8. R_2^N for the related case in space group $P\bar{1}$.

treatment of the primary intensity data in the way proposed by French & Wilson (1978) could be reliable enough to eliminate the enforced signal-to-noise threshold. Also, a sufficiently large value of a restores the usefulness of R_2^N and makes the preferences between R_2 and R_2^N a matter of personal taste. In fact, in a forthcoming paper (Van Havere & Lenstra, 1981) it will be proved that the discriminating power of R_2 and R_2^N as a function of a is exactly the same.

It should be stated clearly that one is only allowed to eliminate low intensities from the data set when checking the correctness of an atom newly added to the model. Only then is one allowed to take advantage of the reduction in computing time. Structure refinement, scaling procedures, *etc.* should be carried out with all the observed intensities as is recommended by Hirshfeld & Rabinovich (1973) and Wilson (1978). Naturally, any systematic exclusion of weak reflections in, for example, a least-squares procedure will cause asymmetry in the error distribution of the observed data and will inevitably introduce systematic errors in the parameters.

One of us, GHP, wants to thank the Belgian organization Instituut voor Wetenschappelijk Onderzoek in Nijverheid en Landbouw, IWONL, for a predoctoral grant. We also thank Drs H. Krabbendam and Professor Dr H. J. Geise for their critical reading of the manuscript.

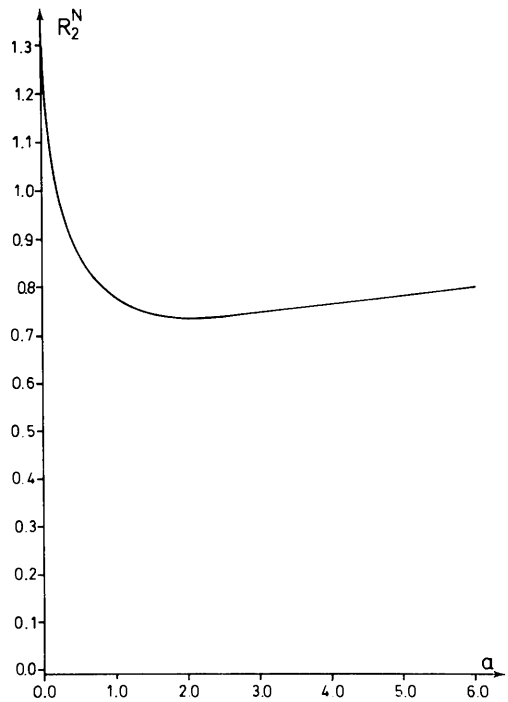

Fig. 9. R_2^N for the unrelated case in space group $P1$.

APPENDIX *A*

In this section, we describe the procedure to calculate the moments given by equation (10), with the joint probability distribution function $P(y_N, y_P)$ for the non-centrosymmetric case. The normatlization integral in the denominator of (10) can also be written as

$$K_a = \int_{\sqrt{a}}^{\infty} P(y_N)\,dy_N. \qquad (A1)$$

So we find that $K_a = e^{-a}$. $\qquad (A2)$

Next, we work out the integral

$$\int_{\sqrt{a}}^{\infty}\int_{0}^{\infty} y_N^2 y_P^2 P(y_N, y_P)\,dy_P\,dy_N. \qquad (A3)$$

Substituting equation (8) in (*A*3) and using the series expansion of the modified Bessel function of the first kind (Abramowitz & Stegun, 1968), given by

$$I_0\left(\frac{2\sigma_1 y_N y_P}{\sigma_2^2}\right) = \sum_{k=0}^{\infty} \frac{\sigma_1^{2k} y_N^{2k} y_P^{2k}}{\sigma_2^{4k}(k!)^2}, \qquad (A4)$$

we find that the integral is given by

$$\frac{4}{\sigma_2^2}\sum_{k=0}^{\infty}\frac{\sigma_1^{2k}}{\sigma_2^{4k}(k!)^2}\int_0^{\infty} y_P^{2k+3}\exp\left[-\frac{y_P^2}{\sigma_2^2}\right]$$

$$\times \int_{\sqrt{a}}^{\infty} y_N^{2k+3}\exp\left[-\frac{y_N^2}{\sigma_2^2}\right] dy_N\,dy_P. \qquad (A5)$$

Writing this expression in terms of the normalized variables $z_N(=y_N^2)$ and $z_P(=y_P^2)$ and using the solutions of the standard integrals

$$\int_a^{\infty} x^m e^{-nx}\,dx = e^{-na}\sum_{r=0}^{m}\frac{m!\,a^{m-r}}{(m-r)!\,n^{r+1}}, \qquad (A6)$$

$$\int_0^{\infty} x^n e^{-bx}\,dx = \frac{n!}{b^{n+1}} \qquad (A7)$$

for $b > 0$ and n positive, we find a double series expansion, which can be shown to be equal to

$$e^{-a}[\sigma_1^2(a^2 + a + 1) + 1 + a]. \qquad (A8)$$

With this technique, we find for the other moments the following expressions

$$\langle y_N^2\rangle_a = 1 + a \qquad (A9)$$

$$\langle y_N^4\rangle_a = a^2 + 2a + 2 \qquad (A10)$$

$$\langle y_P^2\rangle_a = 1 + \sigma_1^2 a \qquad (A11)$$

$$\langle y_P^4\rangle_a = \sigma_1^4 a^2 + 2\sigma_1^2(2-\sigma_1^2)a + 2 \qquad (A12)$$

$$\langle y_N^2 y_P^2\rangle_a = \sigma_1^2 a^2 + (1+\sigma_1^2)a + 1 + \sigma_1^2. \qquad (A13)$$

APPENDIX *B*

The moments presented by equation (10) will be calculated for the centrosymmetric case using the joint probability distribution function (9). The procedure for calculating the moments is somewhat different to the one treated in Appendix *A*.

The normalization integral K_a of (10) is given by

$$K_a = \operatorname{erfc}\sqrt{a/2}. \qquad (B1)$$

Let us work out the integral

$$\int_{\sqrt{a}}^{\infty}\int_{0}^{\infty} y_N^2 y_P^2 P(y_N, y_P)\, dy_P\, dy_N. \qquad (B2)$$

Substituting (9) in ($B2$) and using the definition

$$\cosh x = (e^x + e^{-x})/2, \qquad (B3)$$

we find the integral becomes

$$\frac{\sigma_2^5}{\pi}\int_{\sqrt{a}/\sigma_2}^{\infty} u^2 e^{-u^2/2}\int_0^{\infty} v^2 e^{-v^2/2}\,(e^{\sigma_r uv} + e^{-\sigma_r uv})\,dv\,du \qquad (B4)$$

in which $u = y_N/\sigma_2$ and $v = y_P/\sigma_2$.

By applying the following standard integrals

$$\int_0^{\infty} e^{-(at^2+bt+c)}\,dt = \frac{1}{2}\sqrt{\frac{\pi}{a}}\,e^{(b^2-ac)/a}\operatorname{erfc}\frac{b}{\sqrt{a}} \qquad (B5)$$

and

$$i^n \operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^{\infty}\frac{(t-z)^n}{n!}e^{-t^2}\,dt, \qquad (B6)$$

the numerator of (10) can be written as

$$(1 + 2\sigma_1^2)\operatorname{erfc}\sqrt{\frac{a}{2}} + (1 + 2\sigma_1^2 + a\sigma_1^2)\sqrt{\frac{2a}{\pi}}\,e^{-a/2}, \qquad (B7)$$

so for the other moments we find

$$\langle y_N^2\rangle_a = 1 + Q \qquad (B8)$$

$$\langle y_N^4\rangle_a = 3 + (3 + a)Q \qquad (B9)$$

$$\langle y_P^2\rangle_a = 1 + \sigma_1^2 Q \qquad (B10)$$

$$\langle y_P^4\rangle_a = 3 + \sigma_1^2[6 + (a-3)\sigma_1^2]Q \qquad (B11)$$

$$\langle y_N^2 y_P^2\rangle_a = (1 + 2\sigma_1^2) + [1 + (2 + a)\sigma_1^2]Q \qquad (B12)$$

with

$$Q = \sqrt{\frac{2a}{\pi}}\,e^{-a/2}/\operatorname{erfc}\sqrt{\frac{a}{2}}. \qquad (B13)$$

References

ABRAMOWITZ, M. & STEGUN, I. A. (1968). *Handbook of Mathematical Functions*. New York: Dover Publications.

FRENCH, S. & WILSON, K. (1978). *Acta Cryst.* A**34**, 517–525.

HIRSHFELD, F. L. & RABINOVICH, D. (1973). *Acta Cryst.* A**29**, 510–513.

LENSTRA, A. T. H. (1973). *Acta Cryst.* A**29**, 419–423.

LENSTRA, A. T. H. (1974). *Acta Cryst.* A**30**, 363–369.

LENSTRA, A. T. H. (1975). *Acta Cryst.* A**31**, S12.

LENSTRA. A. T. H. (1979). *Bull. Soc. Chim. Belg.* **88**, 359–368.

PARTHASARATHI, V. & PARTHASARATHY, S. (1975). *Acta Cryst.* A**31**, 38–41.

PARATHASARATHY, S. (1975). *Acta Cryst.* A**31**, 375–380.

PARTHASARATHY, S. & PARTHASARATHI, V. (1972). *Acta Cryst.* A**28**, 426–432.

PETIT, G. H., LENSTRA, A. T. H. & GEISE, H. J. (1978). *Bull. Soc. Chim. Belg.* **87**, 659–666.

SRINIVASAN, R. & PARTHASARATHY, S. (1976). *Some Statistical Applications in X-ray Crystallography*. Oxford: Pergamon Press.

TOLLIN, P. & ROSSMANN, M. G. (1966). *Acta Cryst.* **21**, 872–876.

VAN DE MIEROOP, W. (1979). PhD Thesis (in Dutch), Univ. of Antwerp, Belgium.

VAN DE MIEROOP, W. & LENSTRA, A. T. H. (1978). *Acta Cryst.* A**34**, 860–863.

VAN HAVERE, W. & LENSTRA, A. T. H. (1981). In preparation.

VERSICHEL, W., VAN DE MIEROOP, W. & LENSTRA, A. T. H. (1978). *Acta Cryst.* B**34**, 2643–2645.

WILSON, A. J. C. (1969). *Acta Cryst.* B**25**, 1288–1293.

WILSON, A. J. C. (1974). *Acta Cryst.* A**30**, 836–838.

WILSON, A. J. C. (1976). *Acta Cryst.* A**32**, 53–56.

WILSON, A. J. C. (1978). *Acta Cryst.* A**34**, 474–475.